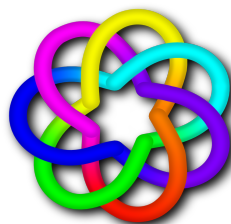


CONFERENCES IN RESEARCH AND PRACTICE IN
INFORMATION TECHNOLOGY

VOLUME 129

HEALTH INFORMATICS AND KNOWLEDGE MANAGEMENT 2012

AUSTRALIAN COMPUTER SCIENCE COMMUNICATIONS, VOLUME 34, NUMBER 7



HEALTH INFORMATICS AND KNOWLEDGE MANAGEMENT 2012

Proceedings of the Fifth Australasian Workshop on
Health Informatics and Knowledge Management
(HIKM 2012), Melbourne, Australia,
31 January – 3 February 2012

Kerryn Butler-Henderson and Kathleen Gray, Eds.

Volume 129 in the Conferences in Research and Practice in Information Technology Series.
Published by the Australian Computer Society Inc.



Published in association with the ACM Digital Library.

Health Informatics and Knowledge Management 2012. Proceedings of the Fifth Australasian Workshop on Health Informatics and Knowledge Management (HIKM 2012), Melbourne, Australia, 31 January – 3 February 2012

Conferences in Research and Practice in Information Technology, Volume 129.

Copyright ©2012, Australian Computer Society. Reproduction for academic, not-for-profit purposes permitted provided the copyright text at the foot of the first page of each paper is included.

Editors:

Kerryn Butler-Henderson

School of Public Health

Curtin University

Perth, WA 6845

Australia

Email: K.Butler-Henderson@exchange.curtin.edu.au

Kathleen Gray

Health and Biomedical Informatics Research Unit

Melbourne Medical School and Department of Information Systems

University of Melbourne

Melbourne, VIC 3010

Australia

Email: kgray@unimelb.edu.au

Series Editors:

Vladimir Estivill-Castro, Griffith University, Queensland

Simeon J. Simoff, University of Western Sydney, NSW

Email: crpit@scm.uws.edu.au

Publisher: Australian Computer Society Inc.

PO Box Q534, QVB Post Office

Sydney 1230

New South Wales

Australia.

Conferences in Research and Practice in Information Technology, Volume 129.

ISSN 1445-1336.

ISBN 978-1-921770-10-4.

Printed, January 2012 by University of Western Sydney, on-line proceedings

Printed, January 2012 by RMIT, electronic media

Document engineering by CRPIT

The *Conferences in Research and Practice in Information Technology* series disseminates the results of peer-reviewed research in all areas of Information Technology. Further details can be found at <http://crpit.com/>.

Table of Contents

Proceedings of the Fifth Australasian Workshop on Health Informatics and Knowledge Management (HIKM 2012), Melbourne, Australia, 31 January – 3 February 2012

Preface	vii
Acknowledgement	viii
Programme Committee	ix
Organising Committee	x
Welcome from the Organising Committee	xi
CORE - Computing Research & Education	xiii
ACSW Conferences and the Australian Computer Science Communications	xiv
ACSW and HIKM 2012 Sponsors	xvi

Contributed Papers

The Role of Emotional Intelligence on the Resolution of Disputes Involving the Electronic Health Record	3
<i>Emilia Bellucci, Sitalakshmi Venkatraman, Nial Muecke, Andrew Stranieri and Jemal Abawajy</i>	
Detection of Evidence in Clinical Research Papers	13
<i>Patrick Davis-Desmond and Diego Mollá</i>	
Evaluation of Web 2.0 Technologies for Developing Online Telehealth Systems	21
<i>Jaspaljeet Singh Dhillon, Czarina Ramos, Burkhard C. Wünsche and Christof Lutteroth</i>	
Achieving Acceptable Structured eReferral Forms	31
<i>Yulong Gu, Jim Warren, Karen Day, Malcolm Pollock and Sue White</i>	
A Bayesian Analysis on Historical Clinical Data Concerning Treatment Change for HIV/AIDS Patients	41
<i>Khairul A. Kasmiran, Albert Y. Zomaya, Ali Al Mazari and Roger J. Garsia</i>	
Inducing and Storing Generalised Evidences using Semantic Web Formalisms	49
<i>Razan Paul, Tudor Groza and Jane Hunter</i>	
Gaining Insight from Patient Journey Data using a Process-Oriented Analysis Approach	59
<i>Lua Perimal-Lewis, Shaowen Qin, Campbell Thompson and Paul Hakendorf</i>	
Effects of Physician Collaboration Network on Hospital Outcomes	67
<i>Shahadat Uddin and Liaquat Hossain</i>	
Author Index	75

Preface

We are pleased to present the papers from the Fifth Australasian Workshop on Health Informatics and Knowledge Management (HIKM), held on 1 February 2012 at RMIT University in Melbourne as part of Australasian Computer Science Week.

Research and development in Health Informatics and Knowledge Management has increasingly wider relevance to real-world settings because of the rise of e-health and the renewed pace of information technology-enabled healthcare system reforms of many kinds, not only in the Australasian region but around the world.

HIKM seeks to showcase scholarly and scientific investigation of this wave of change and innovation in healthcare, and to provide an annual forum for selective, in-depth reporting of new research and development in the field. HIKMs call for papers this year attracted 22 manuscripts that explored a fascinating array of current challenges and opportunities for Health Informatics and Knowledge Management, by researchers based in Australia and New Zealand and beyond.

Many of the manuscripts we reviewed demonstrated ideas and innovations that are still coming to fruition. Although we were unable to include them in HIKM 2012, we hope that the authors will persist with their work and will seek opportunities to present it in future – perhaps at HIKM 2013. Work in this field is an important part of the Australasian and international research endeavour. Application of the ensuing new knowledge should have a larger role that it presently does in efforts to improve and transform healthcare.

Following rigorous peer review by four independent expert reviewers, eight papers were chosen for inclusion in the proceedings of HIKM 2012 (a 36% acceptance rate). We extend our congratulations to the authors. These eight papers offer diverse approaches to and insights from research on a cross-section of key topics in Health Informatics and Knowledge Management: electronic health records, clinical research, telehealth, e-referral, clinical data, clinical decision-support, patient journey data and physician collaboration networks. They are published here as Volume 129 of *Conferences in Research and Practice in Information Technology*.

Kerryn Butler-Henderson
Curtin University

Kathleen Gray
University of Melbourne

HIKM 2012 Programme Chairs
January 2012

Acknowledgement

We wish to acknowledge all the authors who submitted manuscripts for review. Our special thanks to the Australasian Computer Science Week 2012 Organising Committee, the School of Public Health at Curtin University and the Health and Biomedical Informatics Research Unit at the University of Melbourne for supporting and sponsoring HIKM 2012. In particular we wish thank Professor Fernando Martin Sanchez, University of Melbourne, for giving the opening address to HIKM 2012 and Professor Sue Fyfe, Curtin University, for the Schools support. We wish to acknowledge in particular the work of Ms Jodi Burgess, Curtin University, in web hosting and providing ICT administrative support. We also wish to thank the reviewers of HIKM 2012 submissions for their time and their valuable feedback to authors; this was essential in ensuring that the proceedings of HIKM 2012 represent high-quality research outcomes.

Programme Committee

Chairs

Kerryn Butler-Henderson, Curtin University, Australia

Kathleen Gray, University of Melbourne, Australia

Members

Vicki Bennett, Australian Institute of Health and Welfare, Australia

Yulong (Helen) Gu, Auckland University, New Zealand

David Hansen, Commonwealth Scientific and Industrial Research Organisation (CSIRO), Australia

Anthony Maeder, University of Western Sydney, Australia

Tony Sahama, Queensland University of Technology, Australia

Louise Schaper, Health Informatics Society of Australia, Australia

Andrew Stranieri, University of Ballarat, Australia

Jim Warren, University of Auckland, New Zealand

Gerald Webster, Auckland University, New Zealand

Sallyanne Wissmann, Health Information Management Association of Australia, Australia

Dennis Wollersheim, La Trobe University, Australia

Organising Committee

Members

Dr. Daryl D'Souza
Assoc. Prof. James Harland (Chair)
Dr. Falk Scholer
Dr. John Thangarajah
Assoc. Prof. James Thom
Dr. Jenny Zhang

Welcome from the Organising Committee

On behalf of the Australasian Computer Science Week 2012 (ACSW2012) Organising Committee, we welcome you to this year's event hosted by RMIT University. RMIT is a global university of technology and design and Australia's largest tertiary institution. The University enjoys an international reputation for excellence in practical education and outcome-oriented research. RMIT is a leader in technology, design, global business, communication, global communities, health solutions and urban sustainable futures. RMIT was ranked in the top 100 universities in the world for engineering and technology in the 2011 QS World University Rankings. RMIT has three campuses in Melbourne, Australia, and two in Vietnam, and offers programs through partners in Singapore, Hong Kong, mainland China, Malaysia, India and Europe. The University's student population of 74,000 includes 30,000 international students, of whom more than 17,000 are taught offshore (almost 6,000 at RMIT Vietnam).

We welcome delegates from a number of different countries, including Australia, New Zealand, Austria, Canada, China, the Czech Republic, Denmark, Germany, Hong Kong, Japan, Luxembourg, Malaysia, South Korea, Sweden, the United Arab Emirates, the United Kingdom, and the United States of America.

We hope you will enjoy ACSW2012, and also to experience the city of Melbourne.,

Melbourne is amongst the world's most liveable cities for its safe and multicultural environment as well as well-developed infrastructure. Melbourne's skyline is a mix of cutting-edge designs and heritage architecture. The city is famous for its restaurants, fashion boutiques, café-filled laneways, bars, art galleries, and parks.

RMIT's city campus, the venue of ACSW2012, is right in the heart of the Melbourne CBD, and can be easily accessed by train or tram.

ACSW2012 consists of the following conferences:

- Australasian Computer Science Conference (ACSC) (Chaired by Mark Reynolds and Bruce Thomas)
- Australasian Database Conference (ADC) (Chaired by Rui Zhang and Yanchun Zhang)
- Australasian Computer Education Conference (ACE) (Chaired by Michael de Raadt and Angela Carbone)
- Australasian Information Security Conference (AISC) (Chaired by Josef Pieprzyk and Clark Thorburn)
- Australasian User Interface Conference (AUIC) (Chaired by Haifeng Shen and Ross Smith)
- Computing: Australasian Theory Symposium (CATS) (Chaired by Julián Mestre)
- Australasian Symposium on Parallel and Distributed Computing (AusPDC) (Chaired by Jinjun Chen and Rajiv Ranjan)
- Australasian Workshop on Health Informatics and Knowledge Management (HIKM) (Chaired by Keryn Butler-Henderson and Kathleen Gray)
- Asia-Pacific Conference on Conceptual Modelling (APCCM) (Chaired by Aditya Ghose and Flavio Ferrarotti)
- Australasian Computing Doctoral Consortium (ACDC) (Chaired by Falk Scholer and Helen Ashman)

ACSW is an event that requires a great deal of co-operation from a number of people, and this year has been no exception. We thank all who have worked for the success of ACSE 2012, including the Organising Committee, the Conference Chairs and Programme Committees, the RMIT School of Computer Science and IT, the RMIT Events Office, our sponsors, our keynote and invited speakers, and the attendees.

Special thanks go to Alex Potanin, the CORE Conference Coordinator, for his extensive expertise, knowledge and encouragement, and to organisers of previous ACSW meetings, who have provided us with a great deal of information and advice. We hope that ACSW2012 will be as successful as its predecessors.

Assoc. Prof. James Harland

School of Computer Science and Information Technology, RMIT University

ACSW2012 Chair

January, 2012

CORE - Computing Research & Education

CORE welcomes all delegates to ACSW2012 in Melbourne. CORE, the peak body representing academic computer science in Australia and New Zealand, is responsible for the annual ACSW series of meetings, which are a unique opportunity for our community to network and to discuss research and topics of mutual interest. The original component conferences - ACSC, ADC, and CATS, which formed the basis of ACSW in the mid 1990s - now share this week with seven other events - ACE, AISC, AUIC, AusPDC, HIKM, ACDC, and APCCM, which build on the diversity of the Australasian computing community.

In 2012, we have again chosen to feature a small number of keynote speakers from across the discipline: Michael Kölling (ACE), Timo Ropinski (ACSC), and Manish Parashar (AusPDC). I thank them for their contributions to ACSW2012. I also thank invited speakers in some of the individual conferences, and the two CORE award winners Warwish Irwin (CORE Teaching Award) and Daniel Frampton (CORE PhD Award). The efforts of the conference chairs and their program committees have led to strong programs in all the conferences, thanks very much for all your efforts. Thanks are particularly due to James Harland and his colleagues for organising what promises to be a strong event.

The past year has been very turbulent for our disciplines. We tried to convince the ARC that refereed conference publications should be included in ERA2012 in evaluations – it was partially successful. We ran a small pilot which demonstrated that conference citations behave similarly to but not exactly the same as journal citations - so the latter can not be scaled to estimate the former. So they moved all of Field of Research Code 08 “Information and Computing Sciences” to peer review for ERA2012. The effect of this will be that most Universities will be evaluated at least at the two digit 08 level, as refereed conference papers count towards the 50 threshold for evaluation. CORE’s position is to return 08 to a citation measured discipline as soon as possible.

ACSW will feature a joint CORE and ACDICT discussion on Research Challenges in ICT, which I hope will identify a national research agenda as well as priority application areas to which our disciplines can contribute, and perhaps opportunity to find international multi-disciplinary successes which could work in our region.

Beyond research issues, in 2012 CORE will also need to focus on education issues, including in Schools. The likelihood that the future will have less computers is small, yet where are the numbers of students we need?

CORE’s existence is due to the support of the member departments in Australia and New Zealand, and I thank them for their ongoing contributions, in commitment and in financial support. Finally, I am grateful to all those who gave their time to CORE in 2011; in particular, I thank Alex Potanin, Alan Fekete, Aditya Ghose, Justin Zobel, and those of you who contribute to the discussions on the CORE mailing lists. There are three main lists: csprofs, cshods and members. You are all eligible for the members list if your department is a member. Please do sign up via <http://lists.core.edu.au/mailman/listinfo> - we try to keep the volume low but relevance high in the mailing lists.

Tom Gedeon

President, CORE
January, 2012

ACSW Conferences and the Australian Computer Science Communications

The Australasian Computer Science Week of conferences has been running in some form continuously since 1978. This makes it one of the longest running conferences in computer science. The proceedings of the week have been published as the *Australian Computer Science Communications* since 1979 (with the 1978 proceedings often referred to as *Volume 0*). Thus the sequence number of the Australasian Computer Science Conference is always one greater than the volume of the Communications. Below is a list of the conferences, their locations and hosts.

2013. Volume 35. Host and Venue - University of South Australia, Adelaide, SA.

2012. Volume 34. Host and Venue - RMIT University, Melbourne, VIC.

2011. Volume 33. Host and Venue - Curtin University of Technology, Perth, WA.

2010. Volume 32. Host and Venue - Queensland University of Technology, Brisbane, QLD.

2009. Volume 31. Host and Venue - Victoria University, Wellington, New Zealand.

2008. Volume 30. Host and Venue - University of Wollongong, NSW.

2007. Volume 29. Host and Venue - University of Ballarat, VIC. First running of HDKM.

2006. Volume 28. Host and Venue - University of Tasmania, TAS.

2005. Volume 27. Host - University of Newcastle, NSW. APBC held separately from 2005.

2004. Volume 26. Host and Venue - University of Otago, Dunedin, New Zealand. First running of APCCM.

2003. Volume 25. Hosts - Flinders University, University of Adelaide and University of South Australia. Venue - Adelaide Convention Centre, Adelaide, SA. First running of APBC. Incorporation of ACE. ACSAC held separately from 2003.

2002. Volume 24. Host and Venue - Monash University, Melbourne, VIC.

2001. Volume 23. Hosts - Bond University and Griffith University (Gold Coast). Venue - Gold Coast, QLD.

2000. Volume 22. Hosts - Australian National University and University of Canberra. Venue - ANU, Canberra, ACT. First running of AUIC.

1999. Volume 21. Host and Venue - University of Auckland, New Zealand.

1998. Volume 20. Hosts - University of Western Australia, Murdoch University, Edith Cowan University and Curtin University. Venue - Perth, WA.

1997. Volume 19. Hosts - Macquarie University and University of Technology, Sydney. Venue - Sydney, NSW. ADC held with DASFAA (rather than ACSW) in 1997.

1996. Volume 18. Host - University of Melbourne and RMIT University. Venue - Melbourne, Australia. CATS joins ACSW.

1995. Volume 17. Hosts - Flinders University, University of Adelaide and University of South Australia. Venue - Glenelg, SA.

1994. Volume 16. Host and Venue - University of Canterbury, Christchurch, New Zealand. CATS run for the first time separately in Sydney.

1993. Volume 15. Hosts - Griffith University and Queensland University of Technology. Venue - Nathan, QLD.

1992. Volume 14. Host and Venue - University of Tasmania, TAS. (ADC held separately at La Trobe University).

1991. Volume 13. Host and Venue - University of New South Wales, NSW.

1990. Volume 12. Host and Venue - Monash University, Melbourne, VIC. Joined by Database and Information Systems Conference which in 1992 became ADC (which stayed with ACSW) and ACIS (which now operates independently).

1989. Volume 11. Host and Venue - University of Wollongong, NSW.

1988. Volume 10. Host and Venue - University of Queensland, QLD.

1987. Volume 9. Host and Venue - Deakin University, VIC.

1986. Volume 8. Host and Venue - Australian National University, Canberra, ACT.

1985. Volume 7. Hosts - University of Melbourne and Monash University. Venue - Melbourne, VIC.

1984. Volume 6. Host and Venue - University of Adelaide, SA.

1983. Volume 5. Host and Venue - University of Sydney, NSW.

1982. Volume 4. Host and Venue - University of Western Australia, WA.

1981. Volume 3. Host and Venue - University of Queensland, QLD.

1980. Volume 2. Host and Venue - Australian National University, Canberra, ACT.

1979. Volume 1. Host and Venue - University of Tasmania, TAS.

1978. Volume 0. Host and Venue - University of New South Wales, NSW.

Conference Acronyms

ACDC	Australasian Computing Doctoral Consortium
ACE	Australasian Computer Education Conference
ACSC	Australasian Computer Science Conference
ACSW	Australasian Computer Science Week
ADC	Australasian Database Conference
AISC	Australasian Information Security Conference
AUIC	Australasian User Interface Conference
APCCM	Asia-Pacific Conference on Conceptual Modelling
AusPDC	Australasian Symposium on Parallel and Distributed Computing (replaces AusGrid)
CATS	Computing: Australasian Theory Symposium
HIKM	Australasian Workshop on Health Informatics and Knowledge Management

Note that various name changes have occurred, which have been indicated in the Conference Acronyms sections in respective CRPIT volumes.

ACSW and HIKM 2012 Sponsors

We wish to thank the following sponsors for their contribution towards this conference.



CORE - Computing Research and Education,
www.core.edu.au



RMIT University,
www.rmit.edu.au/



**AUSTRALIAN
COMPUTER
SOCIETY**

Australian Computer Society,
www.acs.org.au



**THE UNIVERSITY OF
MELBOURNE**

The University of Melbourne
www.unimelb.edu.au



Curtin University

Curtin University
www.curtin.edu.au

CONTRIBUTED PAPERS

The role of Emotional Intelligence on the resolution of disputes involving the Electronic Health Record

Emilia Bellucci¹, Sitalakshmi Venkatraman², Nial Muecke², Andrew Stranieri²,
Jemal Abawajy¹,

¹ Deakin University,

Australia

² Centre for Informatics and Applied Optimisation, University of Ballarat,
Australia

emilia.bellucci@deakin.edu.au, s.venkatraman@ballarat.edu.au, nmuecke@gmail.com, a.stranieri@ballarat.edu.au,
jemal.abawajy@deakin.edu.au

Abstract

Numerous authors have expressed concerns that the introduction of the Personally Controlled Electronic Health Record (PCEHR) will lead to an escalation of disputes. Some disputes will concern the accuracy of the record whereas others will arise simply due to greater access to health care records. Online dispute resolution (ODR) programs have been successfully applied to cost-effectively help disputants resolve commercial, insurance and other legal disputes, and can also facilitate the resolution of health care related disputes. However, we expect that health differs from other application domains in ODR because of the emotional engagement patients have with their health and those of loved ones. In this study we will be looking at whether the success of an online negotiation is related to how people recognise and manage emotions, and in particular, their Emotional Intelligence score.

Keywords: Dispute resolution, Electronic health record, Emotional intelligence, EHR.

1 Introduction

An Electronic Health Record (EHR) is a virtual record of every health related event (e.g hospital admission, general practitioner visit, allergies) experienced by an individual from in-utero to death. The establishment of an EHR is a national priority in Australia, because it is believed it will improve the quality and efficiency of health care delivery by empowering Health Care Professionals (HCP) with a full description of a patient's history. Challenges to the establishment of an EHR centre around adoption issues and may include data inter-operability, security, privacy and terminological challenges.

A key component of the United States recent health care reforms involves the expenditure of many billions to establish an EHR by 2013. Microsoft and Google have established EHR systems that currently can accept entries automatically from many general practitioner and hospital information systems. In Australia, over \$400 million dollars has been earmarked for the establishment of an EHR, named the PCEHR, due for release in July 2012.

Halamka (2009) has forecast that the emergence of an EHR will bring with it disputes regarding access and use of information in a record. Some of the disputes will be serious and obvious breaches of privacy or security and require resolution by recourse to Courts. However, many disputes will not be so serious or involve jurisdictional issues that render legal action impractical. For instance, disputes involving the accuracy of data in a patient's EHR are unlikely to be litigated. Different protocols for sharing health data across disciplines or countries are similarly unlikely to be resolved by litigation unless breaches are serious. Recently, Online Dispute Resolution (ODR) approaches have emerged to help disputes between buyers and sellers in online markets, (e.g SquareTrade www.squaretrade.com) and insurance claims (e.g SmartSettle www.smartsettle.com; Cybersettle www.cybersettle.com). Bellucci et al (2008) demonstrates the successful use of AssetDivider as software to help divide assets in a divorce. These models appear to successfully deal with financial issues without making provisions for emotion management. However, when dealing with disputes related to EHR, we feel the emotions of the concerned parties, such as patients and health care professionals will play a vital role.

Borland et al, (2010) and Foo et al, (2010) have found that the success of a negotiation is dependent upon the Emotional Intelligence (EI) of the participants. Goleman (1995) defines EI as the ability to recognise and manage one's own emotions and read and deal effectively with other peoples' feelings. Recently, studies have been carried out to investigate if EI could influence informal mediation (Boland and Ross, 2010). In our paper, we postulate that EHR disputes may be influenced by the emotional intelligence of the disputants. Using Argumentative Theory of Reasoning (Mercier and Sperber, 2011), we propose the important role of EI in resolving

Copyright © 2012, Australian Computer Society, Inc. This paper appeared at the 5th Australasian Workshop on Health Informatics and Knowledge Management (HIKM 2012), Melbourne, Australia, January-February 2012. Conferences in Research and Practice in Information Technology (CRPIT), vol. 129. K. Butler-Henderson and K. Gray and, Eds. Reproduction for academic, not-for profit purposes permitted provided this text is included.

disputes related to EHR.

Our paper will describe a laboratory study we are conducting to support our research. We have asked 80 participants to complete two online activities. The first involves an EI test, which will provide us with an indication of Emotional Intelligence Scores. Participants will be given a hypothetical case detailing a dispute in the area of Electronic Health Records. The next activity is to attempt to resolve the dispute using an Online Dispute Resolution system, Re-consider (Muecke and Stranieri 2006), where participants will be acting on behalf of either the patient or medical practitioner.

The rest of the paper is organised as follows. Section 2 describes the emotions that play an important role in any negotiation process. The various types of emotions are discussed and we arrive at a definition of emotion that relates to negotiation. We briefly describe negotiation theory and how to deal with emotions during negotiations that lead us to the concept of Emotional Intelligence. In Section 3, we explain EI and its importance in the negotiation of disputes. Next, in Section 4, we identify the types of disputes that could be associated with issues related to EHR. This section also gives a brief overview of the dispute resolution models found in the literature and the recent trends in dispute resolution mechanisms. In Section 5, we introduce the importance of a disputant's EI to the dispute resolution process and postulate the link between EI and health care disputes. We employ Argumentative Theory of Reasoning in arriving at our conclusions, and finally, in section 6 we summarise our research.

2 Emotions in Negotiation

When it comes to managing our health, we often have very strong feelings, particularly if we, or someone we care about has been given a poor prognosis, or is the victim of communication breakdowns. Whilst the implementation of EHR promises efficiencies in patient care and accountability, it is also expected to bring with it new disputes related to privacy, accuracy, ownership (property) and accessibility (Mason, 1986). In Section 4.1 we detail an adaption of Mason's theory to arrive at seven different types of disputes related to implementation and use of the EHR. We postulate emotion as an important facet to dispute resolution, and one we feel has the potential, if recognised and properly managed, to contribute to the success of e-health related disputes.

2.1 Definition of Emotion

Emotion is difficult to define without reference to an experience (Fisher and Shapiro, 2005). We experience emotions, and as such define emotion using words that express feelings, ie we feel anger, sadness, happiness, joy. Lazarus (1991) gives a more theoretical definition by defining emotion using three distinct features: physiological reactions, action tendencies and subjective experience.

Emotions are also defined by what they are not; Li and Roloff (2006) make a clear distinction between emotions

and moods. Emotions are usually discrete, of high intensity, short duration and are directed to an object, person or event (Li and Roloff, 2006; Van Kleef et al, 2004). Moods on the other hand are generally "more pervasive, enduring and less intense than emotions, and may not have an identifiable target" (Barry, 1999). Moods are identified in broad terms, such as good or bad (Li and Roloff, 2006), while emotions come in a larger variety of experiences – happiness, sadness, anger, disappointment (Li and Roloff, 2006). It is generally understood the broader concept of affect encompasses both emotion and moods (Li and Roloff, 2006, Barry and Oliver 1996, Van Kleef et al, 2004, Borland and Ross, 2010).

For the better part of the 1980s and 1990s, emotion in negotiation has been largely ignored, whilst rational decision making was 'treasured in negotiation' (Li and Roloff, 2006 p170). Early negotiation texts (Lax and Sebenius, 1986) do not acknowledge the presence of emotions in negotiation. Ogilvie and Carsky, (2002) came to a similar conclusion when reviewing fourteen books on negotiation published between 1970 and 1990 and found none with the word 'emotion' in their indexes.

Since the late 1990's, negotiation research has acknowledged emotion as an important part of the negotiation process. It was realised we can make better collaborative decisions if we acknowledge and better understand emotions. Shapiro (2002) and Adler et al, (1998) were among the first researchers to attest emotions cannot and should not be ignored in a negotiation. This increased interest in negotiation and emotion has resulted in work on the role of emotion and moods in negotiation (Van Kleef and De Dreu, 2004); (Li and Roloff, 2006); emotions in cultural negotiations (Brett, 2000, Adam et al, 2010) and active negotiation around which emotions we experience and how intensely we experience them (Shapiro, 2002).

Models of negotiation that involve emotion are a very recent addition to the negotiation landscape. Martinovski and Mao (2009) describe a model which takes the active role of emotions in decision making, and uses it as a modifier to theory-of-mind models, goals and strategies. Findlay and Thagard (2011) used cognitive-affective mapping to track emotional changes in the 1978 Camp David negotiations.

Whilst emotions cannot be ignored in negotiation, neither should they be allowed to 'flood' negotiations to the extent they drive the agenda away from substantive goals, reveal information we would have preferred not to diverge, and disrupt our thinking (Fromm, 2008). Fisher and Shapiro (2005) understand this by advocating methods by which we can manage emotion, and state simply negotiation involves both reason and emotion.

2.2 Types of Emotion

The emotions most likely to arise in dispute resolution centre on cognitive, affective and behaviour aspects (Ogilvie and Carsky, 2002).

Cognitive aspects apply to emotions we experience as a result of participant action or inaction. These emotions can

be grouped into negative and positive emotions, based on whether our negotiation goals are blocked or not (Bodtker and Jameson, 2001).

Affective aspects centre around mood inducing responses and in particular the interplay between emotion and moods. Moods and emotion are often considered to be interdependent (Davidson, 1994). We are more likely to respond positively and creatively to a situation (such as negotiation) if we are in a happy mood. Conversely, we will also respond negatively (angry or fearful) if we are in a bad mood (Forgas, 1998, Baron, 1990, Carnevale and Isen, 1986, Isen et al, 1987).

Behavioural aspects dictate the manner in which emotional experience is expressed. For example, when we are emotional, we find our voices rise louder, our heart rate is faster. If negotiators can recognise both the non-verbal and verbal cues for emotion, they are better equipped to understand and manage emotions in themselves and others. Several authors (Li and Roloff, 2006, van Kleef et al, 2004) distinguish between intrapersonal effects and interpersonal effects of emotions (and moods) on negotiators. Intrapersonal effects refer to the influence of one's emotions on an opponent's behaviour, while interpersonal effects refer to the influence of one's emotions on their own behaviour.

Much of the literature has focused on the effect intrapersonal emotions can have in negotiations (Forgas, 1998; Isen et al, 1987, Baron, 1990; Carnevale and Isen, 1986). Van Kleef et al, 2004 in their laboratory study found a disputant's emotions were affected by their opponents' emotions. For example, negotiation with an angry opponent reported more anger than did disputants who had negotiated with a happy or non-emotional opponent. Similarly, (Shapiro 2002) argues emotions can positively affect our ability to reach negotiation goals. Positive intrapersonal effect has also been attributed to increases in concession making, preferences for cooperation and increases the use of cooperative negotiation strategies (Bellucci, 2004).

So what kinds of emotions do negotiator's experience? Fisher and Shapiro (2005) present the "core concerns" as a model of five areas that stimulate the many emotions that arise in negotiations. These are:

1. Appreciation –through feelings or actions,
2. Affiliation – sense of belonging,
3. Autonomy – respect of your freedom,
4. Status - your standing in negotiation; and
5. Role – defining your role and its activities.

The above emotion groups relate to how negotiators feel, and as such relate to "human interests". Psychologists argue the most powerful interests for well-being are human interests. Wertheim et al., (1992) groups human interests into: security, economic well-being, a sense of belonging, recognition, and control over one's life. We argue the emotions felt during a negotiation are dependent on how much we feel appreciated, our affiliation (sense of belonging), autonomy (control over one's life), status (recognition) and role.

Negotiations can bring rise to the whole myriad of emotions. The following table from Fisher and Sharpiro (2005 p13), give us an indication of the types of emotions experienced in negotiation.

Positive Emotion	Negative Emotion
Excited, glad, amused, Enthusiastic, cheerful, jovial, delighted, ecstatic	Guilty, ashamed, humiliated, embarrassed, regretful,
Proud, gratified, happy, jubilant, thrilled, overjoyed, elated	Envious, jealous, disgusted, resentful, contemptuous
Relieved, comforted, content, relaxed, patient, tranquil, calm	Impatient, irritated, angry, furious, outraged, Intimidated, worried, surprised
Hopeful, in awe, wonder	Fearful, panicked, horrified, Sad, hopeless, miserable, devastated

Table 1. Emotion Words (Fisher and Sharpiro, 2005, p13)

Whilst it is important to recognise emotions in negotiation, it is equally important to learn how to manage emotions to facilitate successful outcomes. The following negotiation theories will provide us with cues to frameworks that support the management of emotions.

2.3 Negotiation theory

Negotiation is a process by which two or more parties conduct communication or conferences with the view of resolving differences between them (Bellucci, 2004). Ponte and Cavenagh (2004) define negotiation as a process of refining and agreeing to the issues requiring resolution, establishing a range of compromise options from which to choose and selecting the appropriate option for settlement. Negotiation is often the first method of dispute resolution that is called upon to resolve social conflicts, and is often preferred to harsher substitutes such as court trials (Guillemin, 2011) or war.

There are a number of theories to describe how stakeholders negotiate; the two major, as proposed by Walton and Mckersie (1965) are: distributive (zero-sum) and integrative (collaborative) theories.

In distributive approaches, the problems are seen as zero sum and resources are imagined as fixed. In integrative approaches, problems are seen as having more potential solutions than are immediately obvious and the goal is to expand the pie before dividing it. An example of a distributed approach is Positional Negotiation. Positional Negotiation is based on the premise that one takes a position in a dispute and argues it. Occasionally concessions will be made in order to avoid a stalemate and ultimately any solution from the negotiation will reflect a win-lose (one disputant will win, while the other loses).

Integrative (cooperative) negotiation describes the communication of parties when the outcomes are the result of coordinated behaviour of both participants (Robertson et al., 1990). Parties are more likely to be

satisfied with (and most importantly adhere to) suggested solutions if they participated in reaching the solution. Whilst reaching a solution indicates success in negotiation, it is how well the parties adhere to the solution which truly makes a negotiation resolved. Our research is based on the assumption that participants desire to co-operate in negotiation (or can be persuaded to), hence increasing the likelihood that solutions arrived by negotiation are successfully implemented. An example of integrative approaches includes Principled Negotiation, developed by the Harvard Negotiation Project. It advocates proposing an argument based on the disputant's interests which support their position. It also promotes cooperation among disputants by advocating a joint search for options and use of objective criteria. Parties attempt to accommodate as many interests of the parties as possible, leading to the so called win-win or all gain approach.

As Kersten (2001) notes, although Walton and McKersie (1965) did not suggest one type of negotiation as superior to the other; over the years, it has become conventional wisdom that the integrative type allows for better compromises, win-win solutions, value creation and expanding the pie. Fisher and Ury (1981) and Lax and Sebenius (1986) discuss these issues in detail.

Ogilvie and Carsky (2002) state most negotiations have elements of both distributed and integrative components, and as such negative emotions from goal blocking and conflict are often inevitable. However, once participants find ways to work collaboratively, positive emotions can result as well. Fisher and Shapiro (2005) advocate meeting the 'core concerns' briefly mentioned in Section 2.2 as a way to promote positive emotion. Positive emotion promotes a collaborative environment; which "signal cooperativeness and trustworthiness, elicit cooperation, trust, and concession from others, and promise rewards for others" (Li and Roloff, 2006 p 172).

We believe negotiation using a collaborative approach and supporting the importance of emotions is likely to result in successful outcomes. We now look to the field of EI for insight on how to recognise, use, understand and manage emotions; which in turn may positively affect the resolution of health care related disputes more effectively.

3 Emotional Intelligence

Emotional Intelligence (EI) is the ability to recognise and manage one's own emotions and those of others. It is defined as the ability to understand and use emotions adaptively in everyday life (Mayer and Salovey, 1997).

The concept of Emotional Intelligence relates to theories on how humans express emotion. It states emotional responses to our environment differ from person to person. There are several measures to quantify one's emotional intelligence, usually referred to an EQ (Emotional Quotient). It is generally understood emotionally intelligent (high EQ) individuals are better equipped to deal with emotional responses and in general with human interaction. Those with a low EQ are less likely to recognise emotions and act on them appropriately. It is generally understood that negotiators need appropriate

levels of cognitive and emotional intelligence in order to negotiate successfully (Thompson, 2001; Fulmer and Barry 2004). A negotiators' cognitive intelligence (measured through IQ) is particularly important if the negotiation is complex and variables constantly change (Fulmer and Barry, 2004). Whilst our paper does not describe the role of cognitive intelligence in any detail, we assume participants in our study will have above average and relatively similar IQ levels, consistent with the IQ of third year university students studying similarly rigorous courses.

3.1 Instruments to measure EI

There are a number of different instruments available today to elicit a respondent's EI; including Bar-On EQ-i (Bar-On, R. 2000), ECI (Boyatzis and Sala, 2004), Goleman's Emotional Intelligence Appraisal (Goleman 1995) and the MSCEIT (Mayer and Salovey, 1997), among others. Additionally many EI instruments have been designed specifically for a certain use, for example, Team analysis and building (Druskat and Wolff, 2001), The EQ Map for the workplace and leadership (Cherniss and Goleman, 2001), business education (Tucker et al 2000), and in e- procurement (Higgs and Reynolds, 2002).

Most researchers agree the MSCEIT (Mayer and Salovey, 1997) is the most appropriate and commonly used measure in negotiation and social interaction (Fulmer and Barry, 2004; Mueller and Curhan 2006). Mayer et al, (2003) and Mueller and Curhan (2003) also claim the MSCEIT is of high reliability, which in turn produce studies of relatively high levels of internal validity.

The MSCEIT involves understanding around four major branches, "perceiving, facilitating, understanding and managing emotion" (Mayer et al, 2002), (Caruso and Salovey, 2004).

- **Perceiving emotion** refers to identifying emotions in ourselves and others.
- **Facilitating emotion** refer to how we use emotions. It asks us to reflect on how moods impact our thinking; and the relationship between various sensations to emotions.
- **Understanding emotion** relates to how emotions can change over time and how we define emotion.
- **Managing emotion** relates to emotion management and emotional relations. We are asked to reflect on which solutions would be most effective in resolving internal problems (to self), as well as problems involving other people.

Ogilvie and Carsky (2002) claim the MSCEIT appears most suited to the study of negotiation, as all levels of the model have relevance and application to negotiation. For instance, negotiation involves information acquisition, reaching consensus and decision making, performed collaboratively among disputants. Whilst Fisher and Sharpiro (2005) make the relationship between emotions and negotiation clear, they also state it is not enough to

recognise the emotions, but also to know how to manage and leverage emotion in negotiation. Emotional intelligence can help disputants understand the emotions they feel whilst negotiating how others feel and the management of these emotions (Ogilvie and Carsky,

2002). A number of studies have been conducted analysing the affect of one's EI on dispute resolution, which we detail in the next section.

3.2 EI and Dispute resolution

Numerous studies have been conducted analysing the relationship between EI and negotiation. Boland and Ross (2010) in their research into leadership, negotiation and EI, state a leader's EI as a good indicator of how well they lead or manage others. Conclusions from their study suggest high EQ disputants were more likely to seek mutually satisfying agreements; while those with a Low EQ would try to put a stop to conflict by not addressing underlying issues (for example by compensating or putting pressure on disputants to settle). Fulmer and Barry (2004) describe how Cognitive Intelligence and Emotional Intelligence helps negotiators succeed in the areas of Information Acquisition and decision making soundness. Mueller and Curhan (2006) establish a disputant's EI can affect the satisfaction of the counterpart negotiator. They state a participants' ability to understand emotion positively predicts their counterpart's outcome satisfaction.

We postulate that whilst the presence and management of emotion is important in all disputes, it is in health we expect emotions are expressed most passionately. Whilst research has been conducted in the relationship and subsequent affect of EI in ADR, little has been discussed by the way of disputes relating to the EHR. In addition, in disputes over the management of an individual's health, the EHR will bring additional disputes regarding its management and use, including the longevity, manipulation, security, access and maintenance of the electronic record.

4 Disputes and EHR

The purpose of this paper is to lay the foundation for our premise that the much established role of emotions in negotiation could be employed as a vehicle to facilitate resolving disputes that are possible in EHR. Since scientific measures could be associated with one's EI, it is now important to forecast and identify what could be the types of disputes related to EHR and the level at which EI could play a role in resolving each of these dispute types. Our literature survey and consultations with practice professionals related to EHR and disputes have resulted in seven major types of disputes described in this section. We have also summarised the existing dispute resolution models, their trends and the inherent impact of emotions on disputants going through any dispute resolution process.

4.1 Types of disputes

In this paper, we consider the definition of disputes given by Felstiner, Abel & Sarat (1980) as a particular form of conflict between two parties, where one party (called filer) makes a claim that is rejected by another party (called respondent). In the EHR related disputes, normally the filer is the patient and the respondent is predominantly the doctor or health care practitioner, who is the creator of the record.

With information abundantly available on the Internet, networked databases, and other electronic sources, disputes or questioning the truth or validity of information is inevitable, and affects the way people interact and businesses operate, for example disputes due to self-diagnosis (Ryan and Wilson, 2008). Disputes over health care records are no exception and are gaining more attention of late due to their storage and access in electronic form.

With paper-based health care records, such records are under the control of the health care providers. Gaining access to records can be time consuming and difficult. However, with health care records available in electronic form and becoming easily accessible to patients as well as health care authorities, disputes could take a wider scope and may even lead to legal challenges. With health care records in electronic form, researchers and health-related practice professionals have recently started to predict the kind of EHR related disputes possible in the years to come. Mason (1986) outlined the following ethical considerations of the information age - privacy, accuracy, ownership (property) and accessibility. It is likely disputes will arise if the above issues are not properly met. We have identified such disputes and have grouped them under seven main EHR issues listed below:

- i. **Privacy** – data could be misused due to interoperability among insurers, doctors, hospitals and other health care providers;
- ii. **Practice Compliance and Trust** – how authentic is the source or the creator;
- iii. **Integrity of data** - data could be inaccurate, incomplete or unclear;
- iv. **Availability of data** – data could be unavailable due to compromise of system or system problem;
- v. **Data Accuracy** – Software flaws in updating records or in the storage of data;
- vi. **Access Control of data** – transparency in the rules of data access and control mechanisms, namely who are given access, who controls the data, what happens after the completion of treatment or death of a person; and
- vii. **Data Security** – encryption mechanisms, breach of access by unauthorised persons.

Although most of the above types of disputes are generic and could be applied to any electronic system, such as online buying and selling disputes, the effect of these disputes have a greater impact when it comes to disputes over health care records. This is because, such disputes not only lead to financial impact and lack of trust in the system, but more importantly could have a range of impacts, from the quality of care to, the patient's very life. In addition, among these seven types of EHR

disputes, our premise is that the first two, namely i) Privacy and ii) Practice Compliance and Trust would most benefit from the management of emotions.

4.2 Dispute resolution models

A more generic term used with regard to dispute resolution is conflict management and early conflict management models (Blake and Mouton, 1964). These consist of five types of handling interpersonal conflicts, namely forcing, withdrawing, smoothing, compromising, and problem solving. Subsequently, models based on intentions and interests of parties involved have emerged with classifications using assertiveness and cooperativeness levels (Thomas, 1976; Pruitt, 1983). They consider problem solving models as those yielding assertiveness and high cooperativeness, which could be a preferred model as it could be mutually beneficial.

With more and more disputes evolving specific discipline areas such as e-commerce, supply-chain, education, health, environment, intellectual property, land use, residential tenancy, labour and the like, one could find the evolution of four conflict resolution models, professional, bureaucratic, legal, and mediation models (Neal and Kirp, 1985). The professional model has a bias towards the professional expertise, and the bureaucratic model is less discretionary of circumstances, following only regulatory standards. Hence, the more popular models for dispute resolution are legal models, where the focus shifts from agency compliance towards individual rights and entitlements, and the negotiation model, where non-adversarial joint problem solving processes are adopted. In the negotiation model, if the process involves the development of a mutual agreeable outcome, then there could be a win-win relationship between the filer and the respondent of a dispute (Goldberg and Huefner, 1995). Some dispute resolution models were later developed to address group conflicts (Khun and Poole, 2000), where the approach is either distributive or integrative of the needs and concerns of the two groups.

4.3 Trends in dispute resolution

In the past couple of decades, disputes that were requiring legal resolution have resorted to Alternative Dispute Resolution (ADR). Due to delays and high costs involved in resolving disputes in court, legal practitioners have been adopting ADR (Landerkin and Pirie, 2003), where an arbitrator mediates between the parties informally and resolves the issues by going through a clarification process. In many situations such as supply chains, ADR is considered as a valued approach and well accepted by parties to opt in for a faster settlement of a dispute, thereby avoiding costs associated with court-based litigation and delays. ADR approaches are utilising advances in IT to reflect an online presence, and these approaches are called Online Dispute Resolution (ODR). ODR has been developed more recently by adapting ADR principles onto the Internet by making use of Web-based software systems. The parties and arbitrators work with facts and documents made available online and make use of web

meetings for the negotiation process (Rabinovich- Einy and Poblet, 2008). Such an approach to ODR is a simple adaptation of ADR. For disputes related to electronic health records, ODR is much easier to approach as the patient's health related documents and facts are already available online.

The majority of the present ODR related studies are concentrating on such simple adaptations of ADR. This paper argues that more sophisticated modelling could be incorporated once IT comes into the scene. One of the main concerns of whether the parties are emotionally prepared to cope with the risk and uncertainty involved in the mediation process have not been addressed so far. This could be addressed with ODR since simple online emotional checks could be performed.

According to Gross (2002), the parties do carry negative emotional experience during and after the negotiation process. They may require cognitive energy and physiological restraint to suppress their emotions. These emotional factors have a greater impact when it comes to disputes related to electronic health records. Hence, in this paper we give importance to a new emotional dimension that has been overlooked in literature. Next, we employ Argumentative Theory to establish the link between EI and dispute resolution related to EHR.

5 Emotional Intelligence and disputes relating to the EHR

In this section, the main aim is to establish the importance of the connection between EI and health care disputes using Argumentative Theory of Reasoning, the theory suggested by Dan Sperber. We adopt an intuitive inference as a mechanism to arrive at arguments used in reasoning (Mercier and Sperber, 2011), for accepting the conclusion that EI and disputes are very much linked to each other.

We have argued through the literature survey that recent models of negotiation make use of emotions in decision making to arrive at negotiations successfully (Li and Roloff 2006; Martinovski and Mao, 2009). According to Fisher and Shapiro (2005), the five 'core concerns' that arise in any negotiation are used as a guide to recognise emotions in ourselves and our opponents.

Since the capability of recognising and managing emotions in ourselves and others has already been coined as Emotional Intelligence (EI) by Mayer and Salovey (1997), we infer with other similar consensus in literature that EI has a role to play in negotiations. According to some recent studies (Ogilvie and Carsky, 2002; Mueller and Curhan, 2006; Boland and Ross, 2010), EI not only helps disputants to manage their emotions during the negotiation process, but also in achieving satisfaction of the outcome. Hence, through the inference mechanism of Argumentative Theory of Reasoning, we have established from existing research studies that the EI of disputants has an effect on the outcome of a negotiation. Can we generalise this premise to health care disputes, in particular disputes related to EHR?

Our analysis of literature suggests that recent health related dispute resolution trends are more inclined towards adopting negotiation models as patients are quite sensitive about their care records that are personal and private (Washington, et al. 2009). Following the EI theory of Goleman (1996), EI provides the inter-personal skills, which when fostered in health care organisations, could result in establishing lasting relationships with patients and partnerships (Morse 1991; McQueen 2004). With EHR systems, there is a higher need to provide a means of communication and negotiation between consumers and health providers as dispute resolution means to deal with issues related to EHR (Washington, et al. 2009). Not only do patients go through emotional distress due to their personal experiences, the health care professionals also experience emotional responses to their patients' suffering and need to adopt EI skills to deal with such situations.

Hence, through the inference mechanism of Argumentative Theory of reasoning, we postulate that the link between EI and EHR related disputes is much more profound, and should be fostered in designing systems that facilitate negotiations in dispute resolution. Many research studies consider EI as a core competency in health care organisations to prevent disputes as much as possible, and to deal with issues directly with care and quality of service for building trust and ongoing consumer relationship (Semple & Cable 2003; Freshman & Rubino, 2002).

While other industries make extensive use of ODR tools (ie ebay), we believe EHR systems should also incorporate ODR functionality to facilitate dispute resolution. Notwithstanding the success of ebay's ODR model, we believe incorporation of an ODR system within an EHR framework has the following advantages:

- Online system functionality will help stakeholders in understanding the informational content of their EHR records, and would aid in better addressing issues relating to health literacy. For example, there could be requests to make an amendment to a patient's EHR, which, depending on the experience and confidence of the health professional, combined with legal ramifications from the modification of EHR; make EHR more difficult to manage than with paper-based records.
- Medical practitioners, administrators and other stakeholders may also have additional stress in dealing with overheads involved with EHR management systems and in understanding other opportunities for error.
- Since patients and doctors are most likely involved in an ongoing relationship, the need to increase satisfaction and good will in both parties is essential.

Therefore, it is our premise that EI awareness when incorporated with EHR systems will result in a higher quality of service, and facilitate a satisfied dispute resolution through better management of disputant emotions during a collaborative negotiation process.

5.1 Research plan

Whilst the field of ODR is well advanced in tools

made available, little work has been done in discovering the traits of a successful ODR process. Borland et al, (2010) and Foo et al, (2010) have found that the success of a negotiation is dependent upon the Emotional Intelligence (EI) of the participants which is the ability to recognise and manage one's own emotions and read and deal effectively with other peoples' feelings (Goleman 1995). In addition, EI involves the ability to use this information to guide one's thinking and action (Salovey and Mayer, 1990). We hypothesize the EI of a negotiator will have an impact on the success of ODR processes.

This project will assess the Emotional Intelligence of participants and whether there is a strong correlation between one's Emotional Intelligence and success in a negotiation relating to the area of e-health.

We will be asking a large number of voluntary participants (80) to complete two online activities. Participants will be recruited from undergraduate students completing their major sequence in Health Informatics. Our proposed study has ethics approval by our universities (Number: BL-EC 41-11 Bellucci).

Participants will be asked to complete two online activities. The first relates to an online version of the MSCEIT test. The test is in the form of multiple choice questions which will assess how participants identify emotions present, use emotions to help a participant think and solve problems, understand the causes of the emotions and manage the emotions to obtain a positive result (Salovey & Mayer, 1990). After participants have worked through the MSCEIT inventory, an EI score will be provided. Next participants will be directed to commence working through a case study in EHR using Re-Consider (Muecke and Stranieri 2006), our online dispute resolution program. Reconsider, as we will discuss next, negotiates disputes by allowing disputants to re-evaluate their claims against a hierarchy of possible claims.

5.2 Re-Consider

The Reconsider ODR approach utilises a model of the important factors of the dispute and protocol which guides users through said model. In our proposed study, a dispute between a doctor and patient is represented by a hierarchical tree of factors. At the top most level of the tree sits the root node, this node represents the most general factor of the dispute. In Figure 1, the root node indicates the extent to which communication issues played in causing a dispute between the doctor and patient. Below the root node are the most important factors (referred to as nodes) required to determine the state of root node. These nodes cover such themes as: the method of communication used by the doctor/patient and ability of the doctor/patient to accurately convey important information to one another. Below each of these nodes are additional nodes, which hold significance in determining the node above them. Each level of the hierarchical tree become progressively more refined, until base facts can be established (known as leaf node).

Each node in the dispute is presented as statements and the possible assertions disputants can make about the

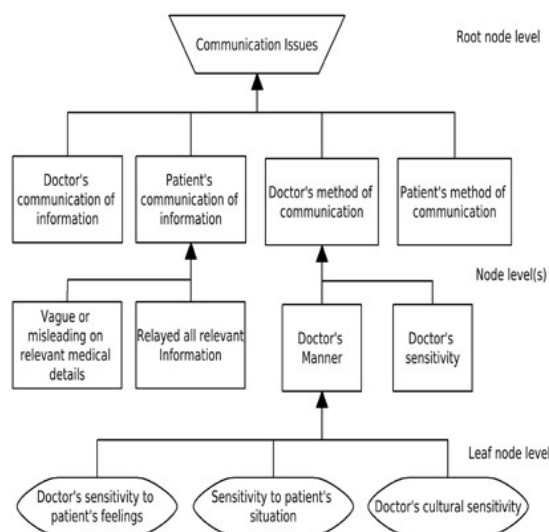


Figure 1. An example of part of a dispute model for communication issues

statement. For example, for the statement “You believe that the information provided by the doctor to the patient was...”, disputants can claim one of the following: a) “quite flawed”, b) “flawed”, c) “not ideal”, d) “neither good nor bad”, e) “good”, f) “very good” and g) “excellent”.

During a dispute, a structure such as the one just described is used by the ReConsider protocol to guide the disputants through the relevant factors of the dispute. Disputants assert their beliefs for each node, progressing from the first level of nodes below the root node to the leaf nodes, whenever a difference in opinion is found to exist. For example, if the doctor and patient agree that the doctor communicated all relevant information, then there is no further need to explore that particular branch of the tree. Once the tree has been fully explored to points of agreement or leaf nodes, the disputants work their way back up the tree toward the root node. This aids the disputants to reconsider the claims they asserted, now with a better understanding of the issues in the dispute.

Lastly, once the disputants have worked their way back to the initial node they asserted claim on, Bayesian inference is used to determine the likely claim value of the root node. This claim is then presented to the users as the recommended solution to the dispute. Agreement on the root node will end the dispute, while disagreement leads the users back to reconsider their positions on the factors of the dispute. The users are prevented from asserting a claim for the root node throughout the dispute, so as to prevent a fixation on the outcome without due consideration of the factors involved.

The case study will involve a dispute arising from the omission of a health professional to properly update a patient's electronic health record. Participants will be asked to act on behalf of either the patient or the doctor involved in the dispute. Before its use, Re-Consider has been coded with the tree (similar to that in Figure 1) that inclusively captures all possible disputes relating to the case study presented. We are currently in the process of conducting the research plan discussed in this section.

We can envisage use of an ODR tool such as Re-Consider to manage disputes in EHR. For example, when the patient and the doctor go through the negotiation process of answering questions at each level of the hierarchy where disagreement exists, their EI could be leveraged to make them understand the facts and information involved in EHR, resulting in a satisfied outcome. We believe the outcome of our research plan will affirm there is a strong correlation between a person's high and low EQ and their capability to negotiate a dispute well.

6 Conclusion

It is recognised that disputant emotions are much more involved in dealing with disputes related to EHR due to the inherent privacy and sensitivity issues associated with one's health. Only recently the concept of EI is finding its place in literature related to health care studies. This paper has adopted Argumentative Theory of Reasoning to suggest the important link between EI and disputes related to EHR that could result in positive emotion management during a collaborative negotiation process. The arguments suggest that EI awareness in EHR systems will lead to better understanding of technological and human issues, and in improving the quality of service in the health care industry. However, further investigation needs to be conducted in line with the issues and discussions we have raised in this paper. We postulate the presence of higher EI stakeholders in health care will contribute to a reduction of EHR related disputes. By alleviating the negative emotions (and moods) that may be involved in the dispute resolution process, it is possible the negotiation will lead to a greater number of satisfied outcomes. With ODR becoming popular as an alternate dispute resolution method, not only do we see potential in the use of ODR in EHR, but also in the recognition and use of EI to facilitate more successful outcomes.

7 References

- Adler R. S., Rosen, B., Silverstein E. M., (1998). “Emotions in Negotiation: How to Manage Fear and Anger”, *Negotiation Journal*, April: 161 - 179.
- Barry B. (1999). The tactical use of emotion in negotiation. *Research on negotiation in organisations*, 7: 93-121.
- Bar-On, R. (2000). Emotional and social intelligence: Insights from the Emotional Quotient Inventory (EQ-i). In R. Bar-On and J.D.A. Parker (Eds.), *Handbook of emotional intelligence: Theory, development, assessment and application at home, school and in the workplace*. San Francisco: Jossey-Bass, pp. 363-88.
- Bellucci, E., (2004). Developing Compensation Strategies for the construction of Negotiation Decision Support Systems. PhD thesis, La Trobe University, Bundoora 3086, Victoria, Australia
- Bellucci, E., (2008). Development of a Mediation Tool in Family Law: AssetDivider. Legal Knowledge and Information Systems Ed. Enrico Francescori, Giovanni

- Sartor and Daniela Tiscornia. IOS Press, Netherlands. 88-97.
- Blake, R. R., & Mouton, J. S. (1964). *The managerial grid*. Houston, TX: Gulf.
- Bodtker A.M., Jameson, J.K. (2001). Emotion in conflict formation and its transformation: Application to organisational conflict management. *International Journal of Conflict Management*, 12, 259-275.
- Boland M.J. and Ross W.H. (2010), Emotional intelligence and dispute mediation in escalating and de-escalating situations, *Journal of Applied Social Psychology*, Vol. 40, No.12, pp. 3059–3105.
- Boyatzis, R. E., Sala, F. (2004). The Emotional Competence Inventory (ECI). *Measuring emotional intelligence: Common ground and controversy*. Geher, Glenn (Ed), (pp. 147-180). Hauppauge, NY, US: Nova Science Publishers, xiv, 277 pp.
- Brett, J. M. (2000). Culture and Negotiation. *International Journal of Psychology*, 35, 97-104.
- Caruso D, Salovey, P. (2004). *The Emotionally Intelligent Manager*. San Francisco: Jossey-Bass.
- Cherniss C, Goleman, D. (2001). The emotionally intelligent workplace: how to select for, measure, and improve emotional intelligence in individuals, groups and organisations. Jossey-Bass, New York.
- Druskat V. U., Wolff, S. B. (2001). Building the emotional intelligence of groups. *Harvard Business Review* pp 81 – 90.
- Felstiner, W. R., Abel, R., Sarat, A. (1980). The emergence and transformation of disputes: Naming, blaming, and claiming. *Law and Society Review*, 15, 3-4, 631-654.
- Fisher and Shapiro, 2005. *Beyond Reason: using emotions as you negotiate*. Penguin Books New York.
- Fisher, R. and Ury, W., (1991). *Getting to Yes: Negotiating Agreement without Giving In*. Boston: Houghton Mifflin
- Foo, M. D.; Elfenbein, H. A.; Tan, H. H., Aik, V.C., (2004), 'Emotional Intelligence and negotiation: the tension between creating and claiming value', *International Journal of Conflict Management*, Vol. 15, No. 4, pp.411 - 429
- Freshman B., Rubino L. (2002) Emotional intelligence: a core competency for health care administrators. *Health Care Management* 20(4), 1–9.
- Fromm, D. (2008). Dealing with Strong Negative Emotions: Emotions in Negotiation. *The Theory and Practice of Representative Negotiation*, Edmond Montgomery Publications Limited.
- Fulmer, I. S., Barry, B. (2004). The Smart Negotiator: Cognitive Ability and Emotional Intelligence in Negotiation. *The International Journal of Conflict Management*, 15, 245-272.
- Guillemin, J.F. (2011). Reasons for choosing alternative dispute resolution in Ingen-Housz A. (Ed.) *ADR in Business - Practice and Issues Across Countries and Cultures* Vol. II, Kluwer, pp 13-48
- Goleman, D. (1995). *Emotional intelligence*. New York: Bantam Books.
- Goleman, D. (1998). What makes a leader? *Harvard Business Review* pp90 – 103
- Gross, J. J. (2002). Emotion Regulation: Affective, Cognitive, and Social Consequences. *Psychophysiology*, 39 (3), 281–291.
- Hajo, A. Shirako A. and Maddux W.W. (2010). Cultural variance in the interpersonal effects of anger in negotiations. *Psychological Science*. 21: 882-889.
- Halamka J. (2009), Life as a Healthcare CIO, Dispute Resolution in Healthcare, April 27, 2009, <http://geekdoctor.blogspot.com/2009/04/dispute-resolution-in-healthcare.html>.
- Kersten, G. (2001). Modeling Distributive and Integrative Negotiations: Review and Revised Characterization. *Group Decision and Negotiation*. 10(6), 493-514.
- Kuhn, T., & Poole, M. S. (2000). Do conflict management styles affect group decision making? *Human Communication Research*, 26, 558-590.
- Landerkin, Hugh F., and Andrew J. Pirie. (2003). "Judges as Mediators: What's the Problem With Judicial Dispute Resolution in Canada?" *Canadian Bar Review* 82(2): 249-298.
- Lax, D. A. and Sebenius, J. 1986. *The Manager as Negotiator*. New York, The Free Press
- Lazarus R. S. (1991) *Emotion and Adaptation*. New York: Oxford University Press.
- Li S., Roloff M. E. (2006) Strategic Emotion in Negotiation: in *Cognition, Emotion, and Culture in Communication to Presence: Cognition, Emotions and Culture towards the Ultimate Communicative Experience* G. Riva, M.T. Anguera, B.K. Wiederhold and F. Mantovani (Eds.) IOS Press, Amsterdam, pp 169-188.
- Martinovski and Mao (2009) Emotion as an Argumentation Engine: Modeling the Role of Emotion in Negotiation. *Group Decision and Negotiation*. Vol 18: 235-259. Springer Science and Business Media.
- Mason, R.O. 1986. Four ethical issues of the information age. *MIS Quarterly* 10(1): 5- 12.
- Mayer, J.D., Salovey, P., Caruso, D.R., Sitranios, G., (2003), Brief report: measuring Emotional Intelligence with the MSCEIT 2.0, *Emotion*, Vol 3 (1), 97–105
- McQueen A.C.H. (2004) Emotional intelligence in nursing work *Journal of Advanced Nursing* 47(1), 101–108.
- Mercier H and Sperber D. (2011) Why do humans reason? *Arguments for an argumentative theory. Behavioural and Brain Sciences*, 34(2):57-74
- Morse J.M. (1991) Negotiating commitment and involvement in the nurse–patient relationship. *Journal*

- of *Advanced Nursing* 17(7), 809–821.
- Mueller, J. S. & Curhan, J. R. (2006). Emotional intelligence and counterpart mood induction in a negotiation. *International Journal of Conflict Management*, 17, 110 - 128.
- Muecke, N. and Stranieri, A. (2006). 'Predictions with Uncertainty to Support Fair Outcomes in Online Legal Disputes'. In: Proceedings of CIMCA06, the International Conference on Computational Intelligence for Modelling. Sydney, pp. 218-220. IEEE
- Neal and Kirp (1985). The Allure of Legalization Reconsidered: the Case of Special Education, *Journal of Law and Contemporary Problems*, vol. 48: (1) 63-87.
- Pruitt, D. G. (1983). Strategic choice in negotiation. *American Behavioral Scientist*, 27, 167-194.
- Ryan, A. & Wilson, S. (2008). Internet healthcare: do self-diagnosis sites do more harm than good? *Expert Opinion on Drug Safety*, 7, 227-229.
- Rabinovich-Einy, O. Poblet, M. (2008) Reflecting on ODR: The Israeli Example Expanding the Horizons of ODR. *Proceedings of the 5th International Workshop on Online Dispute Resolution*.
- Salovey, P. & Mayer, J. D. (1990). Emotional intelligence. *Imagination, Cognition, and Personality*, Vol. 9, 185–211.
- Semple M. & Cable S. (2003) The new code of professional conduct. *Nursing Standard* 17(23), 40–48.
- Shapiro D.L., (2002). "Colloquy: Appreciating Emotion in Conflict Processes: Negotiating Emotions". *Conflict Resolution Quarterly*, Vol 20, no 1: 67 – 82.
- Thomas, K. W. (1976). Conflict and conflict management. In M. D. Dunnette (Ed.), *Handbook in industrial and organizational psychology* (pp. 889-935). Chicago: Rand McNally.
- Tucker, M. L., Sojka, J.K., Barone, F. J. McCarthy, A. M. (2000). Training tomorrow's leaders. Enhancing the emotional intelligence of business graduates. *Journal of Education for Business*, 75, 331-338.
- van Kleef G, A., De Dreu C. K. W., Manstead A.S.R., (2004), "The interpersonal Effect of Anger and Happiness in Negotiations", *Journal of Personality and Social Psychology*, Vol 86, No.1: 57-76.
- Walton, R. and Mckersie, R. (1965). *A Behavioral Theory of Labor Negotiations*, McGraw - Hill, New York.
- Washington, L., Katsh, E. and Sondheimer, N. (2009) "Dispute Resolution: Planning for Disputed Information in EHRS and PHRS." *Journal of AHIMA* 80, no.11: 25-30.
- Yuan Y., Archer N. (2004). Online Dispute Resolution, Using technology to provide dispute resolution services and facilitate business and trade. McMaster e-Business Research Centre (MeRC). McMaster University.

Detection of Evidence in Clinical Research Papers

Patrick Davis-Desmond

Diego Mollá

Department of Computing
Macquarie University
Sydney 2109, NSW, Australia
Email: diego.molla-aliod@mq.edu.au

Abstract

When appraising published clinical research, medical doctors and researchers often need to know whether the clinical outcomes presented had statistical evidence. In this paper we present a study for the detection of expressions of such statistical evidence. An effective rule-based classifier has been developed that uses regular expressions and a list of negation phrases to automatically classify documents as either showing evidence of effect in the results or not. The classifier performed with an accuracy between 88% and 98% at 95% confidence intervals, and it also outperformed a set of baselines using bag-of-word features in several statistical classifiers. The rule-based system is written in Python and is available as open-source code.

Keywords: evidence-based medicine, appraisal, text classification.

1 Introduction

On-line medical databases, such as PubMed¹ and PubMed Central² maintained by the US National Library of Medicine, publish thousands of clinical papers yearly. These free full-text reports can be retrieved to find relevant information for clinical and research purposes. Searching through this research literature to find relevant articles and the best evidence for questions posed by medical practitioners can be a daunting and time-consuming task. Many of the articles retrieved will be irrelevant in cases where the research hypothesis has been rejected due to lack of clinical evidence and statistical proof. The reports most needed by medical practitioners to help find answers to clinical questions are the ones where clinical evidence has been found and the research hypothesis has been accepted.

Current information retrieval systems used to research these on-line databases fail to differentiate be-

tween articles that have clinical evidence or not. As a result, many articles retrieved are not relevant to the medical practitioner. Detection of lack of evidence will allow the bulk of non-relevant articles to be excluded from the search results. Medical practitioners can then focus on reading articles with proven clinical evidence in order to find information of benefit to their patients.

In this paper we present an initial study on the detection of evidence in published medical research papers. We focus on Randomised Controlled Trials (RCT) and show that a rule-based approach that targets the detection of specific expressions of negation in the text gives an accuracy between 88% and 98% at 95% confidence intervals.

The structure of the following sections of this paper is as follows. Section 2 presents work on aspects related to the detection of clinical evidence. Section 3 details the methodology that we have followed. In particular, Section 3.1 shows how the corpus of RCTs has been gathered, Section 3.2 focuses on how the corpus was annotated, Section 3.3 presents a set of baselines that we developed using machine learning methods on bag-of-word features, and Section 3.4 details our rule-based system. Finally, Section 4 presents the conclusions.

2 Related Work

We are not aware of any work that specifically targets the detection of clinical evidence, but there has been work on the detection of polarity of clinical outcomes and substantial work on various tasks related to the detection of negation in clinical texts.

Niu et al. (2006) showed an improvement of the results of a multi-document summarisation approach over clinical trials by incorporating information on the polarity of clinical outcomes. Their polarity detection system classified the clinical studies into one of four types according to whether the outcomes improved the patient outcomes: “positive”, “negative”, “neutral”, “no outcome”.³ They applied SVM on a corpus of 197 abstracts and obtained a maximum accuracy of 82.5%.

Preliminary visual inspection of our corpus of research papers revealed the occurrence of a substantial number of negation expressions that indicate lack of evidence in their findings. We therefore focused our work on detecting those expressions automatically. There is a number of approaches researching negation detection applied to medical texts. One of the better known studies is NegEx (Chapman et al., 2001). NegEx is based on regular expressions and its

The work presented here is based on an extension of coursework of the first author at the Masters of Information Technology, Macquarie University.

Copyright ©2012, Australian Computer Society, Inc. This paper appeared at the 5th Australasian Workshop on Health Informatics and Knowledge Management (HIKM 2012), Melbourne, Australia, January-February 2012. Conferences in Research and Practice in Information Technology (CRPIT), Vol. 129, Kerry Butler-Henderson and Kathleen Gray, Ed. Reproduction for academic, not-for-profit purposes permitted provided this text is included.

¹<http://www.ncbi.nlm.nih.gov/pubmed/>

²<http://www.ncbi.nlm.nih.gov/pmc/>

³The difference between “neutral” and “no outcome” is whether the results indicated no significant evidence (“neutral”) or whether the study did not provide the results (“no outcome”).

focus is on detecting negated findings and diseases in discharge summaries. The algorithm uses several phrases indicating negation and filters out sentences containing “pseudo-negations”, that is phrases that falsely appear to be negation phrases such as double negatives and ambiguous phrasing (e.g. *unremarkable*). The system also limits the scope of negation phrases to a context window with size of five words either side of the concept. Their algorithm uses a predefined set of pseudo-negation phrases, a set of negation phrases, and two simple regular expressions. NegEx performed with an accuracy of about 84% and a recall of about 78%. The rule-based negation classifier developed for our project is a modified and simplified version of NegEx.

There have been several other development efforts based on NegEx. Skeppstedt (2010) evaluated NegEx on clinical health records in Swedish and achieved a precision of 70% and a recall of 81% for sentences containing the negation phrases. Although the results are not as high as results achieved by NegEx, Skeppstedt comments that “the comparison between the English and Swedish evaluations is complicated by the fact that the Swedish test data had lower inter-rater agreement”, which is likely to have affected the Swedish results. The author also notes certain language specific examples such as *icke* (meaning *not* or *non-*), which commonly appears at the start of the disease names like *icke allergisk astma*. NegEx interprets this as negation, thus affecting precision. Goryachev et al. (2006) evaluated four different methods of negation detection, including regular-expression-based algorithms, syntactic-processing-based algorithms, and statistical classifiers including Naïve Bayes and SVM. They modified NegEx and another negation algorithm called NegExpander, the latter developed by Aronow et al. (1999). The study reported a result of 92% accuracy with their modified version of NegEx, which they applied to hospital outpatient reports. Their conclusion was that rule-based classifiers performed better than statistical classifiers. Meystre and Haug (2005) created a modified version of NegEx, so as to detect negation when extracting medical problems from medical records. The negation detection algorithm of NegEx was also modified to match UMLS concepts contained in a keyword table. The aim of Meystre’s work was to develop a natural language processing tool that would harvest potential “problem list” entries from the electronic documents in the LDS Hospital (Salt Lake City, UT) Electronic Medical Record system. These documents comprised free-text documents of patient medical history and reports of medical interventions or clinical progress. The best results were obtained using their MMTx2 tool (75.3% precision and 89.2% recall).

There has been additional work on the detection of negation besides NegEx and its variants. NegExpander (Aronow et al., 1999) was designed to detect expressions of observed evidence⁴ in radiology reports and uses syntactic processing techniques to identify noun phrases or conjunctive phrases that define negation boundaries. NegFinder (Mutalik et al., 2001) uses regular expressions and a parser to identify negation in hospital discharge summaries and surgical notes. Mutalik et al. noted that “MEDLINE indexing uses sophisticated syntactic and semantic processing techniques, but does not incorporate explicit distinctions between positive and negative terms”. The algorithm finds negated concepts in discharge summaries and surgical notes with 91.8% accuracy

and 95.7% recall. ChartIndex (Huang and Lowe, 2007) uses regular expressions and parse trees to locate negated medical concepts in radiology reports. ChartIndex was developed in order to automatically extract meta-data as part of the STRIDE (Stanford Translational Research Integrated Database Environment) system from the over one million full-text pathology reports stored in the STRIDE Clinical Data Warehouse (CDW). ChartIndex achieved accuracy results between 85% and 92%.

Recent work by Uzuner et al. (2009) compared machine learning and rule-based approaches to classifying discharge summaries, and achieved better results using a statistical classifier (SVM) compared to their own extended version of NegEx. Contextual features, including simple lexical information and more complex syntactic information, are extracted from the text and then used by the statistical classifier. A limited word window of 4 words either side of the target is used to limit the scope of the classifier. For example, the verb *showed* preceding a problem suggests that the condition is present, whereas *cured* after a problem suggests that the condition is absent. Uzuner et al. showed that their statistical classifier, StAC, can capture what they termed “assertion classes” on discharge summaries and radiology reports by making use of the information contained in the immediate context of target problems. Uzuner achieved F-value results of 98% for the positive class and 95% for the negative class using StAC, and 93% for the positive class and 90% for the negative class when using their extended version of NegEx.

Our work differs from those of related work in that we aim specifically at detecting the existence of *statistical evidence* by means of detecting specific negation phrases. Our difference from Niu et al. (2006)’s work is subtler. Whereas Niu et al. Focused on detecting outcomes that were beneficial (“positive”) or not (“negative”) for the patient, we focus on whether there are any outcomes at all, and, in subsequent work, we will look at detecting the direction of the outcomes. We believe that providing a first classification between outcome/no outcome before detecting its polarity has the potential of giving better results, since the two classifiers can focus on different types of information.

3 Method

3.1 Corpus Gathering

All medical research articles used in this research were sourced from PubMed Central. We selected a specific type of clinical study named Randomised Control Trial (RCT) since they are frequently used to report the results of clinical studies. RCTs are high-quality studies focusing on the generation of measurable outcomes. In a RCT the subjects are randomly allocated to one of two groups. In the “active” group, the subjects are given the treatment that is the object of study; in the “control” group, they are given a placebo. The outcome of a RCT would normally indicate whether there is a statistically significant difference between the results of the active group against the control group.

The research articles were gathered from PubMed Central and stored into a database. Since PubMed Central does not identify RCTs, we did a first pass using PubMed.

The entire procedure was as follows.⁵

⁴Observed evidence is not to be confused with statistical evidence. Our work focuses on the detection of statistical evidence.

⁵Note that the data were gathered around September 2010. The interfaces to PubMed and PubMedCentral may have changed since

1. Go to PubMed and visit the “Limits” section.⁶
2. Select “Published in the last 180 days” in the Dates list, select “Randomized Controlled Trial” as the type of article, select “English” as the language, and select “Links to Free Full-text” as the text option. Click “Search”.
3. When the results appear, click the link to “Free Full Text”.
4. Visually inspect the list of results to find those that are marked as completed RCTs.
5. Copy the PMID to the database and then click the “Free Text” link to open the article.
6. Identify the PMCID and the PICO details and copy them to the database (see below).
7. Save the record to the database.
8. Go to PubMed Central.⁷
9. Enter the PMCID in the search box, click the “Search all Articles” option, and then click “Search”.
10. Click the “Display” link and then select XML to change to display to XML format.
11. Click the “Send To” link and then select “File” to save the file.
12. Save the file using the PMCID as the filename. Also change the extension from .txt to .xml.

Step 6 involved the extraction of the PICO details. These include four key aspects of patient care (Gosall and Gosall, 2009):

1. **P**roblem or patient;
2. The main **I**ntervention, exposure, test or prognostic factor under consideration;
3. A **C**omparative intervention used in treatment; and
4. The **O**utcomes achieved or measured.

The PICO information was extracted by visual inspection of the text and was stored in the database. The purpose of keeping this information was to help the annotators understand the abstracts quicker.

3.2 Corpus Annotation

The abstracts of the RCTs were visually inspected by annotators to determine the type of evidence presented. Three annotators from two medical institutions⁸ were recruited as domain experts and the annotation was performed using a web-based annotation tool designed for this annotation task.

The abstract and the PICO details were uploaded to the annotation tool. The annotation tool was designed to allow an arbitrary number of domain experts to annotate the texts.

The annotators were instructed to examine the statistics reported in the abstracts in order to decide whether the research hypothesis had been rejected or not. In particular, if no difference is found between

	Accepted	Rejected	Total
(1) Training	66	61	127
(2) Test	33	34	67
(1)+(2) Total	99	95	194

Table 1: Dataset used for training and testing

the results of the active and the control group in a RCT, then there is no evidence to support that the intervention under study has any effect in the measured outcomes, and therefore the research hypothesis is deemed as rejected. If, however, statistical differences are found that are not due to chance alone, then the research hypothesis is accepted.

The annotators were therefore required to read the abstract, and then to select either “Accepted”, “Rejected”, or “Unknown” according to these criteria:

Accepted: A difference is reported between the intervention and the control group.

Rejected: No difference is reported.

Unknown: Unable to tell (e.g. because the RCT does not provide any results).

Figure 1 shows a screen-shot of the summary page presented to the annotators. The page contains the PICO information together with the result of the annotation. The annotators had access to the full abstract as shown in Figure 2.

Abstracts marked as “Unknown” were discarded for this study. The dataset included additional annotations, including annotations about whether there were secondary outcomes and their types. This information will be used in future studies but was discarded in the present study.

Whenever there was disagreement between annotators we asked them to review the articles. The annotators were not influenced to select any class or to change their original classification.

To measure the final agreement between annotators we computed Fleiss’ Kappa (κ). The Kappa statistic can be interpreted as expressing the extent to which the observed amount of agreement among annotators exceeds what would be expected if all annotators made their ratings completely randomly. If agreement is no more than expected by chance, then $\kappa = 0$. With perfect agreement, $\kappa=1$. The exact formula is

$$\kappa = (P_O - P_E)/(1 - P_E)$$

where P_O is the observed agreement, and P_E is the agreement expected by chance.

We found a Kappa value of 70.6%. This falls within the range of values that is usually termed as “good agreement beyond chance”. For the final dataset, we chose the decisions that corresponded to the majority of the annotators’ individual decisions.

Table 1 shows the numbers of articles used for training and testing for each type. We can observe that the ratio between evidence (“Accepted”) and no evidence (“Rejected”) is roughly equal, which approximates the ratio observed in our pilot studies.

3.3 Baselines

We ran several statistical classifiers using word-based features. We partitioned the corpus into a training set as shown in Table 1. The size of the corpus, though similar to that of related work such as by Niu et al. (2006), was small and we would expect better results

then.

⁶<http://www.ncbi.nlm.nih.gov/pubmed/limits>

⁷<http://www.ncbi.nlm.nih.gov/pmc/>

⁸Kolling Medical Research Institute and Royal North Shore Hospital, Sydney.

Outcome of Medical Intervention Survey - Mozilla Firefox

http://localhost/surveys/surveylist.php

MACQUARIE UNIVERSITY SYDNEY ~ AUSTRALIA

Survey
Change Password
Logout

Outcome of Medical Intervention Survey

TABLE: Survey

Search

Exact phrase All words Any word

Page 1 of 1 Records 1 to 2 of 2

	Survey Id	Pub Med Id (*)	Xml Filename (*)	Problem (*)	Intervention P (*)	Comparison P (*)	Outcome P (*)	Comment P (*)	Polarity P	Intervention S 1 (*)
1	19074218	BMJ-2-10-28-2769033	relatively inactive women over a two year period	effectiveness of a primary care based programme of exercise on prescription	usual activity	Physical activity assessed at baseline and 12 and 24 months	increased physical activity and quality of life	Accepted		
2	19858174	BMJ-2-10-27-2767482	enhance completion of treatment for tuberculosis	effectiveness of the provision of whole food	usual diet	Completion of treatment (including cure)	did not improve	Rejected		

Page 1 of 1 Records 1 to 2 of 2

Figure 1: Summary listing page

MACQUARIE UNIVERSITY SYDNEY ~ AUSTRALIA

Survey
Change Password
Logout

Outcome of Medical Intervention Survey

View TABLE: Survey

Back to List Add Edit Copy Delete

Page 1 of 2

Survey Id	1
Article	File No: 2505091
	PubMed Id: 18687718
Abstract	
Objective	To compare the effectiveness of clomifene citrate and unstimulated intrauterine insemination with expectant management for the treatment of unexplained infertility.
Design	Three arm parallel group, pragmatic randomised controlled trial.
Setting	Four teaching hospitals and a district general hospital in Scotland.
Participants	Couples with infertility for over two years, confirmed ovulation, patent fallopian tubes, and motile sperm.
Intervention	Expectant management, oral clomifene citrate, and unstimulated intrauterine insemination.
Main outcome measures	The primary outcome was live birth. Secondary outcome measures included clinical pregnancy, multiple pregnancy, miscarriage, and acceptability.
Results	580 women were randomised to expectant management (n=193), oral clomifene citrate (n=194), or unstimulated intrauterine insemination (n=193) for six months. The three randomised groups were comparable in terms of age, body mass index, duration of infertility, sperm concentration, and motility. Live birth rates were 32/193 (17%), 26/192 (14%), and 43/191 (23%), respectively. Compared with expectant management, the odds ratio for a live birth was 0.73 (95% confidence interval 0.45 to 1.18) after clomifene citrate and 1.46 (0.88 to 2.43) after unstimulated intrauterine insemination. More women randomised to clomifene citrate (159/170, 94%) and unstimulated intrauterine insemination (155/162, 96%) found the process of treatment acceptable than those randomised to expectant management (123/153, 80%) (P=0.001 and P=0.001, respectively).
Conclusion	In couples with unexplained infertility existing treatments such as empirical clomifene and unstimulated intrauterine insemination are unlikely to offer superior live birth rates compared with expectant management.
Pub Med Id	19074218
Xml Filename	BMJ-2-10-28-2769033
Problem	relatively inactive women over a two year period
Intervention P	effectiveness of a primary care based programme of exercise on prescription
Comparison P	usual activity
Outcome P	Physical activity assessed at baseline and 12 and 24 months
Comment P	increased physical activity and quality of life
Polarity P	Accepted

Page 1 of 2

©2010 Macquarie University. All rights reserved.

Figure 2: View annotations details page

with larger training data. Therefore we considered these classifiers only as baselines to improve. We used the following features:

1. All words in the abstract;
2. All words in the conclusion section;
3. Selected words in the abstract; and
4. Selected words in the conclusion section.

All abstracts were originally structured into sections, and, therefore, it was trivial to select the conclusion section. The selection of words was done by visual inspection of the training dataset. The following words were selected: *achieved, decrease, decreased, difference, effect, effective, effects, efficacy, improve, improvement, increase, increased, no, not, provide, provided, reduce, reduced, significant.*

The classifiers selected were decision tree (J48), support vector machine (SVM), and Naïve Bayes (NB). The results are shown in Table 2.

The best results were obtained by using the selected words in the conclusion section by the J48 classifier.

Note, however, the large confidence intervals⁹ that are due to the small size of the test set (67 documents). Still, even with these large confidence intervals we observe statistically significantly better results by focusing on the conclusion section versus using the complete abstracts. The difference between using all words or only a selection was not statistically significant and it would be interesting to repeat the classification with more data to determine whether the difference would become statistically significant.

3.4 Rule-based Classifier

Our rule-based classifier was based on NegEx, which was simplified as follows:

1. A different list of negation triggers was compiled by examining the frequency of triggers in the training dataset (see below).
2. The classes were changed from “affirmed” to “Accepted” and “negated” to “Rejected”. The class “possible” was removed.
3. The function that sorts the negation triggers and the function that uses regular expressions to match a set of negation triggers with text in each sentence were retained with very minor changes.
4. The functions in NegEx related to the detection and use of concept sentences, such as shortness of *breath, headache, chills, fever*, etc., were removed since we integrated the concepts related to evidence in the list of negation triggers.
5. The functions that tag conjunction and pseudo-negation were also removed.

Other minor modifications were made to make the system work with the dataset used for the experiments:

6. The output to the tagger was modified to return the PubMed report ID, the conclusion that was stripped from the abstract, the abstract itself, the current class, the tagged negation phrase when found, and the system class.

⁹Confidence intervals were based on a binomial distribution and were computed using R’s `Hmisc::binconf` function.

been overestimated, cannot endorse, cannot recommend, did not reduce, does not reduce, effectiveness overestimated, failed to, ineffective in, low probability, neither altered, no advantage, no advantageous, no beneficial, no benefit, no certain, no conclusive, no convincing, no definite, no detectable, no difference, no effect, no evidence, no favourable, no findings, no important, no improved, no increase, no irrefutable, no major, no meaningful, no more, no new, no novel, no overall benefit, no overall benefits, no overall effect, no positive, no proof, no reduction, no significant, no statistically, no strong, no substantial, no suggestion, nonsignificant improvement, nonsignificant improvement, nonsignificant reduction, non-significant reduction, nor protected, not affect, not appear to, not appreciate, not associated, not be, not beneficial, not change, not clinically, not confirm, not confirmed, not demonstrate, not differ, not exhibit, not find, not had, not have, not improve, not increase, not influence, not know, not known, not lead, not lend support, not likely, not meaningful, not meaningfully, not met, not necessarily, not observed, not offer, not prevent, not produce, not promote, not prove, not provide, not result, not reveal, not see, not show, not shown, not significant, not significantly, not slow, not statistically, not superior, not suppress, not to, not,, remains unproved, similarly effective, unlikely to

Table 3: List of negation phrases

7. A function was added to parse the XML files downloaded from PubMed Central. The function extracts the text of the abstract element and writes the PMCID and also the abstract to a CSV file that was used as input for the negation program. The annotated class is first entered into this CSV file before being used as input to the negation tagger.
8. A function was added to split the text of the abstract in order to separate the conclusion from the abstract.
9. The classifier incorporated functionality to output the results into a text file in the form of PMCID, Abstract, Conclusion, Tagged Sentence, Current Class, and System Class.

The following feature in NegEx has not been modified in the present version of the negation tagger:

10. The output that calculates the accuracy by comparing the current class with the class found by the classifier.

The list of negation triggers mentioned in item 1 are mostly bigrams and a few trigrams that were extracted by manually inspecting the training data. Phrases such as *no increase, no decrease, no significant, not improve*, and *not found*, often appear in the conclusion of articles where the hypothesis has been rejected. Some trigrams such as *not lend support, no major effect, no overall effect*, and *no overall benefit* also serve to negate the outcome of the intervention. The full list is shown in Table 3.

The rule-based classifier is designed to detect negation in the conclusion section only and then classify the article accordingly. The striking improvement of

	J48	SVM	NB
Baseline 1	49% (37%-61%)	66% (54%-76%)	69% (57%-79%)
Baseline 2	82% (71%-89%)	78% (67%-86%)	71% (59%-80%)
Baseline 3	54% (42%-65%)	63% (51%-73%)	58% (46%-69%)
Baseline 4	84% (73%-91%)	80% (69%-88%)	78% (67%-86%)

Table 2: Accuracy of the baseline classifiers with 95% confidence intervals

results obtained by restricting the analysis to the conclusion section by the statistical classifiers led us to this decision. Limiting the scope of a negation to the conclusion reduces the likelihood of false negatives, that is negation being detected and the article classed as rejected when in fact the research hypothesis has been accepted. All abstracts were structured and therefore it was trivial to select the conclusion section.

The rule-based classifier obtained an accuracy of 95% on the test set, with a 95% confidence interval between 88% and 98%. The better performance of the rule-based classifier with respect to the baselines is encouraging, though given the small amount of data used by our statistical classifiers we cannot rule out the possibility that statistical classifiers trained with more data would outperform the rule-based classifier.

An analysis of the classification errors indicated that often the error was due to the context of the negation. Even after selecting the conclusion section only, sometimes the negation detected referred to a secondary outcome rather than the main outcome. In other cases, the expression related to the evidence appeared in the results section but not in the conclusion section, and therefore it was not detected.

4 Conclusions

We have presented a rule-based classifier that detects whether the abstract of a published clinical RCT indicates whether their research hypothesis is confirmed. We do this by adapting NegEx' negation detector to focus on the detection of expressions of negation of evidence. We have simplified the original NegEx and replaced its original patterns with specific patterns derived from manual inspection of a training set. Experiments on a disjoint test set show an accuracy between 88% and 98% within a 95% confidence interval.

These results are statistically significantly better than those of a set of baselines using statistical classifiers. The amount of training data used in the statistical classifiers is comparable to that of related methods such as the one by Niu et al. (2006), though we believe that they are still too small to draw conclusions about the comparison between rule-based and statistical methods. We therefore plan to repeat the experiments with larger volumes of data. By training on larger volumes of data we expect better results using statistical classifiers. In the process, we will experiment with more complex features to feed the classifiers, including the patterns used in our rule-based system.

We also observed that the results obtained from processing only the words in the conclusion sections of the abstracts are better than those obtained using the complete text. We are considering the possibility of applying automated sentence classification techniques to detect conclusion sentences such as those devised by Demner-Fushman et al. (2006) and Kim et al. (2011).

We also plan to extend the analysis to more varied types of studies. Even though RCTs represent

a relatively large percentage of clinical studies, there are other important types of studies that should be considered, such as meta-analyses.

Further work also includes the detection of secondary outcomes in the papers. By detecting these we hope to reduce the scope of the negation expression and increase the overall accuracy results.

Finally, we plan to perform an extrinsic evaluation by integrating this research into an application system that returns relevant medical papers and ranks them by the quality of their clinical evidence. The research presented here will provide one additional item of information to this system. We will study the impact of this feature in the overall task.

Given that the list of negation triggers discovered in this study does not contain clinical terminology it is possible that this program with the current list of triggers would be useful to classify any experimental-based research paper with no or minor modifications.

Being a modification of NegEx, the classifier is written in Python and is available as open-source code.¹⁰

References

- Aronow, D. B., Fangfang, F. and Croft, W. B. (1999), 'Ad hoc classification of radiology reports', *Journal of the American Medical Information Association* **6**(5), 393–411.
- Chapman, W. W., Bridewell, W., Hanbury, P., Cooper, G. F. and Buchanan, B. G. (2001), 'A simple algorithm for identifying negated findings and diseases in discharge summaries', *Journal of Biomedical Informatics* **34**(5), 301–310.
- Demner-Fushman, D., Few, B., Hauser, S. E. and Thoma, G. (2006), 'Automatically identifying health outcome information in medline records.', *J Am Med Inform Assoc* **13**(1), 52–60.
- Goryachev, A., Sordo, M., Ngo, L. and Zeng, Q. (2006), Implementation and evaluation of four different methods of negation detection, Technical report, DSG.
- Gosall, N. and Gosall, G. (2009), *The Doctors' Guide to Critical Appraisal*, 2 edn, PasTest Ltd., Knutsford Cheshire.
- Huang, Y. and Lowe, H. J. (2007), 'A novel hybrid approach to automated negation detection in clinical radiology reports', *Journal of the American Medical Informatics Association* **14**(3), 304–311.
- Kim, S. N., Martinez, D., Cavedon, L. and Yencken, L. (2011), 'Automatic classification of sentences to support evidence based medicine', *BMC Bioinformatics* **12**(Suppl 2), S5.

¹⁰<http://sourceforge.net/p/clinevidence/>

- Meystre, S. M. and Haug, P. J. (2005), Comparing natural language processing tools to extract medical problems from narrative text, *in* 'AMIA Annual Symposium Proceedings', pp. 525–529.
- Mutalik, P. G., Deshpande, A. and Nadkarni, P. M. (2001), 'Use of general-purpose negation detection to augment concept indexing of medical documents', *Journal of the American Medical Informatics Association* **8**, 598–609.
- Niu, Y., Zhu, X. and Hirst, G. (2006), Using outcome polarity in sentence extraction for medical question-answering, *in* 'Proceedings AMIA'.
- Skeppstedt, M. (2010), Negation detection in swedish clinical text, *in* 'Proceedings NAACL HLT Second Louhi Workshop on Text and Data Mining of Health Documents', pp. 15–21.
- Uzuner, Ö., Zhang, X. and Sibandab, T. (2009), 'Machine learning and rule-based approaches to assertion classification', *Journal of the American Medical Informatics Association* **16**, 109–115.

Evaluation of Web 2.0 Technologies for Developing Online Telehealth Systems

Jaspaljeet Singh Dhillon, Czarina Ramos, Burkhard C. Wünsche, Christof Lutteroth

Department of Computer Science
University of Auckland
Private Bag 92019, Auckland, New Zealand

{jran055,cram021}@aucklanduni.ac.nz, {burkhard,lutteroth}@cs.auckland.ac.nz

Abstract

Telehealth and telecare applications are a promising technology for improving the quality of care while using healthcare resources more effectively. Major obstacles to a more widespread use are the high initial costs and a vendor specific design, which makes it difficult and expensive to add new functionalities. The Internet offers an opportunity to make telehealth applications more accessible, while also adding social aspects and the opportunity for third-party developers to add content. A preliminary user study confirmed that elderly are interested in such an application, and provided guidelines for the user interface design and required functionalities. In this paper, we evaluate technologies for developing online telehealth platforms, and present a first prototype which is extendable and has social networking capabilities. Our results show that a combination of open web standards such as OpenSocial and a CMS such as Drupal represents a suitable design. We illustrate the capabilities of our design and prototype by developing a memory game, which can be submitted by third party developers, similar to a Facebook application, and which utilises the social context of our telehealth application.

Keywords: Web 2.0 technologies, social networks, health informatics, telehealthcare.

1 Introduction

Telehealth and telecare systems are a promising approach to use healthcare resources more effectively. However, usage is constrained by high initial costs and a design often centered on the requirements of the clinical user, healthcare provider, and the equipment vendor. Most existing systems cannot be extended by third parties, require extra costs to add new functionalities, are designed to manage diseases rather than prevent them, and do not address the social and psychological needs of the patient.

A suitable concept to overcome these shortcomings is Web 2.0, commonly known as "the web as a platform" (O'Reilly 2005). The term refers to web applications and

services that facilitate interactive information sharing, rich user experience, dynamic content, and user-centred design. The open-ended nature, interconnectivity and large user community of popular social networks such as Facebook, MySpace and Orkut has enabled third-party developers to offer functionalities and content, which would be difficult to achieve with a stand-alone application.

The use of Web 2.0 in healthcare is rapidly evolving as more applications and services targeting health professionals and patients are being developed. With that trend, the term Health 2.0 is becoming popular with health management systems such as PatientsLikeMe, CureTogether, SugarStats and MyFitnesspal. Existing Health 2.0 applications provide useful functionalities such as diet and exercise monitoring, and formation of support groups. However, they do not offer a comprehensive suite of functionalities and do not replace traditional telehealth platforms (Dhillon et al. 2011).

The above analysis shows that current telehealth technologies represent two extremes. On the one side are very expensive vendor specific telehealth platforms. These are targeting patients with chronic diseases (e.g. diabetes, asthma, COPD, heart failure) requiring a high level of health support, which justifies the high costs of such systems. The systems are doctor centric and in many cases the patient is a passive provider of monitoring data and recipient of doctor advice, with few opportunities to participate in treatment and intervention plans. On the other side are Web 2.0 applications, which are mostly free and generally focus on one functionality only. These applications usually do not involve clinical users, and allow the patient to take control of their health, e.g. by devising diet plans or sharing with fellow patients experiences such as side effects of medications.

In this paper, we analyse different technologies for implementing web-based extendable telehealth systems with social networking capabilities. Specifically, we investigate the potential of popular social networking APIs and web development tools against the requirements for a general affordable telehealth system. We illustrate the results in the design of Healthcare4Life, our Health 2.0 platform, and a memory game application developed for it. We discuss the implementation process and relevant issues which will benefit the development of online telehealth systems in general.

Section 2 summarises important requirements for a general affordable telehealth system. Section 3 and Section 4 present an evaluation of popular Application Programming Interfaces (APIs) of social networking and web development tools for the development of telehealth

systems. Section 5 discusses the design and implementation of Healthcare4Life, a working prototype whose architecture and design is motivated by the requirements specified in Section 2. We also discuss the design and development of an example application (a memory game), and how it is integrated into Healthcare4Life. Section 6 describes the results of implementing the technologies, and Section 7 concludes the paper and gives an outlook on future work.

2 Requirement Analysis

Current telehealth systems do not take into account the importance of patients' social needs. In previous work we showed that social interactions are essential for patients, especially the elderly, to improve their quality of life and to overcome social isolation (Singh et al. 2010b). Social networks can help users to get in touch with their family, make new friends, and discuss medical concerns with peers and support groups. Furthermore, social networks also help with motivating the patient, e.g. by achieving family support, or by patients performing monitoring task and exercises together via a video link or in a virtual environment. Therefore, we suggest that novel telehealth systems should include social networking capabilities.

Existing telehealth systems are predominantly standalone applications, which come with a specific functionality, e.g. a device for measuring vital signs such as blood pressure, weight, pulse, and blood glucose levels. These applications cannot be extended by third parties and require the users to pay more to add new functionalities. In order to intervene early in the development of serious diseases, a larger proportion of the population needs access to telehealthcare services, and a wider range of functionalities must be provided. Examples are support for diet programs or physiotherapy exercises (Dhillon et al. 2011). This also reduces the risk that patients becoming bored with a limited range of content and functionalities. Moreover, for elderly people, it is often difficult to work with many different systems, and a single integrated user interface is necessary. Hence, telehealth systems should be designed with a plug-in architecture to enable third-party application developers to add content or health applications easily.

Current systems are mostly designed for patients to transmit health parameters to clinicians, i.e. do not encourage active participation of users in their healthcare. A general telehealth system should enable the patient to take control of their health. This requires tracking of health parameters (e.g. vital signs, exercise performance), feedback and alerts, and graphical representations to easily track progress and compare it with goals (Singh et al. 2010b). Lee et al. (2011) and Fischer et al. (2011) have shown that the use of patient specific visualizations can alter patient behavior and support rehabilitation.

One of the core requirements of telehealth systems is user-friendliness. The resulting requirements include the use of large font sizes, an easy-to-follow linear structure, and the use of a horizontal menu at the top of the screen to make it easy to identify and choose key functionalities (Dhillon et al. 2011).

It is essential that monitoring data can be shared between different applications. This will avoid the need

for users to re-enter data (e.g. patient parameters), and it allows implementation of more powerful functionalities (e.g. total calories burned or overall fitness). Such features will also help users to keep track of the total amount of time spent in performing the health related activities, and data can more easily be compared with other users to increase motivation.

Another important requirement of telehealth systems is the ability to integrate with consumer level HCI devices, e.g. iPhones and Wii remotes. These devices contain motion sensors such as accelerometers that can measure positions, velocity and direction vectors, which can be leveraged in creating rich applications, e.g. pedometers, fall detection, or guided rehabilitation activities to improve the condition of the user (Dhillon et al. 2011).

3 Evaluation of Social Networking APIs

Social networking APIs enable developers to integrate social features into their systems. The Facebook and OpenSocial APIs are the two most popular examples. In this section, we describe and compare these APIs for the development of telehealth systems based on the requirements discussed in Section 2.

3.1 OpenSocial

OpenSocial provides a set of common APIs for developing web-based solutions, with a focus on social applications. It is currently managed by the non-profit OpenSocial Foundation, is developed by Google along with MySpace, and is supported by a number of other social networks and well known software vendors such as IBM and SAP. The principle idea of OpenSocial is to make applications widely available to more users by enabling application developers to deploy the same application across multiple platforms with no or minimum modification. Nevertheless, developers are increasingly exploring OpenSocial for other development needs, moving from traditional social networking concepts to enterprise-level software.

OpenSocial allows the development of an open platform, also known as an OpenSocial container, where third-party developers can contribute applications written using the OpenSocial API. OpenSocial applications share the same structure as Google gadgets, therefore are also known as OpenSocial gadgets. These gadgets are actually XML documents containing HTML and JavaScript code along with metadata. There are two types of gadgets that can be built using OpenSocial: gadgets that live within the hosting container, and gadgets that rely on an external server. The latter is widely used in realising open platforms, where developers integrate XML specifications located on their own external web servers with the hosting OpenSocial container.

The contents of a gadget can be displayed in the different views supported by the container, e.g. profile, canvas, home and preview (Häsel 2011). Gadgets can be specified to switch between these views to enable the users to interact with applications in different sizes and layouts. Most containers support the canvas view, which displays the rendered gadget by itself in a full screen page within the container.

MySpace, Hi5 and Orkut are some of the popular OpenSocial containers that take advantage of the services provided by the API. Examples are methods to access information about people, friends, and data, within the context of a container.

To become an OpenSocial container that can render remote or embedded gadgets and support social networking features, a system must comply with both the Core Gadget Container Specification and Social Gadget Specification (OpenSocial 2011a). Developers can make use of Apache Shindig, a reference implementation of the OpenSocial standards, to host OpenSocial applications with little effort. It provides the code to render gadgets and proxy requests, as well as handle REST and RPC requests. Communications between the Apache Shindig and the application take place via standardised AJAX requests, defined in the OpenSocial JavaScript API (OpenSocial 2011d). Apache Shindig is currently written in both Java and PHP. The hosting process of the container is made possible through its four components: Gadget Container JavaScript, Gadget Rendering Server, OpenSocial Container JavaScript and OpenSocial Data Server (Shindig 2010). Apache Shindig also provides a variety of security level options to secure requests and responses, i.e. to enable developers to make applications more secure. It uses Shindig user security tokens, two and three way handshakes, OAuth, and various encryption technologies.

Any HTML page can be fetched and displayed in an OpenSocial container using a gadget mechanism called Proxied Content. This implies that developers can specify the Uniform Resource Identifier (URI) of any existing online application in the XML specification to turn it into a gadget that can be rendered by the container (OpenSocial 2011b). However, the outcome of wrapping existing applications in OpenSocial can be less rewarding, as these applications may not be designed to provide social interaction, unless the developer made significant use of the user's social context including friends lists and activity streams (Hinchcliffe 2011).

Although OpenSocial was not ready for productive use when it was launched in November 2007 (Schonfeld 2007), it is rapidly evolving with more improvements and significant features (OpenSocial 2011a). Recently, the OpenSocial Foundation has launched OpenSocial 2.0 (OpenSocial 2011c), which includes features such as embedded experiences, activity streams standardisation, support for mobile devices, OAuth 2.0 and OpenSearch support (Hinchcliffe 2011).

3.2 Facebook

Facebook is the most prominent social network and has nearly 700 million users worldwide (Eldon 2011). Similar to OpenSocial, the API allows applications to utilise profile, friend, photo, and event data to add social context. It also allows the applications to publish activities to the news feed and profile pages of Facebook. Increasingly, it is used by people and company sites as an identity provider with its support for OAuth 2.0. This avoids the need to register or create a new user account on each site individually. The large user base of the

Facebook attracts many third-party developers who build new products and services on this platform.

The API supports the RESTful API and the Graph API (Facebook 2011). The Facebook platform is based on a URL-addressable, RESTlike server API, i.e. it assigns unique IDs to each social object in the system, which can be invoked by a URL. OpenSocial gadgets are rendered by the surrounding container (e.g. Apache Shindig) and can communicate with their backend servers via JavaScript calls, whereas Facebook applications rest entirely on their developers' web server. In contrast to OpenSocial platforms, Facebook restricts developers to proprietary language requirements such as FBML (an evolved subset of HTML), FQL (an SQL-style interface for querying social data), and FBJS (a solution to enable developers to use JavaScript in their FBML applications). The engineering team of Facebook has released and maintains open source SDKs for Android, C#, iPhone, JavaScript, PHP, and Python (Facebook 2011).

The use of the Facebook API for telehealth systems makes it possible to access millions of users of this social network, including family and friends of a patient. According to Norval et al. (2011), this makes it easier to connect people known to a patient to provide care or social support. However, elderly users remain a clear minority and in 2010 only 2% of Facebook users were in the 65+ age bracket (Socialbakers 2011).

Despite the huge success of Facebook, it has been reported recently that traffic is dropping (Eldon 2011) and significant challenges exist to archive the users' personal data (McCown 2009). A growing number of Facebook users are switching over to Google+ due mainly to its apparent integration with a variety of Google services (Sullivan 2011). The latest statistics from SocialBakers (2011) show that the three top categories of applications in Facebook are Games, Entertainment and Lifestyle. Health applications (a sub-category of Lifestyle) has just over 1% of the total available applications. Therefore, although Facebook is known to be the most popular social networking site, it is not necessarily an appropriate platform for health related applications.

3.3 Discussion

Section 2 showed that two major requirements of telehealth systems are to make the systems extendable by third parties and to include social networking capabilities. There are two general approaches to realise such features: 1) creating a new API from scratch and sharing it with health application developers, and 2) deploying and adapting existing tools that support development of an open platform with social interactions. The former is more difficult because it takes time for an API to mature and to be accepted as a standard for developers. When leveraging existing Web 2.0 technologies to create open-ended systems (Dhillon et al. 2011), developers need to make a well-informed choice about the API used.

Social networking APIs are the core technologies needed to realise the open platform and social aspects of such a system. After reviewing the current social networking APIs, it is clear that both OpenSocial and Facebook allow people to keep up with friends, upload an unlimited number of photos, share links and videos, and

learn more about the people they meet. Both APIs support the development of third-party applications. However, the APIs were created with different objectives. Table 1 shows a comparison between the OpenSocialAPI and the Facebook API.

OpenSocial	Facebook
A specification	A social network
Standard API for social applications to run on multiple social networks	Single network API
Open with no proprietary regulations	Strict proprietary regulations
Applications hosted are commonly client-side JavaScript-oriented (gadgets) as well as server-oriented	Applications hosted are all server-oriented
Allows portability of an application into various OpenSocial containers	Applications can only run within the Facebook platform
Uses common languages (e.g. HTML, XML and JavaScript)	Uses proprietary languages (e.g. FBML, FQL and FBJS)
Full control over the social network functionalities and user policies	Little control over the social network functionalities and user policies

Table 1: OpenSocial versus Facebook

OpenSocial is not a technology but a specification. By following the OpenSocial specification any system can be turned into an open platform, which can interact with other applications in a standardised way. The platform and other applications will have a common set of interfaces and processes in order to communicate seamlessly. OpenSocial's social API can be leveraged to incorporate social networking features into a new system. In contrast to OpenSocial, Facebook is a social network and does not provide an open platform. It uses a plug-in architecture to enable developers to create applications, which can only run within the Facebook platform.

Developers using the Facebook API have little control over the social network functionalities and user policies (Norval et al. 2011). Unlike Facebook, OpenSocial also allows the development of web-based telehealth systems without constraining the user with proprietary regulations. Developers will have full control over their system and the freedom to integrate it with other OpenSocial containers. The ability to run applications on various containers will encourage potential developers to contribute health applications. However, OpenSocial allows this to happen only if the applications are programmed to be generic, and do not use their own proprietary API (Häsel 2011).

The reference implementation of OpenSocial, Apache Shindig, enables a telehealth system to be transformed into an OpenSocial container. Development of gadget applications is easy and attractive, since developers are not required to learn new programming languages and specific platform traits associated with its proprietary mechanism (McIlrath 2010). Instead, common languages such as HTML and JavaScript can be used. Developers can create application using a variety of technologies, including CSS, OpenSocial Templates, Flash, PHP, Python, Java, Perl, .NET, and Ruby.

The ability to easily embed existing health applications into an OpenSocial container is a great advantage for telehealth systems. Although these applications may have limited social features, developers will be able to embed proper existing health applications into the telehealth system. In addition, this will allow users, especially the elderly, to interact with existing health related applications within the same interface.

The idea of leveraging OpenSocial for telehealth systems was initiated by Weitzel et al. (2009), who described the use of this Web 2.0 technology in providing extended care networks for chronic disease management and elderly care. Furthermore, Weitzel et al. (2010) have discussed a Web 2.0 model for patient-centred health informatics applications. The suggested model uses open technologies such as OpenSocial, REST, and Open Authentication.

Based on the reviews and analysis, the OpenSocial APIs meet the requirements to develop online telehealth systems which are extendable and contain social aspect.

4 Web Development Tools

Web Development Tools are necessary to construct the web-based systems and its functionalities. The two common approaches are Content Management Systems (CMSs) and Web Development Frameworks (WDFs). In this section, we will highlight the strengths and weaknesses in their ability to design, develop and maintain web-based telehealth systems.

4.1 Content Management Systems

CMSs support developers with setting up rich and dynamic websites. With a CMS, the content is stored in a database and the templates, styles or themes that determine how the content is presented are maintained separately. Most CMSs leverage the power of Cascading Style Sheets (CSS) to easily update or make changes to the look and feel of websites. The main advantage of employing a CMS in developing a web-based system is the variety of ready-made modules, which can be directly used or adopted to add desired features to the website. CMSs can be integrated with existing social networking APIs. For example, Drupal has a module to integrate the OpenSocial API. Most CMSs use popular programming language such as PHP to enable developers to create their customised modules for specific features of their site. The three major open-source CMSs are Drupal, Joomla and WordPress (water&stone and CMSWire 2009).

Drupal is one of the most popular and powerful CMS available to develop dynamic state-of-the-art Web 2.0 sites. Drupal is supported by a large and active community of developers, and offers a large number of open source extensions, modules, and themes. It is based on a customisable framework which enables its site visitors to contribute content. Provided functions go beyond those of a CMS, as it also acts as a framework for developing web applications and is used in a wide variety of deployments. Drupal is increasingly used for developing social networking sites (Purham 2010) and healthcare systems, including sites connecting patients to health services (Drupal 2011a).

Joomla is another strong alternative for rich web development. It is easy to use, but most of the customisations required by the user are built around paid plug-ins and themes. It lacks important features such as a powerful blogging engine, nested categories, a built-in download manager/document repository, Content Construction Kit (CCK) abilities (functionality to easily move content around), and many other features already found in Drupal. WordPress has a strong focus on blogging, although a large number of open source plug-ins are available to extend its functionality. It is ideal for fairly simple blog-style web sites, but is not suitable for more complex site requirements.

4.2 Web Development Frameworks

Web Development Frameworks (WDFs) support developers with building websites, web applications and web services. There are many frameworks available for web development, written in various programming languages, with varying technical and conceptual differences (Singh 2010). For example, Yii, CodeIgniter, Zend, CakePHP and Symfony are just a few of the popular ones from the vast selection of PHP frameworks available to code web-based projects. Although each framework is different, they generally provide a variety of useful features. These frameworks provide functionalities that are common to most web applications, e.g. database access, sessions management and templating systems. WDFs help in providing a basic structure to develop web-based systems, which enables developers to reduce repetition and write code in a shorter amount of time. For instance, a framework enables developers to avoid the need to re-code the same features for each web application they create.

Most WDFs are based on the Model View Controller (MVC) architecture. The MVC implements a “separation of concerns”, i.e. distinct features without overlapping functionality. Examples include isolation of the application logic from the user interface and separation of database access code from the application logic. Separation of tasks, such as web programming from user interface design, allows a development team to focus on specific objectives and use their individual strengths (DocForge 2010). With MVC developers can focus and work on individual elements. Hence, the concept of MVC helps to break the development process of an application into manageable tasks.

4.3 Discussion

Using web development tools can considerably reduce development times, which is essential for complex incremental systems. The choice of using CMSs or WDFs depends on the complexity, requirements and duration of the telehealth project. Table 2 shows a comparison between CMS and WDF. Drupal offers more functionality over other available CMSs and is easier to use, but it has a steeper learning curve.

Although Drupal helps developers to create specific modules for features lacking in their system, it encompasses nearly 8500 contributed modules (Drupal 2011b), including modules to integrate HCI devices such as webcams (Drupal 2011d). Therefore it is likely that the

needed functionality is already available. As it was stated earlier, in order to host OpenSocial applications, Apache Shindig needs to be installed to render these applications. Drupal is preferred over other CMSs and WDFs because it includes the OpenSocial Shindig-Integrator, which can be used to integrate the Apache Shindig container with any Drupal-based system (Drupal 2011c). Hence, Drupal can be used to construct a full-featured and extensible web-based telehealth system that can host OpenSocial health applications.

Content Management Systems	Web Development Frameworks
Low learning curve	High learning curve
Easier system updates	Longer update and upgrade times
Planning is useful	Prior proper planning is essential
More consistent and controlled outcome	Lower quality control over the outcome
Provides back-end support such as modules and themes	Build from scratch using available classes and libraries
Suitable for small to medium projects, including system prototypes	Suitable for large and complex projects
Flexibility of adding customised modules	Unlimited flexibility
Suitable for projects with a small development team	Suitable for projects with a large development team
Reduce development times	Typically requires longer development time

Table 2: Content Management Systems versus Web Development Frameworks

As mentioned earlier, Drupal makes it easier to update a system. This is important for a telehealth system, since patient needs, available knowledge, and required interactions with other health providers can change. The module and theme feature help to set up web interface components. In addition, there are lots of user interface templates available to improve the look and feel of a system. Moreover, Drupal has a large and active user community, which is helpful during the development process.

In comparison to CMSs, WDFs frameworks are suitable for large and complex projects which require maximum flexibility. CMSs provides the user with back-end support (such as modules and themes) to develop and manage a website (front-end). By contrast, users deploying a WDFs framework have to build their sites from scratch by using the readily available classes and libraries. The initial learning curve for some of the WDFs can be quite steep and it requires the user to have solid knowledge of Object Oriented Programming (OOP). Other known shortcomings of employing a WDF such as a PHP framework include: 1) lower quality control over the outcome, 2) longer update and upgrade times, and 3) a stronger need for proper planning (Cheng 2009). Using CMSs makes system updates easier than when using WDFs. This is especially important if future content updates or changes will be done by non-technical users.

5 Design and Implementation

Section 2 summarised requirements for developing a web-based telehealth system, which addresses the

shortcomings of existing telehealth applications. In this section, we explain how we used Web 2.0 technologies in realising a working prototype of a web-based telehealth system called Healthcare4Life.

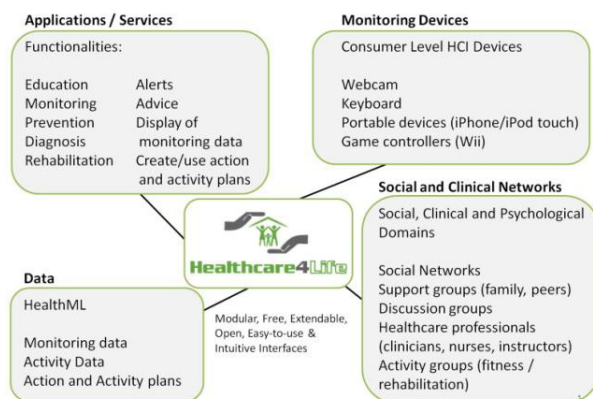


Figure 1: Framework for Healthcare4Life
(Singh, Wünsche and Lutteroth, 2010a)

In previous work, we presented a theoretical framework for Healthcare4Life (Figure 1), a novel web-based telehealth system that combines the power of social media with telehealth systems to enable patients to take charge of their own health (Singh et al. 2010a). Its goal is to transform the restricted nature of traditional telehealth systems by making them widely available, affordable and extendable. The framework has an open OpenSocial-like architecture, which allows third-party providers to add new content and functionalities. This allows users to choose new monitoring and exercise tools if they get tired of existing ones or develop new needs. It also makes it possible to incorporate emerging technologies such as new HCI devices. Functionalities include applications for monitoring, rehabilitation, education, and advice in the form of, e.g. serious games, interactive web pages and expert systems.

Our solution does not put any constraint on the physical location of users. Since we do not provide medical advice, we are not limited by government policies on this. System feedback will contain general health information and advice to contact a user's healthcare provider if unusual patterns in the monitoring data are detected.

Patient privacy is important and users are encouraged not to share clinical information. However, like with existing social networking and patient support websites it is ultimately up to the user to decide what information to store and share. For example, in order to find peers suffering from the same disease it is necessary to specify this information, and other members of such a patient group will hence implicitly gain this information.

Our technology evaluation has shown that OpenSocial and Drupal are suitable technologies to realise a web-based telehealth platform. Both technologies are open-sourced with a large and active community, and access to the resources and support for implementing our system. We therefore decided to use these two technologies for our system. The ideas presented by Weitzel et al. (2009 & 2010), along with our review on current social networking APIs, affirms the value of leveraging

OpenSocial in the development of Healthcare4Life. Since we are using the PHP version of Apache Shindig and Drupal, which is also written in PHP, we adopted PHP as our main programming language.

5.1 Container

The resulting system must support both application users and application developers. Both groups are presented with distinctive functionalities based on their role in the system. The application users or patients are presented with a clear horizontal menu at the top, with six icons and descriptive text: Home, Applications, Profile, Mail, Friends and Search (see Figure 3). Table 3 describes the main purpose of each page and respective Drupal modules used in implementing the features. The core functionality provided to application developers is to embed their gadgets applications into Healthcare4Life. All applications will be listed in the applications directory of the system for users to interact with them.

Page	Purpose	Drupal Module
Home	To share their status, view and comment status of friends within the network.	<i>Facebook-style Statuses</i>
Applications	All health applications added by developers will be listed at this page, as icons. Users are required to click on the respective icon to interact with a health application in canvas view.	<i>Shindig-Integrator</i>
Profile	To view and edit profile information. It will also consist of summary of latest activities such recent health applications used by the user.	<i>Content-profile and Flag</i>
Mail	To send a mail to friends within the network.	<i>Mail</i>
Friends	To access friends' profile page.	<i>Flag</i>
Search	To find new friends within the network.	<i>Search</i>

Table 3: Functionalities provided to application users

We started the implementation process by designing a customised theme for Healthcare4Life using HTML and CSS, which was imported into Drupal. Although Drupal comes with lots of design templates, we opted for a new theme based on the interface design requirements specified by elderly people (Dhillon et al. 2011). We then adapted existing Drupal modules for many of the functionalities of Healthcare4Life. For instance, we employed the Flag module for storing data, e.g. information of friends and activities, and Webforms module for developer and user registration. We also created customised modules for features not supported by the existing modules. For instance, Drupal does not support different types of registration for developers and normal users, which is required for Healthcare4Life. Social networking functionalities currently implemented include the ability to create profile pages, send email to others users, add friends and search for friends using specific keywords (e.g. username, age and hobbies).

Upon implementing the basic functionalities, we integrated OpenSocial with Drupal to transform Healthcare4Life into an open platform for third-party developers. Initially, we installed Apache Shindig within the Healthcare4Life environment and then integrated it with Drupal. Figure 2 depicts the architecture of our system.

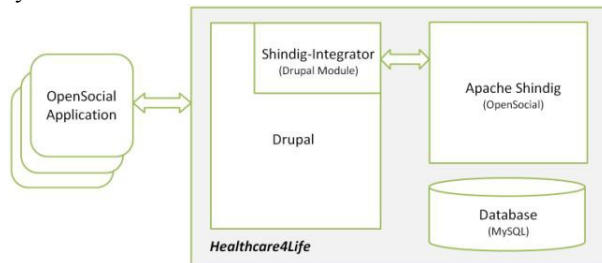


Figure 2: Architecture of Healthcare4Life

In order to make Healthcare4Life an OpenSocial compliant platform, the Shindig-Integrator module was used. This module originally used the User Relationships module for a user's friends' data, and the Profile module of Drupal to store a user's profile (Drupal 2011c). In Healthcare4Life, we have used different modules to achieve the same. Hence, to integrate the Shindig-Integrator with our platform, we had to adjust the code to use the Flag and Content-profile modules for user friends and profile, respectively. The Shindig-Integrator has a class which talks to the database for retrieving a user profile or user friends from the database based on the user ID. We changed the code, which was referring to the Profile and User-relation module's database, and made them retrieve data from the Flag and Content-profile module database.

With OpenSocial, developers are able to add applications residing on their own web servers by simply specifying their URL, i.e. the location of the XML file of the application, in our system. This XML specification of an application will be rendered by Apache Shindig and integrated into our system. The gadget applications added here are automatically listed in the applications directory of the site. We have tested this with existing applications from Labpixies (www.labpixies.com). Apart from specifying the URL of their application, application developers need to select a suitable category for their application. This helps users with selecting applications, e.g. for monitoring, education and rehabilitation.

We have also connected Healthcare4Life with Shindig's OpenSocial Service Provider Interface (SPI) to allow gadget applications to access our site's social data. The SPI implements: 1) retrieving people information, 2) storing and retrieving activities, 3) storing and retrieving persistent data, and 4) sending messages.

5.2 Sample Application

We have designed and developed a simple memory game in order to demonstrate the capabilities of OpenSocial's client side API. This involved transforming an existing JavaScript application into an OpenSocial application that can access the social context of our Healthcare4Life container. Applications that are appropriate for the Healthcare4Life platform must be beneficial to the user's

health. As mentioned earlier, applications plugged into the platform are grouped into specific categories based on their purpose. Rehabilitation is one of the application categories available in Healthcare4Life and will contain applications that aim to improve cognitive or physical functions.

The design and implementation of the new OpenSocial memory game was heavily based on a study by Ijsselsteijn et al. (2007) investigating the needs and motivation of elderly gamers. The original JavaScript memory game did not have social features that the elderly might find interesting, such as personalised, challenging and collaborative game play. In addition to this, the game design did not consider the special requirements of the elderly, e.g. readable fonts, larger images and familiar terminology. For testing purposes, our design aimed to make the new memory game challenging and personal to the user by allowing the application to access the personal data of the user's friends in the Healthcare4Life network.

The original JavaScript memory game is able to run in the OpenSocial container using the Proxied Content technique mentioned earlier. However, this does not provide the memory game with any interaction to the user's social data. The original game simply used cartoon images as objects of the game, which is monotonous and leads to the elderly being less motivated to use the game again. It became obvious that educating users on the benefit of the game to their health is not enough to motivate them to keep playing the game. To make the game more interesting, we have added a new "Card Deck" called "Friends". Terms such as "Card Deck", "Friends" and "Cartoon Images" are used because they are common terminology familiar to the elderly. Choosing the "Friends" game mode allows users to play using their friends' images as objects of the game. This makes the game more personal to the user, giving it a social aspect.

Figure 3 illustrates the memory game running in Healthcare4Life. Developers are free to use their own web servers in hosting their applications, however, we chose to host our OpenSocial application in the iGoogle gadget server and embed it in our platform under the rehabilitation category. The new memory game is an XML document with HTML and JavaScript bodies, much like the architecture of a Google gadget. The main difference is that the game uses OpenSocial to gain access to the social data of the Healthcare4Life container. Developers can use their favourite editor to create their OpenSocial gadget. When developing our memory game, we used the Eclipse IDE, as it provides a readily available OpenSocial plug-in for ease of testing and deployment of applications.

Our game consists of multiple functions which are mixed with calls to the OpenSocial API. Examples are a function for retrieving user's information, a function that randomises and loads the user's friend's thumbnail images into the game, and a function that saves the user's score into the network's persistent data storage. The OpenSocial REST API allows the application to retrieve information about the user by calling the HTTP GET request to our server, in this case, the list of friends they have in the Healthcare4Life network. Healthcare4Life allows users to save small amounts of data about a

particular user, such as the scores achieved in the memory game application, using the storage mechanism of OpenSocial called Persistence API. Likewise, with this mechanism, each of the players is able to load their scores from the last time they used the application. These persistent data are key/value pairs which only accept a string format. Non-string data values such as the scores in our memory game need to be converted into a string format before being stored and then parsed back to its original format. Players are able to compare their achievements with other players by posting on their friends' walls, which effectively adds more social aspects to the game.

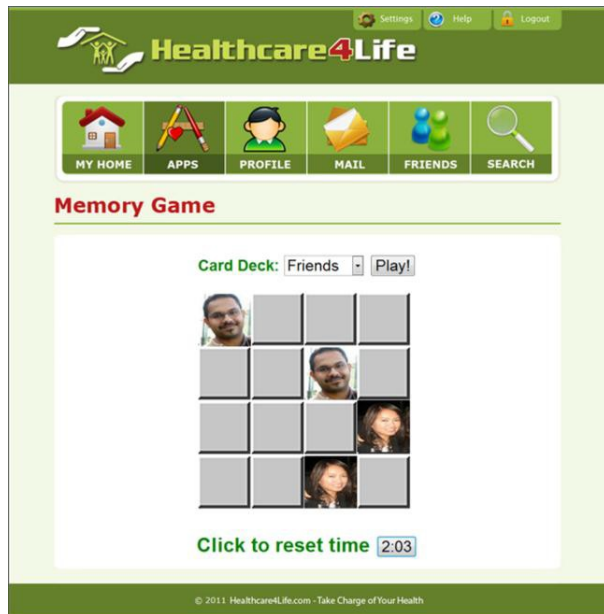


Figure 3: Memory game in *Healthcare4Life*

Healthcare4Life was able to render the game smoothly in canvas view and allowed access to its social data. Using standard HTML, JavaScript and CSS styling, we were able to change the old memory game into a more usable game with larger elements, a more elaborate interface, and more importantly, more interesting game play. The memory game was also tested in other OpenSocial containers, Orkut and iGoogle, which ran the game without requiring any complex modifications. There are a number of other possible social aspects that can be added to the memory game using the OpenSocial API, such as collaborative game play and scoreboards.

6 Results

Our evaluations of functionalities, and the subsequent design of a working prototype and a sample application, showed that OpenSocial and Drupal integrate well and can be used to develop extendable online telehealth systems with social capabilities. Based on our analysis, implementation experience and understanding, almost all of the requirements described in Section 2 can be achieved by leveraging these technologies. Below, we discuss the strengths and related issues in employing OpenSocial and Drupal.

6.1 OpenSocial

The availability of Apache Shindig, the reference implementation of OpenSocial, has helped to turn Healthcare4Life into an OpenSocial container that can host health applications created by external developers. Theoretically, OpenSocial compliant websites can be programmed using any programming language as long as the OpenSocial specifications are satisfied. In practice, however, the Apache Shindig will be used for most projects, which supports only Java and PHP.

It is fairly easy to develop gadget applications since common languages such as HTML and JavaScript are used. There is a lot of documentation available to get started with the development of a gadget application. However, it is difficult to find tutorials to develop rich and complex gadget applications, such as a multi-player memory game. Furthermore, applications must implement a valid version that specifies the features that we want the hosting container to interpret (OpenSocial 2011b). Issues with compliance arise when rendering applications implementing the 0.9.x version of the API, as the Shindig-Integrator of Drupal is only available in version 1.0. Also, it is found that designing a gadget too specific to a container requires more modification when running in other containers. Hence, the idea of running the same application across multiple OpenSocial containers is possible, but it is not as simple as documented.

OpenSocial provides various libraries to support a variety of applications and technologies. For instance, it provides a .NET client library, which enables the communication with the RESTful APIs of the OpenSocial container using Microsoft .NET based technologies (e.g. Kinect based applications). XML, HTML and JavaScript are also the foundations of Silverlight therefore conversion is easy. OpenSocial provides communications between Flash ActionScript and OpenSocial JavaScript API through its External Interface mechanism. Likewise, our investigation so far suggests that it is possible to develop gadget applications to achieve almost anything typically seen in healthcare related applications. For example, patient parameters (e.g. weight) can be stored in a database with the data acquisition date, and plotted using JavaScript and Google chart APIs. Note, however that in our experience, the development of new functionalities is time consuming due to the lack of suitable tutorials and documentation.

The OpenSocial specification does not say anything about data sharing (e.g. patient parameters) between two gadget applications. Since OpenSocial is a specification, it can be extended accordingly, i.e. we can write our own APIs to extend OpenSocial. Care must be taken when deciding whether functionality is provided by extending the OpenSocial API or by writing methods using a web development tool. For instance, sharing of patient parameters between applications can be achieved by using user profile data from CMS instead of using OpenSocial. When data is stored, it is stored in the platform like Healthcare4Life or Facebook. Security is always primary concern for the container and the applications it is hosting. OpenSocial is part of the container and depends on the services and data of the

platform. Therefore, the security aspects need to be implemented on the platform rather than on the applications.

OpenSocial enables developers to embed non-OpenSocial-based applications into containers. However, if an application such as a hand tracking application requires a special plug-in such as Silverlight, this application will not execute within the OpenSocial container without the plug-in. Therefore, it is necessary to develop a testing environment (also known as sandbox) that enables developers to test their gadget application prior to submitting them to appear in the application directory of a system.

OpenSocial specifies a common standard to share social data between two social networks and with OpenSocial applications. Potential developers of Healthcare4Life will be able to host their application on various OpenSocial based container such as MySpace, Orkut and Hi5. However, some of the biggest social networking sites like Facebook do not support OpenSocial.

OpenSocial is a flexible specification that can be treated as a blue print to design large scale enterprise applications, but it mainly focuses on social networking, social media and related entities. If developers are interested to make an enterprise system such as a telehealth system, they have to define their own API and specification according to their requirements if they are not already present in OpenSocial specification.

6.2 Drupal

Drupal reduces the development time for realising our system. The flexibility offered by Drupal and the availability of its contributed modules make it easier to implement functionalities. These modules can be used directly or modified to implement the desired features. Most of the modules are sufficiently documented to understand the source code and its implementation.

The contributed Shindig-Integrator module enables Drupal-based social networking sites to become OpenSocial compliant. However, the Shindig-Integrator module is old and not maintained properly. It depends on other modules, which often get upgraded. It is challenging to understand the Drupal architecture and Shindig-Integrator module code, in order to be able to change it as per Healthcare4Life. Furthermore, the Shindig-Integrator module only supports Apache Shindig release 1.0.x-incubating, i.e. it does not support OpenSocial 2.0. Therefore, we will not be able to integrate the new enterprise and consumer features provided by OpenSocial 2.0 unless the Shindig-Integrator module is upgraded or we invest more time to improve it ourselves.

Drupal has evolved as a significant application development tool, but it suffers from several limitations. Firstly, because it is open source, we cannot rely on open source modules or plug-ins in the long run, as they may not be maintained as is the case for the Shindig-Integrator. Another important issue is selecting the right module for a project. Drupal comes with almost 8500 contributed modules and many of them have similar functionalities. Some modules are easier to use than others, e.g. require the user to write less lines of code to

achieve a desired functionality (Buckman 2011). However, Drupal does not provide any guidelines to select the most suitable one. Users are expected to make their own selection by experimenting how each module fits into their project. Furthermore, Drupal's architecture is quite complicated; it takes a good amount of effort to write new and complex features of a system.

7 Conclusion

We have reviewed popular Web 2.0 technologies for developing web-based telehealth systems. We have specifically investigated the OpenSocial and Facebook APIs and a range of Web Development Tools. Based on our findings, both OpenSocial and Drupal can be used to develop an extensible and dynamic telehealth system. The availability of Apache Shindig (a reference implementation of OpenSocial) and the Shindig-Integrator module of Drupal make it easy to convert ordinary systems to be able to host gadget applications developed by external developers.

We have deployed these technologies to create an online telehealth platform called Healthcare4Life. Results of our deployment show that OpenSocial and Drupal can be integrated successfully to realise a working prototype. We have also developed a gadget application to test the platform. Although, some known minor issues persist, there are many advantages of leveraging such technologies for developing online telehealth solutions. Since our Healthcare4Life prototype is ready, we are all geared to develop more OpenSocial-based health applications and to integrate common HCI devices with the system. We look forward to evaluate our telehealth platform with real users.

8 References

- Buckman, B. (2011): Drupal's increasing complexity is becoming a turnoff for developers. New Leaf Digital's Developer Blog. <http://benbuckman.net/drupal-excessive-complexity>. Accessed 27 August 2011.
- Cheng, M. (2009): PHP CMS vs PHP Frameworks. Philosophy in Programming Society, Singapore. <http://www.scribd.com/doc/14288264/PHP-Meetup-Feb-09-PHP-CMS-vs-Frameworks>. Accessed 10 August 2011.
- Dhillon, J. S., Ramos, C., Wunsche, B. C. and Lutteroth, C. (2011): Designing a Web-based Telehealth System for Elderly People: An Interview Study in New Zealand. The 24th International Symposium on Computer-Based Medical Systems (CBMS 2011), Bristol, United Kingdom.
- Dhillon, J. S., Wunsche, B. C. and Lutteroth, C. (2011): Leveraging Web 2.0 and Consumer Devices for Improving Elderlies' Health. Australasian Workshop on Health Informatics and Knowledge Management (HIKM2011), Perth, Australia.
- DocForge (2010): Framework. DocForge - An Open Wiki For Software Developers, <http://docforge.com/wiki/Framework>. Accessed 24 August 2011.
- Drupal (2011a): Healthcare sites using Drupal. Groups.Drupal homepage, <http://groups.drupal.org/healthcare-sites>. Accessed 10 August 2011.
- Drupal (2011b): Modules. Drupal homepage <http://drupal.org/project/Modules>. Accessed 10 August 2011.

- Drupal (2011c): OpenSocial Shindig-Integrator. Drupal homepage, <http://drupal.org/project/ShindigIntegrator>. Accessed 10 August 2011.
- Drupal (2011d): Webcam trigger. Drupal homepage, http://drupal.org/project/webcam_trigger. Accessed 27 August 2011.
- Eldon E. (2011): Facebook Sees Big Traffic Drops in US and Canada as It Nears 700 Million Users Worldwide. Inside Facebook. <http://www.insidefacebook.com/2011/06/12/facebook-sees-big-traffic-drops-in-us-and-canada-as-it-nears-700-million-users-worldwide/>. Accessed 10 August 2011.
- Facebook (2011): Build the social and personalized web. Facebook Developers. <http://developers.facebook.com/>. Accessed 24 August 2011.
- Fischer, S., Wünsche, B. C., Cameron, L., Morunga, E. R., Parikh, U., Jago, L., Müller, S. (2011), Web-Based Visualisations Supporting Rehabilitation of Heart Failure Patients by Promoting Behavioural Change, Proc. of the 34th Australasian Computer Science Conference (ACSC2011), Perth, Australia, Mark Reynolds Eds., pp. 53-62.
- Häsel, M. (2011), "Opensocial: an enabler for social applications on the web", Communications of ACM, Jan 1 2011.
- Hinchcliffe, D. (2011): OpenSocial 2.0 - Will key new additions make it a primetime player in social apps? Enterprise Web 2.0. <http://www.zdnet.com/blog/hinchcliffe/opensocial-20-will-key-new-additions-make-it-a-prime-time-player-in-social-apps/1603>. Accessed 10 August 2011.
- Ijsselstein, W., Nap, H. H., de Kort, Y., Poels, K. (2007): Digital game design for elderly users. Proc. of Futureplay 2007, Toronto, Canada, 14-18 November 2007), pp. 17-22.
- Lee, T. J., Cameron, L., Wünsche, B. C., Stevens, C. (2011): A Randomised Trial of Computer-Based Communications Using Imagery and Text Information to Alter Representations of Heart Disease Risk and Motivate Protective Behaviour, British Journal of Health Psychology (16) 1, pp. 72-91.
- McCown, F., and Nelson, M. L. (2009): What happens when facebook is gone? Proc. of the 9th ACM/IEEE-CS Joint Conference on Digital Libraries, Austin, TX.
- McIlrath C. (2010): Getting Started with Facebook JS (FBJS). Thinkclay. <http://thinkclay.com/technology/getting-started-with-facebook-js-fbjs>. Accessed 24 August 2011.
- Norval, C., Arnott, J.A., et al. (2011): Purposeful Social Media as Support Platform - Communication Frameworks for Older Adults Requiring Care. ATTACH 2011 Workshop: Advances in Techniques and Technologies Assisting Care at Home, Dublin, Ireland.
- OpenSocial (2011a): OpenSocial Specification Release Notes. <http://opensocial-resources.googlecode.com/svn/spec/2.0/OpenSocial-Specification-Release-Notes.xml>. Accessed 10 August 2011.
- OpenSocial (2011b): OpenSocial Core Gadget Specification 2.0. <http://opensocial-resources.googlecode.com/svn/spec/2.0/Core-Gadget.xml#Versioning>. Accessed 10 August 2011.
- OpenSocial (2011c): OpenSocial 2.0. <http://docs.opensocial.org/display/OS/Home>. Accessed 10 August 2011.
- OpenSocial (2011d): OpenSocial API Reference. <http://code.google.com/apis/opensocial/docs/0.8/reference/>. Accessed 10 August 2011.
- O'Reilly, T. (2005): What Is Web 2.0. Design Patterns and Business Models for the Next Generation of Software. <http://oreilly.com/web2/archive/what-is-web-20.html?page=1>. Accessed 10 August 2011.
- Purham M. (2010): 30 Social Networking Websites That Were Created Using Drupal. Webdevtuts. <http://www.webdevtuts.net/inspiration/30-social-networking-websites-that-were-created-using-drupal/>. Accessed 10 August 2011.
- Schonfeld, E. (2007): OpenSocial Still "Not Open for Business". Techcrunch. <http://techcrunch.com/2007/12/06/opensocial-still-not-open-for-business>. Accessed 10 August 2011.
- Socialbakers (2011): Facebook in 2010: 7.9 new account registrations per second. <http://www.socialbakers.com/blog/109-facebook-in-2010-7-9-new-account-registrations-per-second/>. Accessed 10 August 2011.
- SocialBakers (2011): Health applications Facebook Statistics. <http://www.socialbakers.com/facebook-applications/category/105-health>. Accessed 10 August 2011.
- Shindig (2010): The Apache Software Foundation. Shindig - Welcome to Shindig! <http://shindig.apache.org/>. Accessed 10 August 2011.
- Singh J. (2010): The Best Web Development Frameworks. WebDesignish. <http://www.webdesignish.com/the-best-web-development-frameworks.html>. Accessed 24 August 2011.
- Singh, J., Wünsche, B. C. and Lutteroth, C. (2010a): Framework for Healthcare4Life - A Ubiquitous Patient-Centric Telehealth System. Proc. of CHINZ 2010, Auckland, New Zealand, pp. 41-48.
- Singh, J., Wünsche, B. C. and Lutteroth, C. (2010b): Taxonomy of Usability Requirements for Home Telehealth Systems. Proc. of CHINZ 2010, Auckland, New Zealand, pp. 29-32.
- Sullivan M. (2011): 9 Reasons to Switch from Facebook to Google+. PCWorld. http://www.pcworld.com/article/234825/9_reasons_to_switch_from_facebook_to_google.html. Accessed 10 August 2011.
- water&stone and CMSWire (2009): 2009 Open Source CMS Market Share Report. <http://www.cmswire.com/downloads/cms-market-share/>. Accessed 10 August 2011.
- Weitzel, M., Smith, A., Lee, D., Deugd, S., Helal, S. (2009): Participatory Medicine: Leveraging Social Networks in Telehealth Solutions. Lecture Notes in Computer Science, Series 5597: 40-47. Springer, Heidelberg.
- Weitzel, M., Smith, A., Lee, D., Deugd, S., Yates, R. (2010): A Web 2.0 Model for Patient-Centered Health Informatics Applications. Computer 43(7): 43-50.

Achieving Acceptable Structured eReferral Forms

YULONG GU, JIM WARREN, KAREN DAY, MALCOLM POLLOCK

National Institute for Health Innovation
University of Auckland
PO Box 92019, Auckland, New Zealand
h.gu@auckland.ac.nz

SUE WHITE

Whitehouse Associates, Health Information Management Consultancy
PO Box 12323, Wellington, New Zealand
sue.white@whitehouse.net.nz

Abstract

This paper reports on the implementation of electronic referrals (eReferrals) from community to public secondary services in New Zealand, with a focus on contrast and comparison of the knowledge engineering processes of two distinct regional projects with respect to eReferrals in the colorectal domain. The study data was gathered from project documentation, visits to key sites, analysis of electronic transactional records and stakeholder interviews. Both regional projects demonstrated effective processes of engaging hospital and general practice clinicians in developing investigation-specific eReferral forms: one with an iterative approach, and the other with a structured process. These negotiation processes among the participating clinicians have been effective in presenting, integrating and transferring specialised and locally contextualized knowledge. The resulting structured colorectal eReferral forms have shown sustained uptake and acceptance in both regions. The knowledge embedded in the forms clarifies referring criteria and collects appropriate information for referral triage. Involving both specialists and general practitioners in the form development and refinement has been identified as a key to success in both cases. In conclusion, it is critical to facilitate a negotiation process among secondary and community clinicians in order to achieve acceptable eReferral forms.

Keywords: computer communication networks; electronic referral; knowledge engineering; professional practice.

1 Introduction

Medical referral from a general practitioner (GP) to a specialist is probably the most obvious and familiar kind of referral to a New Zealand health consumer (as compared to, for instance, an administrative referral between two services within a hospital). The key reasons for such GP-to-specialist referrals are for diagnosis or investigation (e.g. special tests), treatment, and reassurance of GP and/or patient. It has been found,

however, that specialists are not generally satisfied with the letters from GPs because some letters do not include enough information: (1) to adequately address the problem; (2) about the reason for consultation, socio-psychological factors, or plans for follow up; or (3) on clinical findings, test results and details of prior treatment (Piterman and Koritsas, 2005, Gandhi et al., 2000, Jenkins, 1993, Williams and Peet, 1994, Tattersall et al., 2002). Overall, the most important information specialists need from GPs regards the problem to be addressed, the clinical questions to be answered, details the patient is unable or unlikely to provide, medical problems and medications (Gandhi et al., 2000).

Research has been conducted into improving GP-to-specialist communication through the use of form letters, that is, a structured or standardised referral letter. One study reported improvements in the quality of referral letters after introduction of a form letter (Couper and Henbest, 1996). It presents a type of knowledge engineering challenge, however, to convey the key criteria for clinical triage assessment using a structured form that GPs will accept. Knowledge engineering, in general, refers to the integration of knowledge into computer systems to solve complex problems normally requiring a high level of human expertise (Feigenbaum and McCorduck, 1983). The knowledge engineering task in developing electronic referral (eReferral) forms appears to focus on the representation of specialists' knowledge about what information is required from the GP in order to adequately address the problem and support a clinical decision.

New Zealand has one of the highest bowel (colorectal) cancer death rates in the world, with a five-year survival rate of 60.8 percent in 2002-2006 (Statistics New Zealand, 2009). Moreover, the investigation of colorectal symptoms such as large bowel symptoms in New Zealand primary care is fraught with a number of issues, including multiple clinical pathways and long waiting lists, motivating exploration of structured eReferrals to improve clinical quality in this domain (Davis, 2009).

eReferrals technology is one of the underpinning technologies in The New Zealand National Health Information Technology (IT) Plan which proposes an improved and rationalised health IT infrastructure that will ultimately support a transformed and more sustainable healthcare system (National Health IT Board, 2010). eReferral systems for GP-to-specialist referral have been evaluated on a small scale in the United Kingdom, where

they were found to improve demographic content of referrals but worsen clinical content (Shaw and de Berker, 2007). We have evaluated four eReferral implementations during 2010-2011: the systems already in operation at Hutt Valley, Northland and Canterbury District Health Boards; and the Auckland Metro region's solution (entering pilot operation at the time of reporting). The Hutt Valley (Warren et al., 2011d), Northland (Day et al., 2011), and Canterbury (Warren et al., 2011b) findings individually, and the results of the four evaluations collectively (Warren et al., 2011a), are available through the Health Innovation Exchange website (the HIVE, <http://hive.org.nz/>).

The current paper concentrates on the learning from the Northland solution and the Canterbury referral management innovations (the Canterbury Initiative, CI) using the example of referral to investigation of colorectal symptoms. Our findings in Northland and Canterbury are compared in terms of their approaches, uptake and acceptance to shed light on the nature of an appropriate process for creation of acceptable eReferral forms.

2 Methodology

Evaluation data was gathered from September 2010 to May 2011 through collection of project documentation, visits to key sites, analysis of electronic transactional records and stakeholder interviews. The overall study protocol was approved by the Multiregional Ethics Committee, approval MEC/10/066/EXP. The qualitative and quantitative study design was based on our evaluation framework for innovative health IT initiatives (Warren et al., 2011c), eReferrals requirements by Health Information Strategy Advisory Board (HISAC) (HISAC, 2005) and the criteria pool utilised by Lau and colleagues for health information systems evaluation (Lau et al., 2007). Quantitative data such as electronic transactional records and system access logs were analysed to assess the uptake and usage pattern of eReferrals as well as the impact of system implementation. Qualitative data was also collected to capture the „voices“ of those involved in and impacted by the innovations, including clinical, operational and management stakeholders. Additional information was gained by review of Northland eReferrals evaluation studies previously conducted (Davis, 2009, Davis, 2011). Using these data, for each project (Northland and Canterbury), we examine the development approach, the knowledge engineering products, and the evidence of uptake and acceptance.

3 Results

3.1 Northland's eReferral Project

Northland grew their project organically on a relatively low budget. The Northland eReferral solution, which was implemented in 2009 and continues to be iteratively developed as part of an ongoing clinical quality improvement initiative, encompasses GPs from 36 general practices referring patients to 29 services at the main 250-bed public hospital.

3.1.1 An Iterative Approach

Origin of the Northland project lies in an attempt by a general surgeon to resolve the issue of inadequate referral information for colorectal problems by collecting more

and better focussed data from referring GPs. The surgeon drafted a scoring referral form in 2008 to address the problem of lack of information in terms of content (what data is required for the specialist to be able to act) and format (the best data structure for easy access to the content). It was suggested this draft be developed as an electronic form attached to the primary care practice management system (PMS). GPs were then invited in an iterative process of development, refinement, and pilot implementation of the colorectal eReferral form which was based on the surgeon's draft. Dialogue continued among specialists, GPs, technology vendors, and management regarding the form design and review in face-to-face meetings, group emails, and the pilot project communications. At the end of this process, the form was released to the wider Northland GP community, achieving immediate acceptance and uptake.

Northland's colorectal form and their other well accepted structured eReferral forms, succeeded from an iterative approach that was often initiated by hospital specialists then involved primary clinicians in later cycles. Learning from the experience with colorectal form development that demonstrated the importance of involving primary care providers, the Northland project team has recognized and implemented open dialogue involving the GP users. The project leader commented, "One of the lessons we have learnt which is valuable is that involving ALL people who work in the workflow (both clinical and non-clinical) is important to ensure the mechanisms of the developed pathway are feasible."

Specialists were enthusiastic to design structured referral forms in conjunction with primary health; they were also actively involved in form training at „road shows“. The analytical capacity around system design brought in by clinical leaders is highly regarded. The specialists expressed a need to provide some form of guidelines for GPs for appropriate referral content, e.g. asking specific questions to prompt the inclusion of certain data. Then input from both specialists and GPs was captured in an iterative process of eReferral design and refinement through open communication. One specialist related a story about creating what he thought of as a well-designed and informative form for patients referred to him. When he met with some GPs to show them the structure, they presented an entirely different perspective that resulted in a very different final version of the form. GP involvement from the design phase of eReferral forms appears essential to the form acceptance.

The Northland experience demonstrates their ability to iteratively refine eReferral forms and for process improvement. The development and implementation of all their forms, and the associated new business processes, occurred as a series of reflective, self-correcting cycles. Each step was analysed carefully for strengths and weaknesses and adjustments were made to the next step of the project. This resulted in a high degree of innovative thinking, but unexpected software development and iterative implementation as each component of the project became ready for use.

3.1.2 Northland's Colorectal Form

Figure 1 is a screenshot of the clinical section of Northland's colorectal eReferral form („Surgery –

Colorectal Referral for Outpatient Appointment”), consisting of three key assessment criteria: symptoms,

family history, and personal history. Figure 2-4 show the expanded views once each criterion is met; these pages

The screenshot shows the 'Surgery - Colorectal Referral for Outpatient Appointment' form in the HealthLink Online system. The form is for Northland District Health Board. The 'Clinical Referral Information' section is currently set to 'No symptoms, No history'. Below this, there are checkboxes for 'Has symptoms?', 'Has Family History?', 'Has Personal History?', and 'Follow-up colonoscopy request?'. There is also a text area for 'Other Information' and 'Comments:'. At the bottom, there are sections for 'Medical History', 'Medications / Allergies', 'Patient Disabilities / Difficulties', 'Diagnostic Reports', and 'Referrer Details', all of which are currently set to 'No medical history specified', 'No long term medications, No allergies/alerts', 'No disabilities specified, No interpreter required', 'No reports selected, 18 available', and 'Sam Entwistle, Millstone Family Practice, NZMC 188984' respectively.

Figure 1: Northland’s colorectal form – clinical section: unexpanded

The screenshot shows the same form as Figure 1, but with the 'Clinical Referral Information' section expanded to 'Has symptoms, No history'. The 'Has symptoms?' checkbox is now checked. Below this, the form is divided into two columns: 'High Risk Symptoms' and 'Low Risk Symptoms'. Each column has a table with 'Has' and 'Criteria' columns. The 'High Risk Symptoms' table includes criteria such as 'Intraluminal rectal mass palpable/visible', 'Unexplained Fe deficiency anaemia with Hb less than 110 Male or Hb less than 100 Female', 'Right sided abdominal mass palpable', 'Positive/concerning imaging', 'Patient is older than 60 changed bowel habit to increased frequency/looseness, more than 6 weeks', 'Patient is older than 60, rectal bleeding with no anal symptoms, > 6 weeks', and 'Patient is between 40 and 60, rectal bleeding and changed bowel habit to increased frequency/looseness, more than 6 weeks'. The 'Low Risk Symptoms' table includes criteria such as 'Changed Bowel habit to constipation in absence of any other alarm symptoms (unresponsive to diet/simple measures)', 'Mucus/blood PR or symptoms suggestive of Inflammatory bowel disease', 'Abdominal pain/bloating with no sinister large bowel disease', 'PR Bleeding with anal symptoms', 'Unexplained weight loss', and 'Other'. There is also a text area for 'Additional relevant clinical information'.

Figure 2: Northland’s colorectal form – clinical section: ‘Symptoms’ expanded

HealthLink Online

Surgery - Colorectal Referral for Outpatient Appointment [Submit](#) [Park](#) [Referral Info](#) [Action](#) [Help](#)

Northland District Health Board

Administrative Details: NDHB, Outpatient Appointment, Non NZ Resident

Patient Information: MICKEY MOUSE, 66yrs, NHI JDR1234

Clinical Referral Information: No symptoms, Has History

Has symptoms? ☐

Has Family History? ☒

Has	Relationship	Number	Age of youngest at diagnosis	Comments
<input type="checkbox"/>	1st Degree			
<input type="checkbox"/>	2nd Degree			

Has Personal History? ☐

Follow-up colonoscopy request? ☐

Other Information

Comments:

Medical History: No medical history specified

Medications / Allergies: No long term medications, No allergies/alerts

[Print](#) [OK](#) [Cancel](#) [Help](#)

Figure 3: Northland's colorectal form – clinical section: 'Family history' expanded

HealthLink Online

Surgery - Colorectal Referral for Outpatient Appointment [Submit](#) [Park](#) [Referral Info](#) [Action](#) [Help](#)

Northland District Health Board

Administrative Details: NDHB, Outpatient Appointment, Non NZ Resident

Patient Information: MICKEY MOUSE, 66yrs, NHI JDR1234

Clinical Referral Information: No symptoms, Has History

Has symptoms? ☐

Has Family History? ☐

Has Personal History? ☒

Had	Condition	Year of diagnosis	Details
<input type="checkbox"/>	Polyps		
<input type="checkbox"/>	Inflammatory bowel disease		
<input type="checkbox"/>	Previous bowel cancer		
<input type="checkbox"/>	Other		

Inflammatory bowel disease

[Print](#) [OK](#) [Cancel](#) [Help](#)

Figure 4: Northland's colorectal form – clinical section: 'Personal history' expanded

collect the most valuable information that specialists and GPs agreed as appropriate and adequate in the referral. The interface design endeavours to capture these key data for clinical triage decision making in a quick and easy fashion for the referrers. The information required includes high risk symptoms and low risk symptoms as a list of tick boxes, and the relevant family and personal histories with tick boxes complemented by free text.

The form is integrated with Medtech Global's Medtech32 Practice Management System (PMS), the market leading product for GPs in New Zealand. The eReferral form is auto-populated with relevant PMS data such as medical history (including long term conditions), medications, and allergies, reducing the data entry burden.

3.1.3 eReferral Acceptance

After its initial iterative development cycles, Northland's colorectal eReferral form achieved steady and sustained uptake. The number of colorectal eReferrals received each month since its implementation in 2009 is recorded in Figure 5. (Note the two major downturns can be attributed to the Christmas / summer holiday period; similar dips were observed for total eReferral volume at Hutt Valley (Warren et al., 2011d)). A survey of 41 GPs shortly after the introduction of the colorectal eReferral form found 57% uptake and universal support for the improved service pathway, as well as 88% support that the form has made the referral task at least somewhat faster (Davis, 2009).

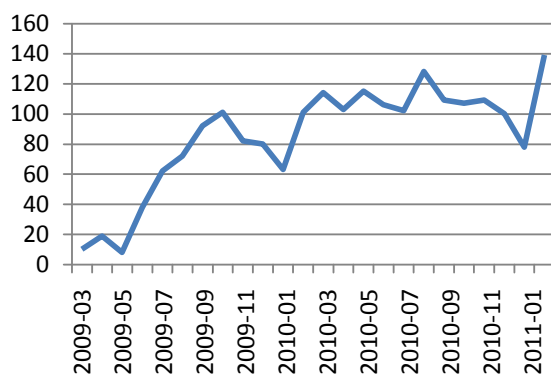


Figure 5: Northland's colorectal eReferrals monthly volume

The content of the colorectal form has shown evidence of significant improvement in terms of the adequacy of referral information leading to more accurate prioritisation and patient quality improvement (Davis, 2009). The evolution of structured colorectal eReferral is such that it provides a knowledge framework to referring GPs so that they can draw their own conclusions about the urgency and complexity of the colorectal problems they refer. As stated by the surgeon, "we've structured the colorectal form so that the GPs do a kind of triage of their own because of the data they insert and how they're asked for it. This is very helpful for both them and us." GPs appeared to find this helpful, as one related, "in situations like a colorectal referral or in a diabetes referral, and in selected other referrals as well, the receiving specialists have identified for us the key bits of information that they require in our referrals. And other fields are automatically

brought up for us to fill in as we're doing the eReferral. So the effect is that the receiving specialists are getting higher quality information more reliably through eReferrals than they would have been otherwise."

The implementation of eReferrals in Northland has promoted and facilitated continuous improvement of referral quality. By the end of February 2011, eReferral has been widely taken up by Northland GPs using six structured forms (Colorectal, Breast, Diabetes-retinal, Diabetes-pregnancy, Diabetes-general, and Acute) and one generic form. Northland specialists have reported that these forms contain "good quality transfer of care data." "The information presentation has changed substantially." "[The structured eReferral form] allows us to address the problem." and "This was a huge opportunity to get GPs to include information that they don't normally think about when writing a referral but that's important to us."

3.1.4 Ongoing Improvement

Northland are also becoming increasingly aware of some unintended negative consequences with customisation in structured forms that requires further investigation to learn and avoid replicating these problems across more forms and over wider areas. For instance, the structured data fields offer GPs the ability to use them as prompts, aiming at completeness that satisfies both GP and specialist. But comments from both GPs and specialists about the structured fields reflected a concern that data becomes inaccurate when a GP has a particular word or phrase in mind and ends up selecting the „best fit“ from the structured fields. Northland has discovered that free text is invaluable and forms need to support the inclusion of the patient story. As put by the project leader, "we have identified that we need to support more free text in general but specifically in relation to the inflammatory bowel referral option as ticking this box gives insufficient information to prioritise.... This is a problem with using symptom complexes rather than binary or limited choice structured answers. We are looking at simple education in the first instance to combat this problem." The resulting lack of clarity potentially puts patient safety at risk, something both sets of doctors are acutely aware of. There is a sense that the structured data "puts words in my mouth" when the narrative would have sufficed in presenting the patient's case. The colorectal form has been in use for three years and the Northland experience has identified weaknesses that could be improved upon, but there is a financial burden to this. The Northland project leader related, "While we are trying to be evidence based in medicine in our clinical interventions, there is a worrying blindness to taking this approach with the „tools“ we are adopting."

3.2 Canterbury Initiative

The „Canterbury Initiative“ (CI) is a healthcare transformation programme that subsumes referral management innovations for Canterbury District Health Board and is notable for its 300 HealthPathways (as of May 2011). These pathways are local agreements between GPs and specialists on the criteria, procedures and faxable templates for appropriate referrals to public secondary services. Extended from these paper-based referral templates, CI's eReferral solution became operational in

July 2010, providing GPs with standardised online referral forms with data pre-population from their desktop PMS (Medtech32).

3.2.1 A Structured Process

CI has established a structured HealthPathway definition process that consists of a maximum of five 90-minute evening meetings, where GPs and specialists have „robust“ discussion regarding the issues, requirements and workflows relevant to the management, assessment, and referral for a condition. These discussions are both informative and conclusive for iterative drafting of the pathways until agreement is reached. The CI facilitator described this process as follows, “at the initial meeting issues and opportunities are identified – a blank white board session. The actions required to address the issues are agreed and assigned. The second meeting doesn’t occur until actions have been progressed. This provides confidence to group members that they are investing their time in activity that actually delivers change. The same applies to subsequent meetings – a maximum of five meetings, which are all in the evenings. All clinicians are paid to attend these hour and a half long meetings. The process takes from 6 to 12 months.” This series of meetings, at its core, is a platform to enable negotiation between GPs and specialists regarding the local way of working with the resources that are available.

By engaging both secondary and primary clinicians, the HealthPathway authoring process, not just the direct product itself, is viewed as part of the reason for CI’s success, particularly because of the trust and relationships it builds. The CI facilitator related, “over time individuals and groups gain confidence and begin to trust each other – realise they are all working within constrained resources and that they need to work together to get the best outcome for patients – the whole of system solution. General practitioners engaged with hospital colleagues in work groups and through education create an environment that enables change. Relationships provide the vehicle to progress.” CI also emphasizes regular „Information Evenings“, with continuing medical education (CME) points awarded for participation, to introduce pathway innovations by face-to-face communication with GPs and get community feedback for further improvement.

3.2.2 Canterbury’s Colorectal Pathway

CI’s „colorectal symptom pathway“ is accessible to Canterbury area health providers via the HealthPathways website (<http://healthpathways.org.nz/>¹) under „Gastroenterology“ or „General Surgery“ or through direct search. This pathway employs quantitative scoring (as shown in Table 1) in its „colorectal symptom flow chart“ which gives the threshold scores for referrals to secondary public services and other management options (see Figure 6).

The „Gastroenterology / Endoscopy CT Colonography / Colonoscopy Diagnostic Request Referral“ form is indicated for if the required score is met, and is available both as a paper-based template and as an eReferral form (a

Symptom	Score
Rectal Bleeding (> 6 weeks)	
Sinister or Outlet	12.5
	5
Change in Bowel Habit (> 6 weeks) Choose one only	
Loose	10
Constipation	5
Weight Loss (> 5 kg)	5
Examination Findings	
Abdominal Mass	20
PR Mass	20
Bloods	
Unexplained Iron Def Anaemia	20
FOB positive (with any symptoms above)	10
If Diarrhoea/Loose Motions	
CRP > 10	10
Family History CRC	
Cat 1	0
Cat 2	5
Cat 3	10
Personal History Adenoma	2.5
Personal History CRC	5
Age	
>=60	10
40 - 59	0
<40	-5

Table 1: The scoring criteria of CI’s ‘colorectal symptom pathway’

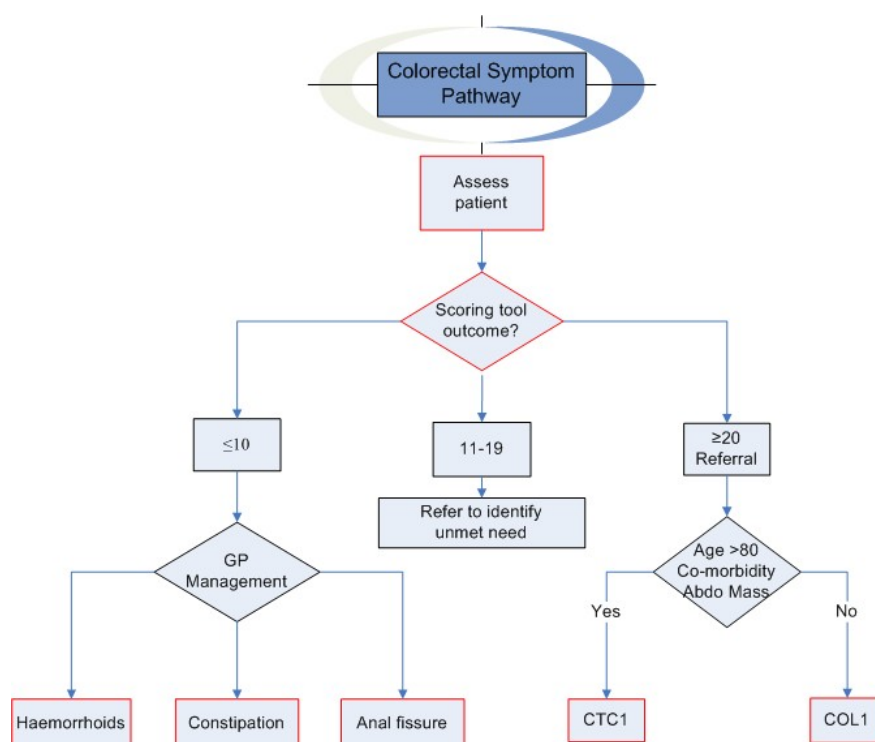
screenshot of the clinical section on electronic version is shown in Figure 7). The interface design applies the same principle as in Northland to minimize referrer’s effort by dropdown buttons and tick boxes.

3.2.3 Canterbury Uptake

The information in CI pathways and the associated referral forms guides GPs in collating assessment, examination, and investigation findings. Figure 8 shows the sustained access to the „colorectal symptom pathway“ pages during a 12-month period from 1 August 2010 to 31 July 2011. These pages were viewed a total of 2,351 times with clear pattern of use during working days (i.e. on a case-by-case basis when doctors are seeing patients, not during out-of-hours times that might be used for discretionary up-skilling).

A sense of improved appropriateness of all referrals as a result of pathway implementation has been repeatedly mentioned in Canterbury interviews. It was described as „right cases with right information most of the time.“ This perceived appropriateness has resulted in fewer declines. As stated by a referral triager, “in terms of just the request for a CTC [computed tomographic colonography], a simple answer is no: we decline very very few, because very very few inappropriate referrals come in anymore.” CI’s approach has transformed care delivery in Canterbury in terms of reducing demand for secondary public services by providing explicit scoring to prioritise the use of the limited resources for publicly funded colonoscopies, with the scoring serving to dissuade referral in cases that will fall below the public service threshold. One of the CI leadership team members stated, “This tool is used for resource allocation in a constrained resource environment and some patients do not receive treatment meeting recommended guidelines.”

¹ CI’s HealthPathways is password protected. CI has indicated that readers may request temporary access by emailing HealthPathways.



The red boxes are clickable; e.g., „Scoring tool outcome?“ links to a referral form with scoring criteria (see Table 1).

Figure 6: Screenshot of CI’s ‘colorectal symptom flow chart’ on the ‘colorectal symptom pathway’

Condition Specific Information Requirements		
(Please specify any conditions (and severity) which may affect the patients ability to tolerate sedation and/or bowel preparation). Tick to answer Yes.		
Artificial heart valve	<input type="checkbox"/>	
On Warfarin or Clopidogrel	<input type="checkbox"/>	
Too frail to tolerate sedation and/or bowel preparation	<input type="checkbox"/>	
<u>Symptoms</u>		Score
Rectal bleeding (>6 weeks)	--- select ---	0.00
Change in bowel habit (>6 weeks)	--- select ---	0.00
Weight loss (>5 kg)	<input type="checkbox"/>	0.00
<u>Examination Findings</u>		
Abdominal mass	<input type="checkbox"/>	0.00
PR mass	<input type="checkbox"/>	
<u>Lab Tests</u>		
Unexplained iron deficiency anaemia	<input type="checkbox"/>	0.00
Enter the value for these tests		
Hb	<input type="text"/>	
MCV	<input type="text"/>	
Ferritin	<input type="text"/>	
FOB positive & no overt bleeding (with any symptoms above)	<input type="checkbox"/>	0.00
If diarrhoea/loose motions CRP > 10	<input type="checkbox"/>	0.00
Family history CRC	--- select --- Help ?	0.00
Personal history adenoma	<input type="checkbox"/>	0.00
Personal history CRC	<input type="checkbox"/>	0.00
Age	< 40	-5.00
Total Score		-5.00

Figure 7: Screenshot of CI’s ‘Gastroenterology / Endoscopy CT Colonography / Colonoscopy Diagnostic Request Referral’ eReferral form – clinical section

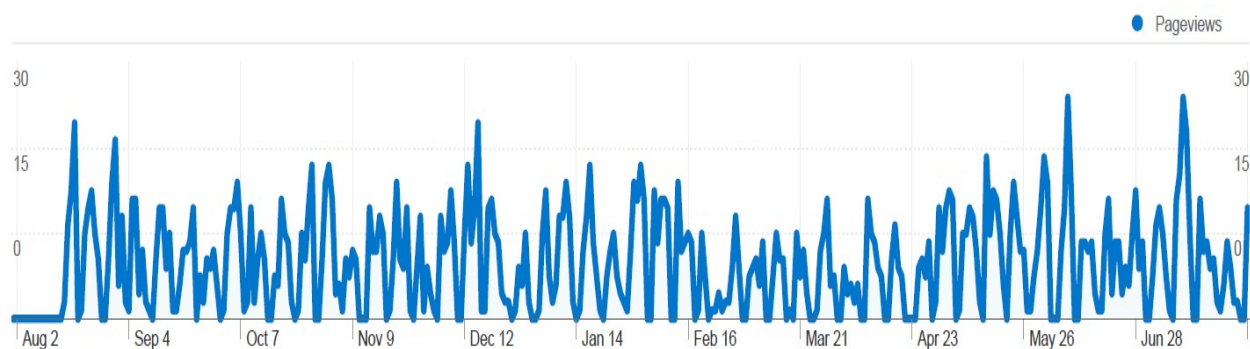


Figure 8: Daily access volume for CI's 'colorectal symptom pathway' pages (1 August 2010 to 31 July 2011)

3.2.4 Ongoing Improvement

While developed with careful consideration of the best evidence, as well as local workflow and demand, the scoring system and thresholds implemented (as per Table 1 and Figures 6 and 7) are as yet unvalidated in the clinical setting. A research study is underway to assess clinical limitations and potential patient risks. The will inform further refinement of the pathway.

It is not yet clear how many pathways and pathway-based referral forms will be required to support all areas of health service delivery in Canterbury where there is stakeholder demand. Established pathways will need maintenance, updating as well as refinement over time. For instance, one interview participant mentioned that the current information on a department pathway is far from adequate. Nevertheless, the content management solution is scaling well, with good enforcement of consistency in presentation that aids readability and gives reuse of successful design components, as well as ongoing enhancement of service-wide features such as online search. Moreover, the HealthPathways website provides a „feedback“ button for users. And the scheduled review for each pathway is at one year after implementation and every two years after that. The review panel includes a clinical editor, clinical director, and the subject expert. One of the CI leadership team members related, “we look at which pathways are the most popular and if there’s any problems appearing because the referrals aren’t being managed – they can’t manage the demand for example, then we can review the pathway and see if there’s any areas that general practitioners could do more of. So we’ll look at them from a resource allocation aspect as well as the evidence base.”

4 Discussion

Organizational knowledge management theories have emphasized knowledge processing, including the codification, retention, integration, coordination, transfer, and sharing of knowledge (Alavi and Leidner, 2001, Argote et al., 2003, Sambamurthy et al., 2003). The negotiation process between specialists and GPs that occurred in both Northland and Canterbury has served as the mechanism that enables knowledge processing, particularly the knowledge codification, integration, and transfer. It is evident in both settings that there is a temptation for some specialists to attempt to bridge the knowledge gap by using structured (electronic) referral forms with decision support tools, such as quantitative

scoring, to contextualise the patient’s condition in the discipline to which the patient is being referred. This works well with the colorectal forms in both regions, but it was not simply a process of codifying the specialists’ knowledge. Rather, their success lies in the negotiation processes undertaken that have been prolific in presenting, integrating and transferring the specialised and locally contextualized knowledge. The result encodes the negotiation, and incorporates both GP and specialist knowledge. Northland started with a knowledge engineering attempt to convert/model specialist’s knowledge into structured referral form; but it soon turned into a negotiation process through GP involvement in form design that is critical to the form acceptance. Subsequently, both specialists and GPs were engaged in an iterative and effective process of form development, refinement and implementation. Similarly, Canterbury’s systematic approach of facilitating negotiation between primary and secondary clinicians aimed for and achieved local agreements on the criteria and procedures for appropriate referrals to public secondary services.

In the referral context, specialists and GPs are two main knowledge sources and users in the „network of relationships“, which according to social capital theory and knowledge management theory is where the collective knowledge is stored (Lindsey, 2002). The engagement of both GPs and specialists during eReferral form or HealthPathway design not only facilitated knowledge processing through their negotiation on the local way of working; it also enhanced the network of relationships among the healthcare providers across the traditional primary-secondary boundary.

In both Northland and Canterbury, condition- or investigation-specific eReferral forms (explicit pathways as well as forms in the CI case) are the result of robust discussion and negotiation between secondary and primary clinicians on the content, format, issues, requirements and workflows regarding referring a patient to a particular service. These knowledge-embedded structured forms clarify the referring criteria and collect appropriate information for referral triage. However, eReferrals should not be viewed as just a „project“ that comes to an end at a set date; it must be an ongoing initiative, with ongoing capacity, because the technology and evidence are constantly changing in medicine. For instance, with respect to evidence, a recent systematic review concluded that the common practice of performing colonoscopies to identify cancers in people with bowel

symptoms is warranted only for rectal bleeding and the general symptom of weight loss, and not for some other commonly considered relevant symptoms such as change in bowel habit (Adelstein et al., 2011); this agrees with general direction of precedence of symptoms in both regions" forms, but the new evidence may be grounds for further differentiating the relevance of the symptoms. Furthermore, it has been identified that unintended negative consequences with customisation in structured forms require further investigation, e.g. when a GP has to select a „best fit" from the structured description of a breast lump and as a result potentially puts patient safety at risk. CI's scheduled review of each pathway appears highly appropriate; along with this protocol review they also examine the demand management from resource allocation as well as evidence-based aspects, which should be emulated in order to achieve substantial transformation of health delivery. In brief, the maintaining, reviewing, and refining of condition- or investigation- specific forms and protocols should be ongoing in order to fully realise the benefits of such tools and minimize unexpected negative consequences. It is well documented that a range of unintended negative consequences can result from introduction of IT in healthcare, allowing clinical information systems to cause their own novel (iatrogenic) errors (Ash et al., 2004) – eReferral is potentially vulnerable to this downside.

Because New Zealand enjoys a high-level of sophistication in PMS usage in its general practice sector (Schoen et al., 2009), eReferrals also offer the opportunity for data to be automatically supplied from the PMS database into the referral form (auto-population) which is realised in both Northland and Canterbury eReferral solutions. Through these features, eReferrals offer the potential for a degree of transformation in health delivery towards a more coherent interface between referring and referred to services – notably, across the GP-specialist and community-hospital boundaries – as compared to an array of freeform posted letters and point-to-point faxes. Moreover, eReferrals can provide a „hook" for electronic decision support and, potentially, for an IT-mediated social network among the stakeholders in the health of the referred patient.

The present study has a number of limitations. It is just the result of two case studies. Despite this limitation, the uptake data are compelling, and difficult to explain in any terms other than end user acceptance (at least by a broad sub-population of the area GPs). The structured eReferral forms in both cases have provided clarification for referral criteria and protocols. This has effectively bridged common barriers to successful coordination of eReferrals such as lack of an institutional referral policy, lack of standardization in referral procedures, and ambiguity in roles and responsibilities, as pointed out by (Hysong et al., 2011). The engagement of end users, i.e. clinicians from primary and secondary settings, in a negotiation process regarding local referral-related knowledge is a key factor for the eReferral success at Northland and Canterbury. This is consistent to eReferral experience that requires collaborating actors – clinicians on both sides of the referral process – to understand each other's needs and work processes (Heimly, 2010), and consistent with the general health information system literature that

emphasises the importance of having developers as users, to integrate decision support and benchmark practices, and to address such contextual issues as provider knowledge and perception (Lau et al., 2010). The present paper has not emphasized quantitative evidence regarding the impact on health outcomes from the implementation of structured eReferral forms beyond the evidence of uptake per se. We have, however, presented interview data, including many direct quotations, to illustrate that the tenor of feedback indicates that the health impacts are positive and promising. These provide pointers for where to direct additional quantitative measurements.

Canterbury provides a fascinating contrast to Northland in the emphasis on explicit pathways. eReferrals were only introduced to CI in July 2010; and their period of introduction coincides with the major earthquake events that have been highly disruptive to the region. The strong uptake of pathways, generally transformed to structured referrals by faxed paper forms rather than eReferral from the GPs software, gives an indication that the negotiation and dissemination of its product is the key element, not the online population of data into a structured form per se.

Last but not least, the issue of form design and development needs to be elevated to a national level, with a library of forms, including information on their usage and related evaluation findings, being maintained for access by new eReferral implementers. One of the biggest barriers might be the lack of a central process for form evolution/iteration that respects and utilises the expertise and experience learnt from implementations such as in Northland and Canterbury. A library per se might not solve this issue. There is much to be gained from the negotiation process being local, in terms of team building, and in terms of tailoring to local preferences and limitations. It is notable that the Northland and CI forms have many differences as well as similarities. Some differences will reflect useful differences in locality appropriateness. However, the differences also serve as a reminder that these forms and protocols, while having passed the initial gates of user acceptance and uptake, are in need of deeper validation for clinical effectiveness and safety. In either case, the opportunity to share learning between regional projects should be developed for its potential to shorten the journey to more widespread and successful evidence-based structured clinical communications.

5 Conclusion and Recommendations

Structured colorectal electronic referral forms have been found in two New Zealand regional implementations to meet with widespread and sustained uptake by clinical users. Such forms, and associated clinical protocols, provide a foundation for clarified referring criteria, more appropriate information for clinical triage and transformation in health service delivery. Experience from both projects demonstrated effective processes of engaging hospital and general practice clinicians in developing investigation- or condition-specific structured forms: one with an iterative approach, and the other with a structured process. The Canterbury case is interesting for emphasising widespread acceptance of online protocols as the primary objective, with computer-based input and transfer of data as secondary. The innovations of both regions show that whatever specific approach is taken, the

importance of user engagement, iterative refinement, and feedback must be emphasized.

Any health innovation process that is knowledge intensive should include elements such as:

- Engagement – It is very easy to leave out the concerns of stakeholders. It is natural for one group to have „blind spots“ regarding the needs of another. GPs and specialists do not have a natural empathy for each other's needs; CI's systematic process of workshopping together in five focused 90-minute sessions for each pathway is worth emulating in similar innovations.
- Iteration – The system will not be perfect on the first effort; revision should be planned. Northland has demonstrated a successful case of iteratively refining and improving the structured forms.

In brief, the processes of engaging both specialists and GPs can effectively achieve presentation and transfer of specialised and locally contextualized knowledge, as well as enhancing the network of relationships among the providers.

6 Acknowledgments

The eReferrals evaluation has been commissioned by the National Health IT Board of New Zealand. We'd also like to record our gratitude for the cooperation of all the people who actively assisted in this review, particularly Dr Dianne Davis, Ian Anderson, Bruce Penny, Dr David Kerr, Dr Graham McGeoch, and Carolyn Gullery.

7 References

- Adelstein, B.-A., Macaskill, P., Chan, S., Katelaris, P. & Irwig, L. (2011) Most bowel cancer symptoms do not indicate colorectal cancer and polyps: a systematic review. *BMC Gastroenterology*, 11, 65.
- Alavi, M. & Leidner, D.E. (2001) Review: Knowledge Management and Knowledge Management Systems: Conceptual Foundations and Research Issues. *MIS Quarterly*, 25, 107-136.
- Argote, L., Mcevily, B. & Reagans, R. (2003) Managing Knowledge in Organizations: An Integrative Framework and Review of Emerging Themes. *Management Science*, 49, 571-582.
- Ash, J.S., Berg, M. & Coiera, E. (2004) Some unintended consequences of information technology in health care: the nature of patient care information system-related errors. *J Am Med Inform Assoc*, 11, 104-12.
- Couper, I.D. & Henbest, R.J. (1996) The quality and relationship of referral and reply letters. The effect of introducing a pro forma letter. *S Afr Med J*, 86, 1540-2.
- Davis, D. (2009) Single Point of Entry Colorectal Service with a Supporting e-Referral Form. Northland District Health Board.
- Davis, D. (2011) Results of an audit of 50 Diabetes General electronic Referral forms. NDHB.
- Day, K., Gu, Y., Warren, J., White, S. & Pollock, M. (2011) National eReferral Evaluation: Findings for Northland District Health Board. Wellington: , Ministry of Health.
- Feigenbaum, E.A. & Mccorduck, P. (1983) *The fifth generation*, Reading, MA, Addison-Wesley.
- Gandhi, T.K., Sittig, D.F., Franklin, M., Sussman, A.J., Fairchild, D.G. & Bates, D.W. (2000) Communication breakdown in the outpatient referral process. *J Gen Intern Med*, 15, 626-31.
- Heimly, V. (2010) Collaboration across organizational borders, the referral case. *Stud Health Technol Inform*, 157, 106-11.
- Hisac (2005) eReferrals (<http://www.hisac.govt.nz/moh.nsf/indexcm/hisac2-actionzone-ereferrals#> last updated 22 May 2009, accessed 16 Feb 2011).
- Hysong, S.J., Esquivel, A., Sittig, D.F., Paul, L.A., Espadas, D., Singh, S. & Singh, H. (2011) Towards successful coordination of electronic health record based-referrals: a qualitative analysis. *Implement Sci*, 6, 84.
- Jenkins, R.M. (1993) Quality of general practitioner referrals to outpatient departments: assessment by specialists and a general practitioner. *Br J Gen Pract*, 43, 111-3.
- Lau, F., Hagens, S. & Muttitt, S. (2007) A proposed benefits evaluation framework for health information systems in Canada. *Healthc Q*, 10, 112-6, 118.
- Lau, F., Kuziemy, C., Price, M. & Gardner, J. (2010) A review on systematic reviews of health information system studies. *J Am Med Inform Assoc*, 17, 637-45.
- Lindsey, K. (2002) Measuring Knowledge Management Effectiveness: A Task-Contingent Organizational Capabilities Perspective. *The Eighth Americas Conference on Information Systems*.
- National Health It Board (2010) National Health IT Plan: Enabling an integrated healthcare model. Wellington (<http://www.ithealthboard.health.nz/content/national-health-it-plan>).
- Piterman, L. & Koritsas, S. (2005) Part II. General practitioner-specialist referral process. *Intern Med J*, 35, 491-6.
- Sambamurthy, V., Bharadwaj, A. & Grover, V. (2003) Shaping Agility Through Digital Options: Reconceptualizing the Role of Information Technology in Firms. *MIS Quarterly*, 27, 237-263.
- Schoen, C., Osborn, R., Doty, M.M., Squires, D., Peugh, J. & Applebaum, S. (2009) A survey of primary care physicians in eleven countries, 2009: perspectives on care, costs, and experiences. *Health Aff (Millwood)*, 28, w1171-83.
- Shaw, L.J. & De Berker, D.A. (2007) Strengths and weaknesses of electronic referral: comparison of data content and clinical value of electronic and paper referrals in dermatology. *Br J Gen Pract*, 57, 223-4.
- Statistics New Zealand (2009) Measuring New Zealand's Progress Using a Sustainable Development Approach: 2008. Wellington, Statistics New Zealand.
- Tattersall, M.H., Butow, P.N., Brown, J.E. & Thompson, J.F. (2002) Improving doctors' letters. *Med J Aust*, 177, 516-20.
- Warren, J., Gu, Y., White, S., Day, K. & Pollock, M. (2011a) Evaluating NZ Health Innovation Projects: eReferrals Evaluation. Wellington, Ministry of Health.
- Warren, J., Gu, Y., White, S., Day, K. & Pollock, M. (2011b) National eReferral Evaluation: Findings for the Canterbury Initiative. Wellington, Ministry of Health.
- Warren, J., Pollock, M., White, S. & Day, K. (2011c) Health IT Evaluation Framework. Wellington, Ministry of Health.
- Warren, J., White, S., Day, K. & Pollock, M. (2011d) National eReferral Evaluation: Findings for Hutt Valley District Health Board. Wellington, Ministry of Health.
- Williams, P.T. & Peet, G. (1994) Differences in the value of clinical information: referring physicians versus consulting specialists. *J Am Board Fam Pract*, 7, 292-302.

A Bayesian Analysis on Historical Clinical Data Concerning Treatment Change for HIV/AIDS Patients

KHAIRUL A. KASMIRAN¹ ALBERT Y. ZOMAYA¹
ALI AL MAZARI² ROGER J. GARSIA³

¹ School of Information Technologies
The University of Sydney, Australia

kkas4050@uni.sydney.edu.au, albert.zomaya@sydney.edu.au

² PSCJ Campus

Alfaisal University, Saudi Arabia

almazari@pscj.edu.sa

³ Department of Clinical Immunology

Royal Prince Alfred Hospital, Australia

roger.garsia@email.cs.nsw.gov.au

Abstract

The advances of biomedical informatics have allowed considerable amounts of historical data to be recorded about an HIV/AIDS patient. By a Bayesian approach coupled with a statistical hierarchical analysis, we evaluate the significance of historical viral load and CD4 data on the HIV treatment change decision. We show that, while there is correlation between the decision and the historical data up to 3 years before the decision, historical viral load data at 12, 18 and 24 months before the decision and historical CD4 data at 30 and 36 months before the decision provide little additional predictive power to what is provided by more recent data. Our results suggest that in a resource-constrained situation, an abridged historical viral load and CD4 archive can provide comparable prognostic support to a fully-populated archive with significantly reduced storage costs.

Keywords: HIV historical clinical data, Bayesian analysis, Information storage

1 Introduction

Despite dramatic advances in medical science, Acquired Immunodeficiency Syndrome (AIDS) still claims around 1.8 million deaths annually (UNAIDS, 2010). Its cause is the Human Immunodeficiency Virus (HIV) and the final outcome of chronic infection is usually death from a complication of the immunodeficiency or its treatment.

Modern multi-agent antiviral treatments can slow down HIV significantly, allowing substantial immunological recovery and prolonging the lives of HIV/AIDS patients. Since there are still no options that can eradicate HIV completely from a patient, they need to be regularly monitored for signs of disease progression.

A considerable number of clinical indicators have been documented and studied to assess their suitability for predicting disease progression in HIV-infected patients (Langford et al., 2007). They include quantitative assessment of the subtypes of T-cells, in particular CD4-bearing T-cells (CD4) and the quantitation of blood levels of HIV-1 viral load (VL) which are well-established as important predictors (DHHS, 2011). CD4 T-cells form a critical part of the human immune system and are the main target for the HIV-1 virus as the virus uses the CD4 molecule on their surface as part of the entry pathway into cells. Early on in the studies of AIDS and HIV, it was noted that a lower CD4 count and/or a higher VL level indicate a greater risk for disease progression in cohort studies and this information has been widely promulgated and used in guiding treatment recommendations (Phillips and Pezzotti, 2004, Goujard et al., 2006).

We study the CD4 and VL indicators from a different perspective – whether historical records of the indicators are predictive of a treatment change. While computerization allows for an abundance of such records and presentation to the clinician in various formats, including graphically longitudinally, the usefulness of comprehensive longitudinal data remains in question. Simply put, decision making with regard to treatment change may be dominated by recent data rather than by contextualizing recent data in the longitudinal historical data. In this paper, we provide an evaluation of the significance of historical CD4 and VL data towards the treatment change decision. The results can be used in informing design of decision support systems for

clinicians as well as informing storage prioritization in a resource-constrained situation.

We perform the evaluation via a novel Bayesian approach coupled with a statistical hierarchical analysis. We employ this approach since it can provide estimates of the direct, indirect and total effects of each historical CD4 and VL variable. The ratios of indirect effects with respect to direct effects are our major concern here, since a high ratio indicates a variable that has most of its influence mediated through other variables.

The evaluation is performed on an HIV/AIDS patient data set from the Royal Prince Alfred Hospital, Sydney, Australia. Our results suggest that an abridged historical archive without CD4 data from 12, 18 and 24 months and VL data from 30 and 36 months prior to the current date contains a similar amount of information compared to a complete archive but with the advantage of lower storage costs.

2 Related Work

To the best of our knowledge, no previous work has studied the relevance of historical CD4 and VL data to the treatment decision using Bayesian statistics, but there are several lines of related work.

There have been surveys that examine the HIV clinician's decision to initiate HIV treatment. Bogart et al. (2001) did an influential study that aimed to determine how gender, disease severity, ethnicity and risk group influence physicians' treatment decisions and judgment of treatment adherence. The study found that in a North American setting, former injection drug users, African American men and patients with less severe disease are judged as less likely to adhere to treatment, and that perceived adherence and disease severity has an effect on treatment decisions. Bogart et al. (2000) also did a survey to determine how 3 medical factors (e.g. CD4 cell count) and 17 nonmedical factors (e.g. patient homelessness) influence prescribing of highly active antiretroviral therapy (HAART) by physicians to HIV/AIDS patients. Their conclusion was that most physicians follow guidelines for HAART prescription, and that several nonmedical factors are weighed as heavily as disease severity. Another survey focused on the effect of a schizophrenia diagnosis on the recommendation of HIV treatment (Himelhoch et al., 2007), and found that the diagnosis does not affect the likelihood of a recommendation for treatment of HIV.

Statistics-based methods have also been used to determine associations between HIV treatment initiation and various factors. Logistic regression has been used to analyze the effects of various non-clinical factors such as age, gender and ethnicity on HIV antiretroviral therapy initiation (Stöhr et al., 2007, Holodniy et al., 2007). For Stöhr et al. (2007), they found that among age, gender, ethnicity and exposure category, only exposure category had a large effect on treatment initiation. For Holodniy et al. (2007), they did a study on treatment-naïve veterans under the US Department of Veteran Affairs and found that region of care has an effect on the treatment received. Cox regression was used by Tegger et al.

(2008) to examine the relationship between HAART initiation and the diagnoses of mental illness and substance use. Among their findings was that patients with depression/anxiety who weren't receiving medication for their disorder, patients with alcohol use disorders and patients who use opiates or amphetamines were less likely to initiate HAART compared to patients without the disorders. Easterbrook et al. (2008) used linear regression to determine factors related to antiretroviral therapy initiation at a lower CD4 cell count. They reported that the most important factors were being male, nonwhite, heterosexual or an injecting drug user. All these studies use commencement of therapy as the valid indicator that a factor has had an effect and are subject to the criticism that patient acceptance of advice to commence treatment is the actual indicator rather than the clinician's assessment of the suitability for treatment initiation.

Support for using Bayesian statistics in decision making comes from Lilford and Braunholtz (1996). They suggest that conventional statistical tests are inappropriate for decision making, since they present results as if there are only two choices: significant/not significant. They also prevent additional evidence outside the data from being handled explicitly. The authors suggest using posterior probability distributions to guide policy, since they present degrees of belief as being continuous rather than dichotomous. Further support is provided empirically by Troiani and Carlin (2004). They compared between classical, Bayesian and heuristic approaches for predicting whether a lung transplant recipient will have an acute disease event, concluding that the Bayesian approach is the best.

A number of papers have used Bayesian statistics to model HIV longitudinal data. Lange et al. (1992) studied the progression of HIV infection in 327 subjects by a high-dimensional (1314 parameters) Bayesian model with a linear growth curve. They focused on CD4 T-cell numbers as progress indicators, and their results suggested that CD4 T-cell numbers alone do not provide much information. A semi-Bayesian approach was used by Faucett and Thomas (1996) to model a covariate over time, simultaneously relating it to disease risk. The particular covariate under study was CD4 count, and the authors found that with the assumption of correct modeling their joint approach produced improved results compared to treating each relation separately.

3 Our Approach

Our approach takes the following steps:

1. Represent the pertinent data in the patient database into the form of a Directed Acyclic Graph (DAG) model.
2. Perform a hierarchical analysis (Cohen et al., 2002) on the model to estimate relevance of historical CD4 and VL variables, using Gibbs sampling (Gelfand and Smith, 1990) to estimate model parameter distributions.

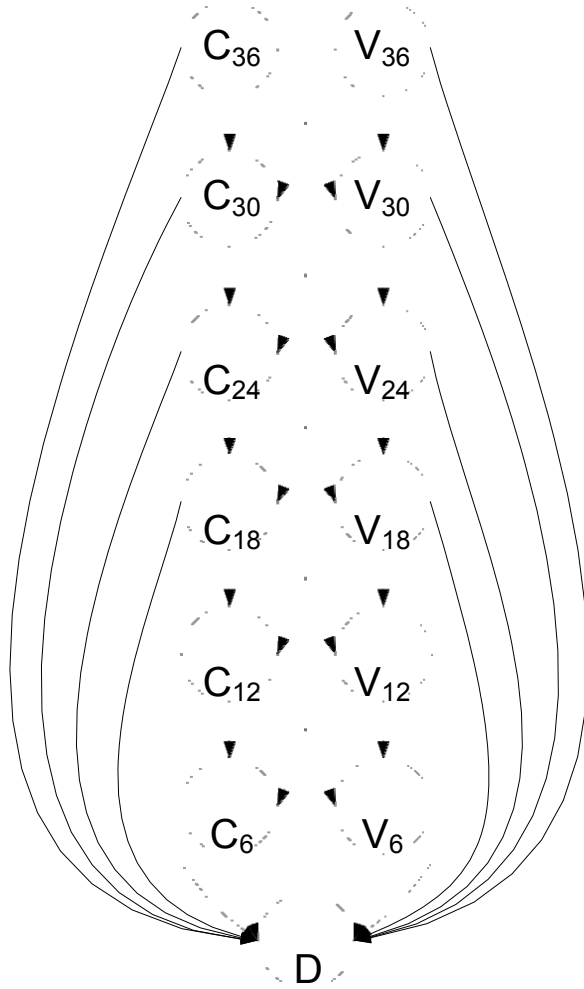


Figure 1: The main relationships between the historical CD4/VL variables and the treatment decision. Not shown are the dependencies between a CD4/VL reading and the readings before it.

A DAG can represent a Bayesian statistical model with many variables in which relationships between variables are represented by directed arrows between them and there is no closed loop in the graph. An example of a DAG is Figure 1. Elements of this model (e.g. the weight of a variable) can have prior probability distributions associated with them, and these distributions have their own associated parameters. In a DAG, it is possible to localize the relationships between distributions and through the Gibbs sampling method that can take advantage of these localized relationships, the posterior joint distribution and the parameters of the overall model can be estimated.

3.1 Data Extraction

For each HIV/AIDS patient P , three data vectors are obtained. The first vector contains the hospital visit dates of each patient, mathematically represented by the date vector $\text{Visit}_P = (t_1^v, t_2^v, t_3^v, \dots)$. We further associate each visit with a treatment variable that is 1 (true) if the visit resulted in treatment change or initiation; 0 (false)

otherwise. The second vector contains the historical CD4 readings for each patient, mathematically represented by the vector $\text{CD4}_P = (c_1, t_1^c), (c_2, t_2^c), (c_3, t_3^c), \dots$, where each pair represents a CD4 reading of c_i cells/mm³ taken at date t_i^c . The third vector contains the historical VL readings for each patient, mathematically represented by the vector $\text{VL}_P = (v_1, t_1^v), (v_2, t_2^v), (v_3, t_3^v), \dots$, where each pair represents a VL reading of v_i log10 copies/ml taken at date t_i^v . The branched-chained DNA (bDNA) method (Nolte and Herbert, 1998) was used to obtain the VL readings.

From the Visit_P , CD4_P and VL_P vectors, the following CD4 and VL readings are extracted: their historical values from 6 to 36 months before the visit date t^{vi} at 6-month intervals. For each reading x months ago, a window of $x - 3$ months is used. If there are no readings within that window, linear interpolation is done between the two closest readings before and after x months. If linear interpolation cannot be done due to the unavailability of data after x months, then that reading is recorded as unavailable and the associated visit date is excluded from further consideration. This ensures that in the extracted data set, the number of data points for each month window is balanced. In total, 6 CD4 variables, 6 VL variables and 1 treatment variable are associated with each visit.

3.2 Model Development

Figure 1 is a model showing the main dependencies between the variables in this study. Each C_i and V_i variable represents the CD4 and VL reading i months ago respectively while the D variable represents the treatment decision. Not shown in the model for brevity are the direct dependencies between a CD4/VL reading and all the readings that come before it.

We represent the relation between the treatment decision and the CD4/VL variables with the logistic model (Equation 2). The treatment decision D can be either 0 (no change of treatment) or 1 (change treatment). The predictors for the decision are represented by the p_i variables, while their associated coefficients (w_i) represent each predictor's influence on D . Noise and the influences of predictors that are not included are represented by ε , the general error term. The inverse-logit (logistic) function constrains the summation to values between 0 and 1.

$$D = \text{logit}^{-1}(w_1 p_1 + w_2 p_2 + \dots + w_n p_n) \quad (2)$$

The full analysis for Figure 1 is performed via a hierarchical analysis (Cohen et al., 2002). Hierarchical analysis is a statistical technique for models without reciprocal causation. It takes into account dependencies between variables, thus the temporal dependencies between CD4/VL variables are taken care of even though the underlying logistic model (Equation 2) treats variables as independent. In hierarchical analysis, each variable is considered one-by-one in order of causal priority. For simplicity, we demonstrate this technique on a smaller model (Figure 2) that considers CD4

variables only. For this model, the first variable considered is C_{36} since it is not causally dependent on any other variable. It is entered into the logistic model as in Equation 3. The estimated value of w then becomes the total effect of C_{36} on the treatment decision.

$$D = \text{logit}^{-1}(wC_{36}) \quad (3)$$

The next variable considered is C_{30} since among the remaining variables it is only causally dependant on the C_{36} variable. It is entered into the logistic model as in Equation 4. The coefficient w_{30} of the newly considered variable C_{30} is the total effect of that variable on the treatment decision while for the previously considered C_{36} variable, the difference between its coefficients in Equation 3 and Equation 4 is the indirect effect of that variable mediated through the C_{30} variable.

$$D = \text{logit}^{-1}(w_{36}C_{36} + w_{30}C_{30}) \quad (4)$$

Generalizing, the estimated coefficient of a variable when it first enters the logistic model is its total effect, while for previously considered variables, the difference of its coefficients before and after a newly entered variable enters the model is its indirect effect mediated through the newly entered variable. When all the variables are in the model as in Equation 5, the coefficient w_i of each variable C_i is the direct effect of C_i on the treatment decision.

$$D = \text{logit}^{-1}(w_{36}C_{36} + w_{30}C_{30} + w_{24}C_{24} + w_{18}C_{18} + w_{12}C_{12} + w_6C_6) \quad (5)$$

We also estimate the zero-order coefficients of every CD4 variable via Equation 6, and the magnitude of any spurious relationships (S_i) through Equation 7 where T_i is the total effect of the C_i variable.

$$D = \text{logit}^{-1}(z_iC_i) \quad (6)$$

$$S_i = z_i - T_i \quad (7)$$

For the main model (Figure 1), there are two parallel tracks that raise the question of whether to add the CD4 or VL variable first to the logistic equation. We experimented both ways and found that adding the VL variable first caused CD4 variables to have unusually large spurious effects. Therefore, we decided to add the CD4 variable first.

OpenBUGS, the successor to the well-known WinBUGS statistical modeling package (Lunn et al., 2009), is used to fit the logistic models and estimate the probability distributions for the coefficients. OpenBUGS can accept complex DAG models and perform Gibbs sampling on those models towards estimation of model parameter distributions. The BUGS software family has been applied in various applications including disease mapping (Lawson et al., 2003) and behavioral studies (Chevrolat et al., 1998).

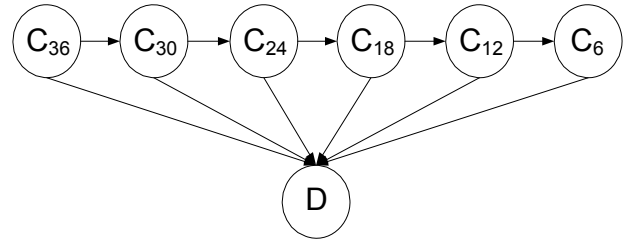


Figure 2: Model that considers CD4 variables only.

All coefficients are assumed to be Gaussian distributed with a prior distribution of $\mathcal{N}(0, 1)$, reflecting a weak initial conservative assumption that the associated variable is not significant. As is custom, the error, ε , is assumed to be Gaussian distributed with mean 0 and unknown variance. For every logistic equation, parameters were estimated from 20000 sampling runs after a burn-in period of 5000 runs. To facilitate comparison, all variables have been normalized to have a mean of 0 and a standard deviation of 1.

4 Experimental Study

4.1 Population

We demonstrate our procedure on a dataset of 528 HIV/AIDS patients from the Royal Prince Alfred Hospital, Sydney who had visited the hospital between January 3, 1991 and March 24, 2009. Ethics approval for usage of the data was obtained from the Sydney South West Area Health Service. A total of 15923 visits were recorded. The number of visits per patient ranged from 1 to 199 visits with a median of 16 visits. The median time between visits is 1 month. Each patient has a history of CD4 and VL samples at irregular intervals depending on whether a sample was taken during a particular visit. Additionally, the database also records when a particular HIV treatment was changed or initiated for a particular patient.

In terms of data quality, the patient samples are processed by quality-assured laboratories that participate in external quality assurance programs. Treatment changes in the database are recorded by the treating physician and do reflect the use of antiretroviral agents towards initiation of a new treatment regimen or change of the current treatment regimen. The host country, Australia, provides wide-ranging access to HIV medicine via a government-subsidized program that provides discounted HIV medicine to qualified recipients based on clinical and laboratory criteria.

4.2 Results and Discussion

The zero-order effects of each historical VL variable (Table 1) show that as expected, viral load has a strong positive correlation with the treatment decision. There is generally a decreasing trend in effect with age, with sharp drops between the VL variables at 6 and 12 months and between the VL variables at 30 and 36 months. It is interesting to note that the VL variable at 30 months appears to buck this decreasing trend.

VL age (months)	Mean	Std dev	95% credible interval
6	0.3672	0.0573	(0.2555, 0.4807)
12	0.2947	0.0581	(0.1811, 0.4096)
18	0.2735	0.0587	(0.1586, 0.3891)
24	0.2508	0.0594	(0.1347, 0.3676)
30	0.2538	0.0597	(0.1371, 0.3712)
36	0.1657	0.0603	(0.0479, 0.2837)

Table 1. Zero-order effects of the historical VL variables.

CD4 age (months)	Mean	Std dev	95% credible interval
6	-0.1638	0.0731	(-0.3121, -0.0246)
12	-0.1016	0.0694	(-0.2431, 0.0291)
18	-0.0970	0.0692	(-0.2389, 0.0335)
24	-0.0896	0.0689	(-0.2309, 0.0399)
30	-0.0578	0.0661	(-0.1951, 0.0661)
36	-0.0327	0.0635	(-0.1642, 0.0862)

Table 2. Zero-order effects of the historical CD4 variables.

VL age (months)	Mean	Std dev	95% credible interval
6	0.3418	0.0970	(0.1526, 0.5320)
12	0.2161	0.0960	(0.0284, 0.4067)
18	0.1994	0.0952	(0.0143, 0.3869)
24	0.1487	0.0938	(-0.0328, 0.3336)
30	0.2847	0.0891	(0.1106, 0.4604)
36	0.1705	0.0630	(0.0464, 0.2938)

Table 3. Total effects of the historical VL variables.

CD4 age (months)	Mean	Std dev	95% credible interval
6	-0.2046	0.1597	(-0.5169, 0.1100)
12	0.0168	0.1589	(-0.3052, 0.3194)
18	-0.0383	0.1490	(-0.3325, 0.2477)
24	-0.1015	0.1430	(-0.3804, 0.1805)
30	-0.0868	0.1332	(-0.3441, 0.1764)
36	-0.0327	0.0635	(-0.1642, 0.0862)

Table 4. Total effects of the historical CD4 variables.

For the zero-order effects of the CD4 variables, the results are inconclusive with a 95% credible interval (Table 2). However, compared to the prior distribution of $\mathcal{N}(0, 1)$, all the posterior distributions have been shifted in the negative direction. This suggests that the CD4 count has a negative correlation with the treatment decision, though the correlation is weaker compared to viral load. Again, there is a decreasing trend with age with the CD4 variable at 6 months having a considerably stronger correlation with the treatment decision compared to the rest.

The total effects of the historical variables obtained after the hierarchical analysis is performed show a different picture. For the VL variables (Table 3), there is still a decreasing trend in effect as the VL variable ages. For the 30-month VL variable however, its bucking of the trend is far more pronounced, with its total effect appearing to be considerably higher than both the 24-month and 36-month VL variables. The CD4 variables (Table 4) also show a decreasing trend in effect, though

compared to the zero-order effects, the total effect of the CD4 variables from 6 months to 24 months follow a different profile as it becomes close to zero for the 12-month mark and then recovering at the 24-month mark.

Considering direct, indirect and spurious effects for the VL variables (Table 5, left side), the 24-month, 18-month and 12-month VL variables have direct effects that are close to zero and a low ratio of direct to indirect effects (0.10, 0.22 and 0.02 respectively). This suggests that their influence on the treatment decision is mostly through variables that are more recent. As for the 36-month VL variable, it has a moderate direct effect but it also has a low ratio of direct to indirect effects (0.40) which suggests that more recent variables can provide much of its total effect on the treatment decision. The 30-month VL variable has both a moderate direct effect and a high ratio of direct to indirect effects (1.01) suggesting that considerable information is provided by that variable that is independent of the other VL and CD4 variables.

For the CD4 variables (Table 5, right side), it is the 36-month and 30-month CD4 variables that have direct effects close to zero and a low ratio of direct to indirect effects (0.44 and 0.19 respectively), thus suggesting their influence on the treatment decision is mostly through more recent variables. The 24-month, 18-month and 12-month CD4 variables all have a high ratio of direct to indirect effects (18.15, 1.52 and 1.18 respectively) which as before suggests that a considerable portion of their total effect towards the treatment decision is independent of all other variables. A case can be made though for removal of the 18-month CD4 variable due to its low direct effect. As for the 6-month CD4 variable, Table 5 shows that its effect on the treatment decision is mostly through the 6-month VL variable, thus suggesting that the 6-month VL variable can be a substitute for it.

Other features of note are that as expected, for a VL variable at a particular period, its strongest indirect effect on the treatment decision is through the VL variable 6 months later, and then the effect drops off considerably at 12 months onwards. This provides additional support for HIV guideline recommendations (DHHS, 2011) that VL readings need to be taken regularly at intervals below 6 months for they appear to quickly become out-of-date. For CD4 variables, the situation is less clear. Table 5 shows that a CD4 variable at a particular period has a considerable indirect negative effect on the treatment decision through the VL variable taken in the same period and then a considerable indirect positive effect through the VL variable 6 months later. A possible explanation is that a high CD4 cell count may initially slow down the progress of infection and appearance of symptoms, but then eventually the CD4 cells are infected by the virus and become virus factories with overall rate of production determined by the initial number of CD4 cells available for infection.

5 Conclusions

For the 24-month, 18-month and 12-month VL variables and the 36-month and 30-month CD4 variables, the

	Zero-order	Spurious	Direct, indirect and total effects		Zero-order	Spurious	Direct, indirect and total effects
36-month VL	0.1657	-0.0048	-0.1144	36-month CD4	-0.0327	-	0.0259
Via 30-month VL			0.2007	Via 30-month CD4			-0.0692
Via 24-month VL			0.0296	Via 24-month CD4			-0.0357
Via 18-month VL			0.0144	Via 18-month CD4			-0.0096
Via 12-month VL			0.0104	Via 12-month CD4			-0.0008
Via 6-month VL			0.0337	Via 6-month CD4			-0.0120
Via 30-month CD4			0.0042	Via 36-month VL			-0.0489
Via 24-month CD4			-0.0040	Via 30-month VL			0.0984
Via 18-month CD4			0.0008	Via 24-month VL			0.0111
Via 12-month CD4			0.0009	Via 18-month VL			0.0139
Via 6-month CD4			-0.0058	Via 12-month VL			-0.0045
Total effect			= 0.1705	Via 6-month VL			-0.0013
				Total effect			= -0.0327
30-month VL	0.2538	-0.0309	0.1430	30-month CD4	-0.0578	0.0290	0.0208
Via 24-month VL			0.0882	Via 24-month CD4			-0.0503
Via 18-month VL			0.0267	Via 18-month CD4			-0.0021
Via 12-month VL			0.0050	Via 12-month CD4			0.0097
Via 6-month VL			0.0086	Via 6-month CD4			-0.0154
Via 24-month CD4			0.0108	Via 30-month VL			-0.1146
Via 18-month CD4			-0.0008	Via 24-month VL			0.0375
Via 12-month CD4			0.0027	Via 18-month VL			0.0019
Via 6-month CD4			0.0005	Via 12-month VL			0.0173
Total effect			= 0.2847	Via 6-month VL			0.0085
				Total effect			= -0.0868
24-month VL	0.2508	0.1021	-0.0157	24-month CD4	-0.0896	0.0119	-0.0962
Via 18-month VL			0.1163	Via 18-month CD4			-0.0184
Via 12-month VL			0.0343	Via 12-month CD4			-0.0001
Via 6-month VL			0.0155	Via 6-month CD4			-0.0380
Via 18-month CD4			0.0037	Via 24-month VL			-0.0562
Via 12-month CD4			-0.0022	Via 18-month VL			0.0649
Via 6-month CD4			-0.0033	Via 12-month VL			0.0008
Total effect			= 0.1487	Via 6-month VL			0.0417
				Total effect			= -0.1015
18-month VL	0.2735	0.0741	0.0367	18-month CD4	-0.0970	-0.0587	-0.0231
Via 12-month VL			0.1237	Via 12-month CD4			0.0105
Via 6-month VL			0.0296	Via 6-month CD4			-0.0089
Via 12-month CD4			0.0007	Via 18-month VL			-0.0891
Via 6-month CD4			0.0087	Via 12-month VL			0.0768
Total effect			= 0.1994	Via 6-month VL			-0.0045
				Total effect			= -0.0383
12-month VL	0.2947	0.0786	0.0039	12-month CD4	-0.1016	-0.1184	0.1107
Via 6-month VL			0.1919	Via 6-month CD4			-0.1061
Via 6-month CD4			0.0203	Via 12-month VL			-0.0987
Total effect			= 0.2161	Via 6-month VL			0.1109
				Total effect			= 0.0168
6-month VL	0.3672	0.0254	0.3418	6-month CD4	-0.1638	0.0408	-0.0200
Total effect			= 0.3418	Via 6-month VL			-0.1846
				Total effect			= -0.2046

Table 5. Direct, indirect, spurious, zero-order and total effects of the historical variables on the treatment decision.

combination of low direct effects and low ratios of direct to indirect effects suggest that these variables add little additional predictive power about the treatment decision and can be dropped with little effect. If more aggressive removal is required, variables that could possibly be dropped include the 18-month CD4 variable due to its

low direct effect, and the 36-month VL and 6-month CD4 variables due to their low ratios of direct to indirect effects. For each of the other variables, they have large direct effects and a considerable proportion of their total effect on the treatment decision is not indirect, thus they should be protected from removal.

Limitations of our method include treating each time point in the HIV patient's history as equal, ignoring changes in the patient and HIV virus, and assuming that the decision making process doesn't change over time with new regulations and knowledge about HIV/AIDS. It is reasonable to assume however that in general terms the biological interactions between the HIV virus, the immune system and treatment is constant over the short time span of a patient's history, and that treatment has always been with intent to eradicate the HIV virus.

Future work will extend the model to include other factors that may affect the treatment decision, for example adherence to treatment and genetic factors from the patient and HIV virus. Another extension approach would be to further particularize the model on the different antiretroviral treatment available and determine whether the results are universal over the different treatment. Research is also needed to determine whether these results are specific to 3-year old patients or are more generally applicable.

This study suggests that in the building of decision support systems for the management of HIV/AIDS the crucial data for triggering a decision to review medication recommendations is the more recent data and little would be lost by omitting more distant data from the dataset for review; it must be acknowledged however that the recommendation for treatment is highly complex as there are more than 20 agents available and generally treatment comprises three or four drugs used concurrently. Thus longitudinal data that demonstrates patterns of response to particular agents in a specific individual as well as from cohorts treated with certain combinations may still be of great value in guiding the final decision. Our analysis suggests that a two step process may be suitable with default data display and decision support to trigger review based on recent data being the first step and detailed evaluation of an archived dataset as the next. Such an approach would be amenable to prospective study and rigorous evaluation.

6 References

- Bogart, L.M., Catz, S.L., Kelly, J.A. & Benotsch, E.G. (2001): Factors Influencing Physicians' Judgments of Adherence and Treatment Decisions for Patients with HIV Disease. *Medical Decision Making*, **21**(1):28-36.
- Bogart, L.M., Kelly, J.A., Catz, S.L. & Sosman, J.M. (2000): Impact of Medical and Nonmedical Factors on Physician Decision Making for HIV/AIDS Antiretroviral Treatment. *Journal of Acquired Immune Deficiency Syndromes*, **23**(5):396-404.
- Chevrolat, J.-P., Golmard, J.-L., Ammar, S., Jouvent, R. & Boisvieux, J.-F. (1998): Modelling behavioral syndromes using Bayesian networks. *Artificial Intelligence in Medicine*, **14**(3):259-277.
- Cohen, P., Cohen, J., West, S.G. & Aiken, L.S. (2002): *Applied Multiple Regression/Correlation Analysis for the Behavioral Sciences*. London, Routledge Academic.
- DHHS (2011): Guidelines for the use of antiretroviral agents in HIV-1-infected adults and adolescents - 10 January, 2011. Panel on Antiretroviral Guidelines for Adults and Adolescents, Department of Health and Human Services.
- Easterbrook, P.J., Phillips, A.N., Hill, T., Matthias, R., Fisher, M., Gazzard, B., Gilson, R., Scullard, G., Johnson, M., Dunn, D.T., Orkin, C., Anderson, J., Schwenk, A., Leen, C. & Sabin, C.A. (2008): Patterns and predictors of the use of different antiretroviral drug regimens at treatment initiation in the UK. *HIV Medicine*, **9**(1):47-56.
- Faucett, C.L. & Thomas, D.C. (1996): Simultaneously modelling censored survival data and repeatedly measured covariates: a Gibbs sampling approach. *Statistics in Medicine*, **15**(15):1663-1685.
- Gelfand, A.E. & Smith, A.F.M. (1990): Sampling-Based Approaches to Calculating Marginal Densities. *Journal of the American Statistical Association*, **85**(410):398-409.
- Goujard, C., Bonarek, M., Meyer, L., Bonnet, F., Chaix, M.-L., Deveau, C., Sinet, M., Galimand, J., Delfraissy, J.-F., Venet, A., Rouzioux, C. & Morlat, P. (2006): CD4 Cell Count and HIV DNA Level Are Independent Predictors of Disease Progression after Primary HIV Type 1 Infection in Untreated Patients. *Clinical Infectious Diseases*, **42**(5):709-715.
- Himelhoch, S., Powe, N.R., Breakey, W. & Gebo, K.A. (2007): Schizophrenia, AIDS and the Decision to Prescribe HAART -- Results of a National Survey of HIV Clinicians. *Journal of Prevention & Intervention in the Community*, **33**(1):109-120.
- Holodniy, M., Hornberger, J., Rapoport, D., Robertus, K., MacCurdy, T.E., Lopez, J., Volberding, P. & Deyton, L. (2007): Relationship Between Antiretroviral Prescribing Patterns and Treatment Guidelines in Treatment-Naive HIV-1-Infected US Veterans (1992-2004). *Journal of Acquired Immune Deficiency Syndromes*, **44**(1):20-29.
- Lange, N., Carlin, B.P. & Gelfand, A.E. (1992): Hierarchical Bayes Models for the Progression of HIV Infection Using Longitudinal CD4 T-Cell Numbers. *Journal of the American Statistical Association*, **87**(419):615-626.
- Langford, S., Ananworanich, J. & Cooper, D. (2007): Predictors of disease progression in HIV infection: a review. *AIDS Research and Therapy*, **4**(1):11.
- Lawson, A.B., Browne, W.J. & Rodeiro, C.L.V. (2003): *Disease Mapping with WinBUGS and MLwiN*. Chichester, Wiley.
- Lilford, R.J. & Braunholtz, D. (1996): The Statistical Basis Of Public Policy: A Paradigm Shift Is Overdue. *BMJ: British Medical Journal*, **313**(7057):603-607.
- Lunn, D., Spiegelhalter, D., Thomas, A. & Best, N. (2009): The BUGS project: Evolution, critique and future directions. *Statistics in Medicine*, **28**(25):3049-3067.

- Nolte, F.S. & Herbert, E.S. (1998): Branched DNA Signal Amplification for Direct Quantitation of Nucleic Acid Sequences in Clinical Specimens. In *Advances in Clinical Chemistry*. 201-235. Spiegel, H.E. (ed). Elsevier.
- Phillips, A. & Pezzotti, P. (2004): Short-term risk of AIDS according to current CD4 cell count and viral load in antiretroviral drug-naïve individuals and those treated in the monotherapy era. *AIDS*, **18**(1):51-58.
- Stöhr, W., Dunn, D.T., Porter, K., Hill, T., Gazzard, B., Walsh, J., Gilson, R., Easterbrook, P., Fisher, M., Johnson, M.A., Delpech, V.C., Phillips, A.N. & Sabin, C.A. (2007): CD4 cell count and initiation of antiretroviral therapy: trends in seven UK centres, 1997-2003. *HIV Medicine*, **8**(3):135-141.
- Tegger, M.K., Crane, H.M., Tapia, K.A., Uldall, K.K., Holte, S.E. & Kitahata, M.M. (2008): The Effect of Mental Illness, Substance Use, and Treatment for Depression on the Initiation of Highly Active Antiretroviral Therapy among HIV-Infected Individuals. *Aids Patient Care and STDs*, **22**(3):233-243.
- Troiani, J.S. & Carlin, B.P. (2004): Comparison of Bayesian, classical, and heuristic approaches in identifying acute disease events in lung transplant recipients. *Statistics in Medicine*, **23**(5):803-824.
- UNAIDS (2010) Global report: UNAIDS report on the global AIDS epidemic.

Inducing and Storing Generalised Evidences using Semantic Web formalisms

Razan Paul, Tudor Groza and Jane Hunter

School of ITEE, The University of Queensland, Australia

{razan.paul, tudor.groza, j.hunter}@uq.edu.au

Abstract

Over the course of the last decade, decision support systems have been used to assist clinicians and researchers in expanding the body of knowledge of particular (bio)-medical areas, as well as in diverse decision-making processes (e.g., diagnosis, treatment). Creating a decision support model (e.g., a rule base) requires a set of well-established medical guidelines built on mature domain knowledge. The absence of such mature domain knowledge has hindered the development of appropriate decision support methods in the skeletal dysplasia domain. In this paper, we make the first step towards providing a solution to this issue by proposing an ontology and associated extraction algorithm that can infer generalized evidences from existing bone dysplasia patient cases. This establishes the foundation for a decision support model based on evidential reasoning, which enables semi-automated diagnosis or key disease feature extraction.

Keywords: Evidence Ontology, Decision support methods, Evidential Reasoning

1 Introduction

Skeletal dysplasias are a heterogeneous group of genetic disorders affecting skeletal development. Currently, there are over 450 recognized bone dysplasias, structured in 40 groups. Patients with skeletal dysplasias have complex medical issues including short stature, bowed legs, a larger than average head and neurological complications. However, since most skeletal dysplasias are very rare (<1:10,000 births), data on clinical presentation, natural history and best management practices is sparse. Another reason for data sparseness is the small number of phenotypic characteristics typically exhibited by patients from the large range of possible phenotypic and radiographic characteristics usually associated with these diseases. Due to the rarity of these conditions and the lack of mature domain knowledge, correct diagnosis is often very difficult. In addition, only a few centres worldwide have expertise in the diagnosis and management of these disorders.

As there are no defined guidelines, the diagnosis of new cases relies strictly on parallels to past case studies.

Molecular genetics research on skeletal dysplasias has advanced considerably over the years – enabling the

genetic defects responsible for more than 200 skeletal dysplasias to be identified. However, a lack of decision support methods and interoperable knowledge bases available to the skeletal dysplasia community – hinders collaborative diagnosis and research in this area. A rich knowledge base, together with associated decision support methods would enable researchers to verify known trends and to discover new, previously unknown trends among clinical attributes associated with this class of diseases, that can be used to assist and inform the decision making process associated with disease diagnosis and characterisation.

The general sparseness and disperse nature of skeletal dysplasia data has limited the development and availability of authoritative databases by the leading clinical and research centres. To make diagnoses, improve understanding and identify best treatments, clinicians need to analyse historical dysplasia patient data, verify known facts and relationships and discover new and previously unknown facts and relationships among the phenotypic, radiographic and genetic attributes associated with existing and new cases. For example, it is currently extremely difficult to recognise skeletal dysplasias that are etiologically related or to identify clinical or radiological characteristics that are indicative of defects resulting from a specific molecular pathway.

In order to do this, researchers and clinicians currently need to query many heterogeneous data sources and to effectively aggregate diverse types of data relating to phenotypic, radiographic and genetic observations. Given the appropriate data integration and reasoning tools, clinicians should be able to deduce, for example, that “mutations of the COL3A1 gene cause Platyspondylic lethal skeletal dysplasia which is characterised by short fingers in 90% of patients”. However, this data integration step represents a significant challenge due to the extreme heterogeneity of the data models, metadata schemas and vocabularies, data formats and inconsistencies in naming and identification conventions.

Semantic Web standards (Berners-Lee, Hendler et al. 2001; Shadbolt, Hall et al. 2006; Allemang and Hendler 2008) encode and formalize data and background knowledge associated with a specific domain by means of standardized metadata schemas, controlled vocabularies and ontologies. These standards are critical to facilitating information sharing and integration. Hence, a key aim is to apply Semantic Web technologies to the data integration challenge described above by formalizing and modelling dysplasia data using ontologies and controlled vocabularies.

The above-mentioned issues also limit the potential of successfully applying existing or traditional knowledge representation, reasoning and decision support methods in

the bone dysplasia domain, such as Rule Based Systems (Hudson 2006), Neural networks (Chan, Ling et al. 2011), Fuzzy cognitive maps (FCMs) (Hudson 2006; Gadaras and Mikhailov 2009; Papageorgiou, Papandrianos et al. 2009; Begum, Ahmed et al. 2010; Chan, Ling et al. 2011)) or Fuzzy Rule based classifications (Gadaras and Mikhailov 2009). Creating a decision support model (e.g., a rule base) requires a set of well-established medical guidelines and mature domain knowledge. However in the skeletal dysplasia domain, clinicians frequently have to diagnose patients with little or no similarity to past cases – this requires the generation of new evidence by combining existing evidence.

Our hypothesis is that by representing the knowledge and data using Semantic Web formalisms, and applying inductive reasoning on the resulting knowledge base – we can induce generalized evidences and store them in an Evidence Ontology. The result is an interoperable generalized evidence store for the skeletal dysplasia domain. Storing generalized evidences in an ontology enables sharing among and aggregation from multiple autonomous systems, thus leading to a distributed decision support approach.

The remainder of the paper is structured as follows: Section 2 provides a comprehensive overview of generically related efforts in the decision support area, while Section 3 introduces the fuzzy terminology used by our approach described in Section 4. Before concluding in Section 6, we briefly describe our evaluation plans in Section 5.

2 Related Work

Most prior work in representing generalized knowledge for medical decision support methods (Hudson 2006; Gadaras and Mikhailov 2009; Papageorgiou, Papandrianos et al. 2009; Begum, Ahmed et al. 2010; Chan, Ling et al. 2011) use some non-standard formalisms or proprietary formats which hinder integration, interoperability and efficient knowledge reasoning. They also lead to unjustified results by fusing all generalized knowledge into a black box system or assume the existence of mature domain knowledge. Moreover, some of these previous methods cannot evolve over time, due to their shallow knowledge representation formalisms. Case-based reasoning (Begum, Ahmed et al. 2010), on the other hand, cannot combine past evidences to form a new evidence for a given problem where no past similar evidence exists. This scenario is typical for rare diseases like skeletal dysplasias. It also uses non-generalized evidences, which does not guarantee correctness.

Rule based systems (Hudson 2006) and fuzzy rule-based classification (Gadaras and Mikhailov 2009) use exact matching on rules built on mature and established domain knowledge – which is inapplicable in a domain that suffers from data sparseness. The neural network approach (Chan, Ling et al. 2011) cannot provide justification for the resulting knowledge because it fuses all the evidence into the internal weights, whereas in the skeletal dysplasia domain, justification is very important to both clinicians and researchers in order to understand the underlying causal elements.

Today's decision support systems require the automatic integration of knowledge from multiple sources. However, the lack of interoperability and standard formalisms impede these systems to take advantage of the connectivity provided by the Web. Decision support systems (Goossen, Jntema et al. 2011; Lee and Wang 2011) using Semantic Web standards are being developed to overcome the above challenges. Semantic Web rule-based reasoning has been used for domain specific decision support methods, for example, in the Ambient Intelligence domain (Patkos, Chrysakis et al. 2010). However, such approaches cannot make use of underlying trends in instance data that have not been encoded as ontological background knowledge and cannot handle probabilistic uncertainties within the knowledge. Moreover, they cannot form new evidence by combining existing evidence via reasoning, where there exist no prior examples.

A recent related effort (Lee and Wang 2011) presents a novel fuzzy expert system for a diabetes decision support application using a 5-layer fuzzy ontology and a semantic decision support agent. However, as with its predecessors, this system also depends on mature and established domain knowledge, and uses fuzzy rule-based reasoning (Straccia 2008), which follows an exact matching approach.

Medical decision support systems have emerged from the co-evolution of research in decision support systems and medical informatics. In (Hussain, Abidi et al. 2007), a Semantic Web based Clinical Decision Support System is presented to provide evidence guided recommendations for follow-up after treatment for Breast Cancer. ControlSem (Andreasik, Ciebiera et al. 2010), a medical decision support system using Semantic Web technologies, was developed with the goal of controlling medical procedures. Similarly, in (Prcela, Gamberger et al. 2008), the authors present a medical expert system for heart failure. These expert systems use general purpose rule base reasoning (deductive reasoning) (Straccia 2008) because the underlying domain has well-defined rules and mature background knowledge.

Reasoning plays a vital role on the Semantic Web and is based on the background knowledge provided by the data model, logic and rules layers. Deductive reasoning is able to derive new knowledge, however, is relies completely on rules and existing ontological background knowledge and, hence, cannot make use of regularities in the instance data that have not been. In contrast, induction can exploit regularities in the instance data to discover new generalised rules or evidences.

Data mining is also applied to discover new information, hidden in patterns emerging from existing information. One of widely used techniques in data mining is finding association rules. The first pioneering work to mine conventional positive association rules using a level wise search algorithm was explained in (Agrawal, Imieli ski et al. 1993). Following this work, many other, improved, algorithms have been proposed, in particular for finding rules that represent decision occurring frequently based on a set of facts (Doddi, Marathe et al. 2001; Pan, Li et al. 2005; Sheela and Shanthi 2009; Weng and Chen 2010).

Finally, Semantic Web Mining (Lisi 2006; Lisi 2006; Stumme, Hotho et al. 2006) is a new research area that aims to discover hidden knowledge from Semantic Web instance data by combining Semantic Web techniques and data mining. The newly discovered knowledge can then be used for enriching the domain model and, hence, possibly improve the future decision making process. Most of the existing work in Semantic Web mining applies existing data mining algorithms in the Semantic Web context. For example, (Lisi 2006) describes a middleware, SWing, to enable inductive reasoning on the Semantic Web. Similarly, (Maedche and Staab 2000) use association mining to extract relations from text.

3 Uncertainty, Fuzzy Set Theory and Ontology

Uncertainty comes in various forms: probabilistic uncertainty (e.g., “There is a 65% chance I will get my promotion”), vagueness (fuzziness – e.g., “Mike is old to some degree”), ambiguity, subjectivity, incompleteness, etc. It is widely accepted that uncertainty is an indispensable part of medical data, and that the first two types of uncertainty play an important role, e. g., a symptom may or may not occur with a disease, it has an uncertain relation with the disease, etc (De, Biswas et al. 2001; Straszecka 2006).

In the Semantic Web world, OWL ontologies and SWRL rules can be used to capture the domain knowledge in a highly expressive manner. However, these cannot model vague and uncertain knowledge, and implicitly concepts, such as “short” Limb, “happy” person or “narrow” chest, because are unable to capture the degree of happiness or the measure of shortness.

Fuzzy set theory and fuzzy logic (Ross 2010) are suitable formalisms to handle imprecise and vague knowledge of a particular domain. In traditional set theory, any element belongs or not to a set, in type-1 fuzzy set theory, any element can belong partially to a set. For example, Tim has “ $short\ limb \geq 0.5$ ” states that Tim has a short limb with a degree of at least 0.5. The traditional set theoretic operations are extended to the Fuzzy set and Fuzzy complement, union, intersection and the logical operation of implication are performed by special mathematical functions over the unit interval, and they are defined as fuzzy complement (c), tconorm (u), t-norm (t or $*$) and fuzzy implication (\Rightarrow) (Ross 2010). There are other uncertainties that type-1 fuzzy set cannot handle, e.g., the *confidence* or *certainty* that Tim has “ $short\ limb \geq 0.5$ ”. Type-2 fuzzy sets (Castillo, Melin et al. 2007) can handle these types of uncertainties by associating uncertainty with the membership function of a type-1 fuzzy set.

A fuzzy linguistic variable defines the terminology required to use a fuzzy concept like age in expressing rules and facts. A Fuzzy value is an instance of a fuzzy concept for a fuzzy linguistic variable, e.g., age is young ≥ 0.8 . A Fuzzy linguistic Term is a word or expression used to facilitate the expression of Fuzzy value for a fuzzy linguistic variable. For example, age may have the Fuzzy terms {young, adult, old}. In our method, we employ type-2 fuzzy sets, fuzzy Linguistic variables, Fuzzy Value and Fuzzy Linguistic Term. Let’s suppose we have to represent short limb with a degree of at least

0.8 and a certainty of 0.5. In this case we define a Fuzzy linguistic variable “limb”, featured by three Fuzzy linguistic terms: {short, medium, long}. Our representation string will then be: 0.5 / (short, 0.8).

There are many concepts in medical domain that are vague and have no clear boundaries, such as “young”, “tall” or “small”. It is widely known that the crisp formalisms such as the one provided by OWL cannot handle vague and uncertain information on the Semantic Web. There are nevertheless, other ways to deal with such data. Firstly, it is possible to extend current Semantic Web languages to cope with fuzzy and uncertain information. Secondly, one can develop a specific, fuzzy, ontology. The World Wide Web Consortium (W3C) has set up a working group to work on representing and reasoning under uncertainty using ontologies. Results of this group can be seen in the existing fuzzy DL reasoners, like fuzzyDL (Simou and Kollias 2007; Bobillo and Straccia 2008) and FiRE (Simou and Kollias 2007). Straccia (Straccia 2010) and Pan (Pan, Stamou et al. 2007) have also described mechanisms for persistent storage and querying of fuzzy and uncertain information in databases.

From the fuzzy ontology perspective, there have been several solutions proposed to date. Gu et al (Gu, Wang et al. 2007) describe a Fuzzy Ontology of edutainment based on reification of relations in OWL, a technique similar to representing n-array relations (Noy, Rector et al. 2006) in OWL. (Bobillo and Straccia 2009) propose an OWL ontology to represent important features of fuzzy OWL 2 statements, via temporary concepts (nodes) like ConjunctionConcept and ConceptAssertion. At a later stage, the same authors also propose a concrete methodology to represent a fuzzy ontology using OWL 2 annotation properties (Bobillo and Straccia 2010). Finally, (Stoilos, Stamou et al. 2005) extend OWL with fuzzy set theory in order to capture, represent and reason with fuzzy and uncertain information, while (da Costa, Laskey et al. 2008) propose the PR-OWL formalism by extending OWL to provide the ability to express probabilistic knowledge.

4 Research Methodology

Figure 1 presents the high level building blocks of our research methodology, represented by the SKELETOME ontology set, the Evidence Extraction Process and the Evidence ontology. In the following sections we detail each of the three building blocks.



Fig. 1. Inducing generalised evidences from the Skeletome Ontology set and storing them into the Evidence Ontology

SKELETOME Ontology set. The main role of the SKELETOME Ontology set is to improve the highly static and rigid format of the ISDS Nosology (Warman, Cormier Daire et al.) by enabling a more flexible classification of the disorders and the integration with

existing Web resources, such as the Human Phenotype Ontology (Robinson, Köhler et al. 2008) or the NCI Thesaurus (Sioutos, Coronado et al. 2007). The set is composed of three ontologies: the Bone Dysplasia ontology that captures the complex relations between the phenotypic, radiographic and genetic elements characterizing all skeletal dysplasias; the Patient ontology that models patient information and the Context ontology maintaining provenance information.

Evidence Ontology. The Evidence Ontology models uncertainty (both fuzzy / vagueness and probabilistic uncertainty) by re-using concepts from probabilistic uncertainty and from Fuzzy Theory, such as fuzzy value, fuzzy variable, fuzzy set, membership value and fuzzy term. It enables the representation of uncertain generalized evidences and helps to simplify uncertain knowledge representation in OWL. The crisp syntax of OWL DL is used within the Evidence ontology to enable the encoding of Fuzzy and probabilistic uncertainty semantics.

Evidence extraction process. The generalized evidence extraction from past patient cases stored via the SKELETOME ontology set is a crucial prerequisite for the implementation of any decision support method, like automated diagnosis or key disease feature inference. Without the extracted evidence, uncertainty reasoning cannot be performed. The actual extraction process uses Machine Learning techniques, and more specifically, a level wise search algorithm (Paul and Hoque 2010), to be able to infer evidences from the instances of the SKELETOME ontology concepts, made available by domain experts.

4.1 The SKELETOME Ontology Set

As already described, the SKELETOME Ontology Set consists of three ontologies that model together the skeletal dysplasia domain knowledge, patient information and context information.

The actual requirements of the ontology set emerged from the needs of the skeletal dysplasia community, and include the following:

Common terminology: The diagnosis and management of skeletal dysplasias depends on highly specialised domain knowledge across a number of disciplines (radiography, genetics, orthopaedics, physiotherapy), which is not easily comprehensible to individual communities or hospitals. In order to enable the exchange of knowledge between experts (across languages and disciplines), patients, their families and medical staff, a common terminology is required, hence leading to a shared conceptualisation of the domain.

Data integration: Large datasets containing rich information on molecules (genes, proteins) already exist and the information relevant to skeletal dysplasias needs to be extracted and cross-referenced with the clinical data and knowledge produced by SKELETOME. This requires integration both at conceptual level, as well as, at actual data / instance level.

Capturing provenance and expertise: The contributed content may take several forms, ranging from personal observations to scientific publications. Independently of the form, SKELETOME requires a mechanism to keep

track of the provenance of the data and knowledge (to ensure proper privacy and access control), to provide a measure of certainty of derived data and to leverage expertise from the content and to streamline the delivery of the most relevant information / queries to the most appropriate person.

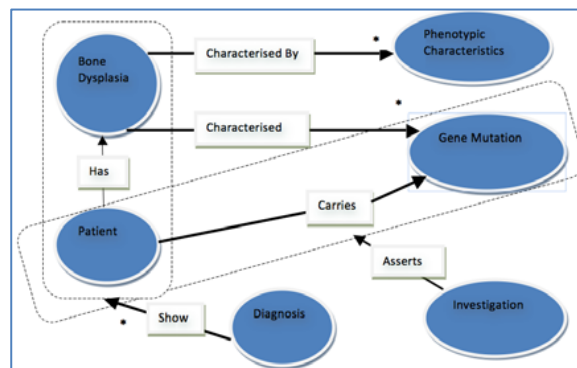


Fig. 2. Core concepts of the SKELETOME Ontology Set

Figure 2 depicts the core concepts of the SKELETOME Ontology Set. *Bone Dysplasia*, *Phenotypic Characteristic* and *Gene Mutation* are concepts defined by the Bone Dysplasia Ontology. The Bone Dysplasia ontology aims to complement the spectrum of existing ontologies and address the specific knowledge representation shortcomings of the ISDS Nosology (Warman, Cormier Daire et al.). None of the existing phenotype ontologies (e.g., the Human Phenotype Ontology) or well-known terminologies (e.g., SNOMED-CT) describe in detail skeletal dysplasias. As a result, our ontology provides a comprehensive, accurate and formal representation of the genotypes and phenotypes involved in skeletal dysplasias, together with their specific and disease-oriented constraints. As opposed to the current ISDS Nosology, this ontology enables a shared conceptual model, formalised in a machine-understandable language. In addition, it is continuously evolving and provides a foundational building block for facilitating further knowledge extraction and reasoning.

On the other hand, *Patient*, *Diagnosis* and *Investigation* are concepts present in the *Patient Ontology*. This ontology has the role to maintain patient data as instances of the domain knowledge, and in particular of associations of particular genotypic or phenotypic characteristics to different bone dysplasias. The graph created by the relationships between the above-mentioned concepts represents the input for the following steps of our research methodology.

4.2 Evidence Extraction from the SKELETOME Ontology set – the Level-wise algorithm

Figure 3 depicts the steps performed to create the decision support model. Firstly, data rows are extracted from the SKELETOME knowledge base. These are then transformed in the quantization process to make them suitable for comparison in the level-wise algorithm. Finally, the level-wise search algorithm is applied to discover generalised evidences, which are the stored as instances in the evidence ontology.

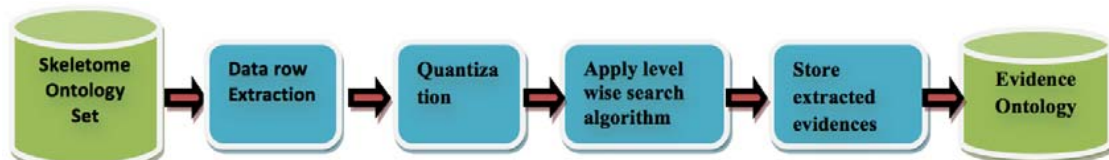


Fig. 3. Decision support model creation steps

As a remark, the evidence extraction process assumes the strict use of positive statements in the data, due to the way in which patient cases (and clinical summaries) are described in this domain. More concretely, the SKELETOME ontology will only contain statements in the form of P implies Q , where P is a set of phenotypes and Q a skeletal dysplasia, without considering negation, e.g., P does not imply Q .

Another, different, remark needs to be made with respect to the evolution of the domain knowledge. The structure of the SKELETOME ontology will naturally evolve in accordance with the advances in the field. This evolution will be reflected both in the instance data (i.e., new patient cases), but also in the evidence extraction. From a technical perspective, we have currently plan to deal with

this evolution by re-generating the evidences as part of a periodical batch process. However, for the future, we will consider incorporating such changes in the generalized evidences in an incremental manner.

Data Row Extraction. This step transforms the instance data present in the SKELETOME knowledge base, which is structured as interconnected graphs, into rows, as required by the level-wise search algorithm. Subsequently, it finds the most appropriate method to perform candidate and item set generation and to find rules within the given dataset, by also taking into account the physical resources associated with such a data-intensive method.

<p>Algorithm: Discovering the trend of the instance data Input: Data Rows from ontology, minimum support, minimum average certainty Output : Itemsets which are strong candidates of Evidences.</p> <ol style="list-style-type: none"> 1. $K=1, S = \{\emptyset\};$ 2. Read the Knowledgebase about which attributes are Dysplasia (action type) and which are Symptoms(Non action). 3. $I_k =$ Select all 1-itemsets which has support and average certainty greater or equal to minimum support and minimum certainty 4. While($I_k \neq \emptyset$) { <ol style="list-style-type: none"> 4.1 $K++;$ 4.2 $C_k =$ Candidate_generation(I_{k-1}) 4.3 CalculateCandidatesSupportAndCertainty(C_k) 4.4 $I_k =$ SelectDesiredItemSetFromCandidates(C_k, S_k, minimum support, minimum average certainty); 4.5 $S = S \cup S_k$ 5. return S <p>procedure Candidate_generation(I_{k-1}) 1. For each Itemset $i_1 \in I_{k-1}$ 1.1 For each Itemset $i_2 \in I_{k-1}$ 1.1.1 Newcandidate, $NC = \text{Union}(i_1, i_2);$ 1.1.2 If Size of NC is k 1.1.2.1 If NC contains one or no action item 1.1.2.1.1 Add it to C_k if every subset of items is frequent. 2. return $C_k;$ </p>

 Procedure CalculateCandidatesSupport(C_k) 1. For each transaction t of Data Rows 1.1 CalculateSupportAndCertaintyFromOneTransactionForCandidates(C_k, t); **procedure CalculateSupportFromOneTransactionForCandidates(C_k, t)** 1. $C_t =$ Find the subsets of t which are candidate 2. For each candidate $c \in C_t$ 2.1 $c.\text{count}++$ 2.2 Calculate Average Certainty **procedure SelectDesiredItemSetFromCandidates(C_k, S_k , minimum support, minimum average certainty)** 1. For each Itemset $c \in C_k$ 1.1 If c contains only non-action items 1.1.1 If $c.\text{support} \geq \text{minimum support}$ and $c.\text{certainty} \geq \text{minimum average certainty}$ 1.1.2 Add it to I 1.2 else if c contains action items 1.2.1 If $c.\text{support} \geq \text{minimum support}$ and $c.\text{certainty} \geq \text{minimum average certainty}$ 1.2.2 Add it to $I \& S_k$ - 2. return I |

Fig. 4. Trend discovery algorithm in instance data

Algorithm : Finding generalized evidences from desired item set.

Input: S(Desired item sets), minimum probabilistic uncertainty

Output: R (set of Evidences)

1. $R = \emptyset$
2. For each $X \in S$
 - 2.1 Symptom set $P = (p_1, p_2, \dots, p_n)\{$
where $p \in X$ and $AC(p) \neq 2\}$
 - 2.2 Dysplasia set $Q = (q_1)\{$
where $q \in X$ and $AC(q) \neq 1\}$
 - 2.3 $CC = \text{CalculateCorrelationCoefficient}(P, Q)$
 - 2.4 $PU = \text{CalculateProbabilisticUncertainty}(CC, X, \text{certainty})$
 - 2.5 if $PU \leq \text{minimum probabilistic uncertainty}$
 - 2.3.1 $P \leftrightarrow Q$ is a valid evidence.
 - 2.3.2 $R = R \cup (P \leftrightarrow Q)$

Fig. 5. Generalized evidence discovery algorithm from the desired item set

Quantization. A second prerequisite to perform evidence extraction using the level-wise search algorithm is to transform the row data into a suitable format. Skeletal dysplasia data types can take multiple forms, ranging from categorical, or Boolean to continuous numerical data, interval, percentage or fraction. Continuous numerical data cannot be compared by direct difference as it may fail in recognizing some of the intrinsic data characteristics. For example, age intervals of equal width (e.g., $0 < \text{age} \leq 10$, $10 < \text{age} \leq 20$) may ignore certain data characteristics due to the ambiguous conventions associated with the patient's age interpretation, i.e., young, adult, or elder. A set of rules is created for each continuous numerical attribute using the knowledge of clinicians and researchers. A rule engine maps continuous numerical data to items using these developed rules. A domain dictionary is used to transform the data for discrete attributes.

Evidence extraction using Level-wise Search. A level wise search algorithm is developed to extract evidences from the SKELETOME knowledge base. The algorithm is based on the following statements:

- A statement ($A \rightarrow B$) is treated as evidence based on the symmetric relationship strength between the antecedent and the consequent.
- Most generalised evidences involve a coherent subset of attributes, instead of implicitly including all possible attributes.
- Symptoms and observations lead to a particular decision and a decision can be a diagnosis or a procedure. All symptoms or observations are part of an antecedent and all diagnoses are part of a consequent.

Once the data has been transformed into a row-based format, the horizontal axis will represent patient instances. Fields composing the horizontal axis will be tagged as Action (representing the diagnosis) and Observation (representing symptoms, lab tests, genetic tests or radiographic features).

In the above-mentioned interval-based crisp quantization, elements near the boundaries of an interval will either be ignored or overemphasized(Kaya and Alhajj 2008). This may lead to losing some of the

underlying meaning of the data. For instance, an interval representing young persons might have a range between 18 and 40 years. In this instance, a person aged 17 would be a 0% representative and an 18 year old person would be 100%. However, the actual difference between these two ages is not that significant. This problem is caused by the sharp boundary between intervals(Kaya and Alhajj 2008). Implementing fuzziness can overcome this problem.

To address this issue, we use fuzzy quantization as an intermediate phase within the overall quantization step. For instance, we partition the values of the Age attribute into three fuzzy sets: low, medium and high. The intervals of low, medium and high could be $\{0-33\}$, $\{27-55\}$ and $\{48-\infty\}$ respectively. In this instance, a person aged 30 years would be a representative of low with a certain degree and a representative of medium with a different degree. The domain experts define the corresponding fuzzy sets and their membership functions.

To have a clear understanding of the final data representation, below we present an example of a fuzzy encoding for a patient who exhibits three symptoms and has been diagnosed with a particular bone dysplasia:

Patient 1: $\{0.8/(s_1, 0.9), 0.8/(s_2, 0.9), 0.8/(s_3, 0.9), 0.8/(D, 0.9)\}$

s_1 = "Symptom X is Low"

s_2 = "Symptom Y is High"

s_3 = "Symptom Z is Low"

D = "Dysplasia BD is Medium"

Generalized evidence represents information inferred from generalized facts. The process of extracting generalized evidences from past patient cases consists of two steps:

1. Discovering the trend of the instance data by finding a desired item set using the level wise search algorithm.
2. Finding generalized evidences from the desired item set.

Step1: This step considers only the fuzzy terms of the fuzzy values, leaving out the membership value of these fuzzy terms. Even so, we consider only the fuzzy values that have membership value greater than a given

threshold (minimum membership value – *mmv*). At the same time, a fuzzy value with more than one fuzzy term will be converted into multiple transactions. For example: {Symptom X{Low, High}, Dysplasia X{High}} will be converted into: {Symptom X{Low}, Dysplasia X{High}} and {Symptom X{ High }, Dysplasia X{High}}.

Figure 4 details the trend discovery algorithm, where K is the size of the item set, S is set of the desired item set and $\{S_1, S_2, \dots, S_y\}$ are the desired item sets of length $\{1, 2, \dots, Y\}$. Also, $\{C_1, C_2, \dots, C_q\}$ are the candidate item sets of length $\{1, 2, \dots, Q\}$ and $\{I_1, I_2, \dots, I_t\}$ are frequent item sets of length $\{1, 2, \dots, t\}$.

Calculating Support and Average Certainty of an item set. If an item set has the items $I = \{i_1, i_2, i_3, \dots, i_n\}$, there are m transactions in the knowledge base, we calculate the support and average certainty of the item set using the formulae below:

$$\text{Support or Probability}(I) = \frac{\sum_{k=1}^m I \in t_k}{\text{total number of transactions}}$$

$$\text{Average certainty}(I) = \frac{\sum_{t=1}^m \prod_{j=1}^n \text{certainty}(i_j)}{\text{total number of transactions contains } I}$$

Step 2: Figure 5 lists the algorithm for finding generalized evidences from the desired item set. Firstly, we reduce the desired item sets $\{S_2, S_3, \dots, S_y\}$ only to those that have a skeletal dysplasia associated.

We then partition the symptoms and dysplasia of each item set into two sets: an action item set containing the dysplasias and a non-action item set containing the symptoms, with the symptom set of each of the initial item sets related to the dysplasia associated with the respective item set. Each of these relationships will represent generalized evidence. AC is the function that determines the type of an item, i.e., action or non-action.

$AC(x) = 2$ if it is non-action item/symptom

$AC(x) = 1$ if it is action item/dysplasia

Subsequently, we calculate the correlation coefficient between the action item set and the non-action item set of the evidence, and the probabilistic uncertainty by multiplying the resulting correlation coefficient and the average certainty. The generalized evidences having the probabilistic uncertainty value greater than a certain threshold will be considered as final result.

Ranking the generalized evidences could have been performed also by using *confidence*, which is another widely adopted interestingness metric. However, confidence does not account for the consequent frequency with the antecedent. In order to rank generalised medical evidences, we need a metric that takes into account frequency in both directions, i.e., the consequent frequency with the antecedent and the antecedent frequency with the consequent.

Correlation coefficient calculation. In a given medical relationship $s \rightarrow t$, s is a group of medical items where each item is constrained to appear in antecedent and t is a

group of medical attributes where each item appears in consequent. Moreover $s \cap t = \emptyset$. For this relationship, the support is defined as $\text{support} = P(s, t)$ and the confidence is defined as $\text{confidence} = P(s, t)/P(t)$, where P is the probability.

The correlation coefficient (also known as the Φ -coefficient) measures the degree of relationship between two random variables by looking at the degree of linear interdependency. It is defined by the covariance between the two variables divided by their standard deviations:

$$\rho_{st} = \frac{\text{Cov}(s, t)}{\sigma_s \sigma_t}$$

Here $\text{Cov}(s, t)$ represents the covariance of the two variables and σ_x and σ_y stand for standard deviation. The covariance measures how two variables change together:

$$\text{Cov}(s, t) = P(s, t) - P(s)P(t)$$

Standard deviation is the square root of its variance and variance is a special case of covariance when the two variables are identical.

$$\begin{aligned} \sigma_s &= \sqrt{\text{Var}(s)} = \sqrt{\text{Cov}(s, s)} = \sqrt{P(s, s) - P(s)P(s)} \\ &= \sqrt{P(s) - P(s)^2} \\ \sigma_t &= \sqrt{P(t) - P(t)^2} \\ \rho_{st} &= \frac{P(s, t) - P(s)P(t)}{\sqrt{P(s) - P(s)^2} \sqrt{P(t) - P(t)^2}} \end{aligned}$$

$P(s, t)$ represents the support of an item set that consists of both s and t . Let the support of the item set be S_{st} . $p(s)$ and $p(t)$ will represent the support of antecedent s (S_s) and antecedent t (S_t), respectively. The value of S_{st} , S_s and S_t are computed during the desired item set generation. Using these values, we can calculate the correlation of every medical relationship among diverse groups of medical items. The correlation value will indicate the medical researcher how strong a medical relationship is from the perspective of historical data.

$$\rho_{st} = \frac{S_{st} - S_s S_t}{\sqrt{S_s - S_s^2} \sqrt{S_t - S_t^2}}$$

Hence, creating an association rule from the values of S_{st} , S_s and S_t provides us with a single metric, correlation coefficient, to represent the degree of relatedness between the antecedent and the consequent. For each medical relationship or rule, this metric is used to indicate the degree of relatedness between different groups of items.

ρ_{st} takes values between -1 and +1. If two variables are independent then ρ_{st} is 0. When ρ_{st} is +1 the variables are considered perfectly positively correlated. A positive correlation represents an evidence of a general tendency of relatedness between a group of attribute values s and a group of attribute values y of a particular patient. The more positive the value is, the stronger the relationship is. When ρ_{st} is -1 the variables are considered perfectly negatively correlated.

Storing the Evidences. The fuzzy evidences discovered in the previous stage are stored in the Evidence Ontology

together with their corresponding probabilistic uncertainty.

4.3 Evidence Ontology

The Evidence Ontology (see Figure 6) is an OWL ontology we have built to represent uncertain generalised evidences. The ontology captures both fuzziness / vagueness and probabilistic uncertainty by re-using concepts from Fuzzy Theory, such as fuzzy value, fuzzy variable, fuzzy set, membership value, fuzzy term and probabilistic uncertainty. It enables the expression of uncertain information / evidence via ontological concepts and helps in simplifying knowledge representation in OWL.

The structure of the ontology comprises two conceptual layers:

- A fuzzy theory layer describing the features of fuzzy theory, such as, fuzzy term and membership value characteristics, and
- A conditional statements/conditional expressions layer modelling past skeletal dysplasia evidences with probabilistic uncertainty.

The SKELETOME ontology set is hence extended, via the Evidence Ontology, to enable analytical uncertainty reasoning on skeletal dysplasia cases. The resulting representation allows a powerful set of uncertainty operations while not introducing any inconsistency in the host ontology.

The syntax of the Fuzzy theory layer is based on OWL 2.0, while the semantics is based on the theory of fuzzy sets. Evidence are of the form antecedent \rightarrow consequent with an associated probabilistic uncertainty, where antecedent consists of a set of Fuzzy Values of skeletal dysplasia symptoms, while the consequent is a set of Fuzzy Values of bone dysplasias. Range, domain, cardinality and functionality axioms are employed in the Evidence Ontology to keep the integrity of the Fuzzy theory, conditional statements and probabilistic uncertainty semantics. The object properties are exposed to establish association/relations between concepts, and data type properties are exposed to describe the attributes of concepts.

The Evidence Ontology has 8 main classes representing different concepts of fuzzy theory and conditional statements, listed below:

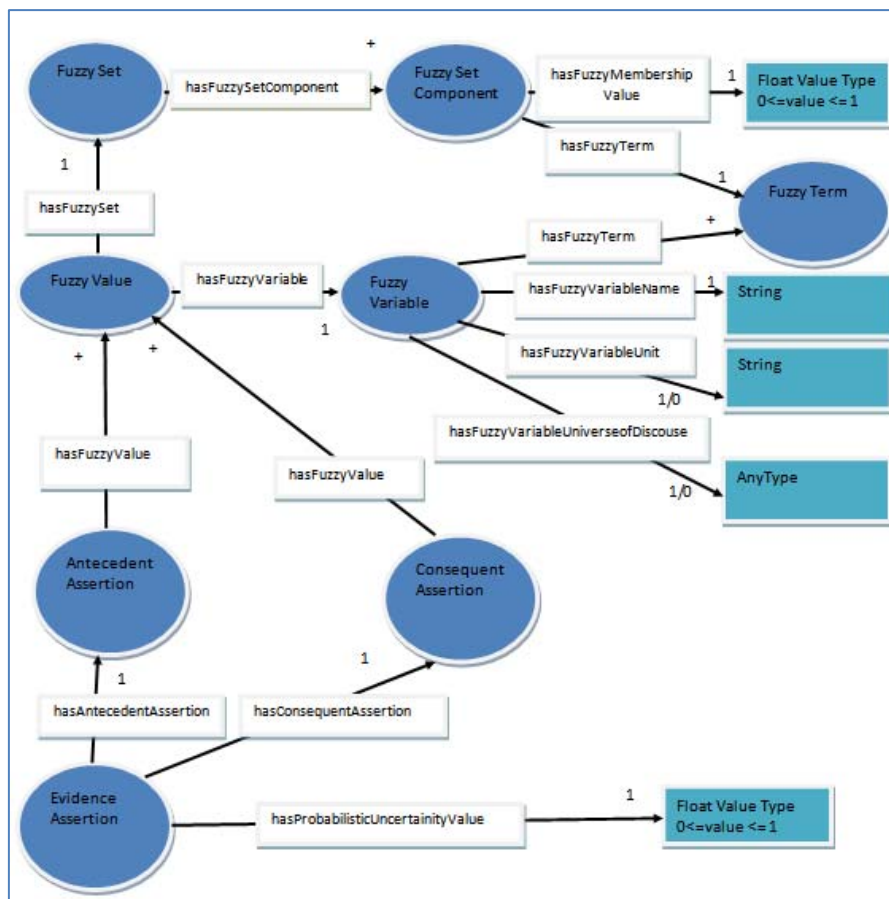


Fig. 6. Evidence Ontology concepts

Evidence Assertion represents evidences extracted from existing patient cases. An Evidence Assertion instance comprises an Antecedent Assertion, a Consequent Assertion and a Probabilistic Uncertainty.

Antecedent Assertion represents the antecedent part of an evidence. Its instances are composed of a set of fuzzy values.

Consequent Assertion represents the consequent part of an evidence. Its instances are composed of a set of fuzzy values.

Fuzzy Variable represents a fuzzy variable from the fuzzy theory. Fuzzy variables usually consist of a *name* (e.g., age), *terms* (e.g., child, young or blue), an *unit* (e.g., years) and the *universe of discourse* (e.g., 0-200).

Fuzzy Term represents a fuzzy term from the Fuzzy Theory, which is used as part of Fuzzy Set and Fuzzy Variable. For example, a fuzzy variable SymptomA may have fuzzy terms like Low, High or Medium. A Fuzzy Set may have the fuzzy term Medium with membership value 0.8. Each Fuzzy Term has a membership function.

Fuzzy Value represents the corresponding fuzzy value from the Fuzzy Theory and is value of a feature in a fuzzy sense instead of the crisp sense. After fuzzification, a numeric value converted in to Fuzzy Value and has two parts: a **Fuzzy Set** (i.e., the numeric value in the Fuzzy Terminology) and a **Fuzzy Variable** (Metadata about the Fuzzy terminology)

Fuzzy Set. Every member of a Fuzzy Set has membership degrees. A Fuzzy set instance is composed of a set of **FuzzySetComponents**.

FuzzySetComponent / FuzzySetMember represents a fuzzy element of the Fuzzy set theory. A FuzzySetComponent instance is composed of a fuzzy term and a corresponding membership value.

Currently, the Evidence Ontology has 8 classes, 10 object properties, 5 data type properties and no instances.

5 Evaluation Plans

To date, we have developed the Evidence ontology and identified a mechanism for inducing evidences. The next phase of the project involves evaluating these two aspects and refining/optimizing them based on the results.

Task-based evaluations (Porzel and Malaka 2004) will be used to assess the capability of the Evidence ontology to represent the generalized uncertain evidence. A set of use-cases, formulated as parameterized test questions and answer keys will be leveraged to characterize the ontology in terms of accuracy, insertion errors, deletion errors and substitution errors.

Similarly, to quantitatively assess the quality of the evidence extraction process, we will measure the evidence retrievability (recall) (Gupta, Fang et al. 2008) and the evidence spuriousness (precision) (Gupta, Fang et al. 2008). Evidence retrievability measures how well the underlying trends in past data have been discovered. Although retrievability provides a good estimate of the fraction of detected patterns in the data, it does not provide an estimate of the quality of the found patterns. The quality of a pattern is measured using spuriousness, which quantifies the number of items in the pattern that are not associated with the matching base pattern.

6 Conclusion

No prior research has investigated the induction of generalised evidences from immature domain knowledge, specifically in the skeletal dysplasia domain. This domain raises two important challenges with respect to

developing decision support methods: (1) the absence of a wealth of background knowledge that would enable deductive reasoning, and (2) the sparseness of skeletal dysplasias data.

In this paper we have proposed a method for inducing generalized evidences from the existing patient cases, via a level-wise algorithm. The resulting knowledge is stored in the Evidence Ontology, which not only provides the foundational model for the development of appropriate decision support methods, but also a means for sharing generalised evidences in an interoperable manner.

Future work will focus on firstly evaluating the Evidence ontology and the level-wise algorithm. Following the evaluation and refinement of these components, the next step involves using instances of the Evidence ontology in conjunction with evidential reasoning, for automated diagnosis and key disease features extraction.

7 References

- Agrawal, R., T. Imieli ski, et al. (1993). Mining association rules between sets of items in large databases, ACM.
- Allemang, D. and J. A. Hendler (2008). Semantic web for the working ontologist: modeling in RDF, RDFS and OWL, Morgan Kaufmann.
- Andreasik, J., A. Ciebiera, et al. (2010). ControlSem—distributed decision support system based on semantic web technologies for the analysis of the medical procedures. 3rd Conference on Human System Interactions (HSI), IEEE.
- Begum, S., M. U. Ahmed, et al. (2010). "Case-Based Reasoning Systems in the Health Sciences: A Survey of Recent Trends and Developments." IEEE Transactions on Systems, Man, and Cybernetics--Part C: Applications and Reviews, IEEE.
- Berners-Lee, T., J. Hendler, et al. (2001). "The Semantic Web - A new form of Web content that is meaningful to computers will unleash a revolution of new possibilities." Scientific American 284(5): 34-+.
- Bobillo, F. and U. Straccia (2008). fuzzyDL: An expressive fuzzy description logic reasoner, IEEE.
- Bobillo, F. and U. Straccia (2009). "An OWL Ontology for Fuzzy OWL 2." Foundations of Intelligent Systems, Proceedings 5722: 151-160.
- Bobillo, F. and U. Straccia (2010). "Fuzzy Ontology Representation using OWL 2." Arxiv preprint arXiv:1009.3391.
- Castillo, O., P. Melin, et al. (2007). "Type-2 fuzzy logic: Theory and applications." Grc: 2007 Ieee International Conference on Granular Computing, Proceedings: 145-150.
- Chan, K., S. Ling, et al. (2011). "Diagnosis of hypoglycemic episodes using a neural network based rule discovery system." Expert Systems with Applications.
- da Costa, P. C. G., K. B. Laskey, et al. (2008). "PR-OWL: A Bayesian Ontology Language for the

- Semantic Web." *Uncertainty Reasoning for the Semantic Web I* 5327: 88-107.
- De, S. K., R. Biswas, et al. (2001). "An application of intuitionistic fuzzy sets in medical diagnosis." *Fuzzy Sets and Systems* 117(2): 209-213.
- Doddi, S., A. Marathe, et al. (2001). "Discovery of association rules in medical data." *Medical Informatics and the Internet in Medicine* 26(1): 25-33.
- Gadaras, I. and L. Mikhailov (2009). "An interpretable fuzzy rule-based classification methodology for medical diagnosis." *Artificial Intelligence in Medicine* 47(1): 25-41.
- Goossen, F., W. IJntema, et al. (2011). News personalization using the CF-IDF semantic recommender, ACM.
- Gu, H. M., X. Wang, et al. (2007). "Building a fuzzy ontology of edutainment using OWL." *Computational Science - ICCS 2007, Pt 3, Proceedings* 4489: 591-594.
- Gupta, R., G. Fang, et al. (2008). Quantitative evaluation of approximate frequent pattern mining algorithms, ACM.
- Hudson, D. L. (2006). *Medical Expert Systems. Encyclopedia of Biomedical Engineering*, John Wiley and Sons.
- Hussain, S., S. R. Abidi, et al. (2007). "Semantic web framework for knowledge-centric clinical decision support systems." *Artificial Intelligence in Medicine, Proceedings* 4594: 451-455.
- Kaya, M. and R. Alhajj (2008). "Online mining of fuzzy multidimensional weighted association rules." *Applied Intelligence* 29(1): 13-34.
- Lee, C. S. and M. H. Wang (2011). "A Fuzzy Expert System for Diabetes Decision Support Application." *Ieee Transactions on Systems Man and Cybernetics Part B-Cybernetics* 41(1): 139-153.
- Lisi, F. A. (2006). "A methodology for building Semantic Web Mining systems." *Foundations of Intelligent Systems, Proceedings* 4203: 306-311.
- Lisi, F. A. (2006). "Practice of inductive reasoning on the Semantic Web: A system for Semantic Web mining." *Principles and Practice of Semantic Web Reasoning* 4187: 242-256.
- Maedche, A. and S. Staab (2000). "Discovering conceptual relations from text." *Ecai 2000: 14th European Conference on Artificial Intelligence, Proceedings* 54: 321-325.
- Noy, N., A. Rector, et al. (2006). "Defining n-ary relations on the semantic web." *W3C Working Group Note* 12.
- Pan, H. W., J. Z. Li, et al. (2005). "Mining interesting association rules in medical images." *Advanced Data Mining and Applications, Proceedings* 3584: 598-609.
- Pan, J. Z., G. Stamou, et al. (2007). Expressive querying over fuzzy DL-Lite ontologies, Citeseer.
- Papageorgiou, E. I., N. Papandrianos, et al. (2009). "Fuzzy Cognitive Map Based Approach for Assessing Pulmonary Infections." *Foundations of Intelligent Systems, Proceedings* 5722: 109-118.
- Patkos, T., I. Chrysakis, et al. (2010). "A Reasoning Framework for Ambient Intelligence." *Artificial Intelligence: Theories, Models and Applications*: 213-222.
- Paul, R. and A. S. M. Hoque (2010). Mining irregular association rules based on action & non-action type data. *Fifth International Conference on Digital Information Management (ICDIM)* Thunder Bay, ON Canada, IEEE.
- Porzel, R. and R. Malaka (2004). A task-based approach for ontology evaluation, Citeseer.
- Prcela, M., D. Gamberger, et al. (2008). "Semantic web ontology utilization for heart failure expert system design." *Studies in health technology and informatics* 136: 851.
- Robinson, P. N., S. Köhler, et al. (2008). "The Human Phenotype Ontology: a tool for annotating and analyzing human hereditary disease." *The American Journal of Human Genetics* 83(5): 610-615.
- Ross, T. J. (2010). "Fuzzy logic with engineering applications." Willy-India.
- Shadbolt, N., W. Hall, et al. (2006). "The Semantic Web revisited." *Ieee Intelligent Systems* 21(3): 96-101.
- Sheela, L. J. and V. Shanthi (2009). "DIMAR - Discovering Interesting Medical Association Rules from MRI Scans." *Ecti-Con: 2009 6th International Conference on Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology, Vols 1 and 2*: 614-618.
- Simou, N. and S. Kollias (2007). Fire: A fuzzy reasoning engine for imprecise knowledge, Citeseer.
- Sioutos, N., S. Coronado, et al. (2007). "NCI Thesaurus: A semantic model integrating cancer-related clinical and molecular information." *Journal of Biomedical Informatics* 40(1): 30-43.
- Stoilos, G., G. Stamou, et al. (2005). Fuzzy OWL: Uncertainty and the semantic web, Citeseer.
- Straccia, U. (2008). "Managing uncertainty and vagueness in description logics, logic programs and description logic programs." *Reasoning Web*: 54-103.
- Straccia, U. (2010). *An Ontology Mediated Multimedia Information Retrieval System*, IEEE.
- Straszeka, E. (2006). "Combining uncertainty and imprecision in models of medical diagnosis." *Information Sciences* 176(20): 3026-3059.
- Stumme, G., A. Hotho, et al. (2006). "Semantic Web Mining - State of the art and future directions." *Journal of Web Semantics* 4(2): 124-143.
- Warman, M. L., V. Cormier Daire, et al. "Nosology and classification of genetic skeletal disorders: 2010 revision." *American Journal of Medical Genetics Part A*.
- Weng, C. H. and Y. L. Chen (2010). "Mining fuzzy association rules from uncertain data." *Knowledge and Information Systems* 23(2): 129-152.

Gaining Insight from Patient Journey Data using a Process-Oriented Analysis Approach

Lua Perimal-Lewis

School of Computer Science, Engineering and Mathematics
Flinders University of South Australia
GPO Box 2100, Adelaide 5001, South Australia
lua.perimal-lewis@flinders.edu.au

Shaowen Qin

School of Computer Science, Engineering and Mathematics
Flinders University of South Australia
GPO Box 2100, Adelaide 5001, South Australia
shaowen.qin@flinders.edu.au

Campbell Thompson

Medicine
The University of Adelaide
5005 South Australia

campbell.thompson@adelaide.edu.au

Paul Hakendorf Clinical

Epidemiology Unit, Flinders
Medical Centre
Flinders Drive, Bedford Park 5042 South Australia

Paul.Hakendorf@health.sa.gov.au

Abstract

Hospitals are continually struggling to cater for the increasing demand for inpatient services. This is due to increased population, aging, and the rising incidence of chronic diseases associated with modern life. The high demand for hospital services leads to unpredictable bed availability, longer waiting period for acute admission, difficulties in keeping planned admission, stressed hospital staff, undesirable patient and family experience, as well as unclear impact on the quality of care patients receive. This study aims to gain insight into patient journey data to identify problems that could cause access block. Process mining techniques combined with statistical data analysis are adapted to discover inpatient flow process patterns and their correlation with patient types, ward types, waiting time and Length of Stay (LOS). Open source process mining software, ProM, is used in this study. The study is done in collaboration with Flinders Medical Centre (FMC) using data from their Patient Journey Database.

Keywords: patient flow, inpatient journey analysis, process mining, length of stay

1 Introduction

Australian Public Hospitals are generally patient-centred organisations. Hospital managements need to constantly look for ways to improve their patient care processes in order to address challenges caused by the ever-increasing demand for hospital services. One such challenge, which is often in media's limelight, is the Emergency Department (ED) overcrowding also known as access block.

Fatovich (2002) describes ED overcrowding as a worldwide problem where an ED is unable to provide timely emergency care and as a consequence ambulances are instructed to divert to another facility. The reason for ambulance diversion is simply because of the lack of capacity to safely attend to newly arrived patients.

O'Connell, Bassham et al. (2008) asserts that ED congestions are intensified by regular failure to manage processes involved in progressing patients through the hospital. According to O'Connell, Bassham et al. (2008) better inpatient management and better patient flow will mitigate some of the issues faced by ED.

The study aims to investigate the Flow of Patient or Care Flow within a hospital with the understanding that smooth flow would reduce bottlenecks in hospitals. Smooth patient flow will have positive influence to the overall hospital capacity that would consequently ease ED over-crowding and other systemic issues in a hospital such as bed availability.

2 Background

Public Hospitals are required to conform to certain Key Process Indicators (KPIs). Such conformances are essential because of the competitive nature of government funding and the need to provide justified information to tax-payers. Hospitals do already have mature processes and the ability to report and monitor performances with statistics. However, such data does not show how to improve a process. The term Clinical Process Re-engineering has become a common term when referring to clinical process improvements. Clinical Process Re-engineering could be considered similar to Business Process Re-engineering (BPR). Both initiatives focus on continuous improvements of business processes or clinical processes to gain “competitive advantage”. However, this “competitive advantage” from a hospital’s perspective is service oriented and increasingly becoming patient-centred, endeavouring to service the health care needs of its population. BPR aims at improving core business process and in the hospital setting one of the core business processes can be regarded as the patient journey.

Clinical process redesign is the use of process redesign and change management to health care (Ben-Tovim, Dougherty et al., 2008). The activities in the redesign process focuses on the patient’s perspective. The aim of clinical process redesign is to harmonise the poorly coordinated patient journey as they move across multiple departments, making them simpler whilst looking at the overall design of the clinical processes. Clinical process redesign takes a holistic approach by looking at a wider area in the redesign process, which requires continuous fine-tuning and adjustments to constantly adapt to the ever-transient nature of the hospital system.

Many hospitals around the world have adopted “lean thinking” methodology to improve patient care and have seen significant improvements in their clinical processes. One such success story, according to Richard C (2011), is experienced at Denver Health in America where over the last 5 years, the lean methodology has contributed to revenue increase and decrease in expenses. FMC’s Redesigning Care program has also adopted the concept of “lean thinking”, which enabled the hospital to provide safer and more accessible care (Ben-Tovim, Bassham et al., 2008). Whilst “lean thinking” has contributed to better understanding of the patient flow process and therefore better co-ordination, implementing this concept alone in isolation might not be enough to relieve the hospital-wide crisis where bed availability is still a concern.

Health Care data analysis is traditionally done using various statistical techniques in order to report and hopefully forecast health care performances. New approaches in health care modelling are emerging where more than one technique and approach are used to discover hidden information that might not be easily discovered from one approach. Combinations of techniques are used to complement each other.

Ceglowski, Churilov et al. (2007) proposed combining Data Mining techniques and discrete event simulation for identifying bottlenecks in the patient flow between ED and a hospital ward by providing insight into the complex

relationship between patient urgency, treatment and disposal and the occurrence of queues for treatment.

The use of Decision Support System (DSS) in health care is wide spread. DSS in Health Care industry could be divided into 2 broad categories. One category of DSS is used to help physicians with their day-to-day decision-makings. An example is a DSS based on clinical practice guideline in the management of diabetic patients (Lobach and Hammond, 1997). The other category of DSS is used by hospital management to make decisions for better hospital resource management. The fundamental information needed for such a system is based on the outcomes of some method of data analysis and modelling. The closer the outcome is in depicting the real scenario the better the DSS output.

Based on this background information and knowledge of common strategies being used by hospitals to better manage their patient journeys, this study aims to build on the existing resources and knowledge gained from some of these practices. The process mining activities proposed under this study will complement the “lean thinking” strategies currently being practiced at FMC. An in-depth evaluation of the patient flow processes using the same Patient Journey Database will aid in identifying process bottlenecks hidden within the statistics.

3 Data

The Patient Journey Database from FMC records information on the journey or movement of a patient from the time of admission to the time of discharge. Therefore, it only contains information on inpatients or officially admitted patients.

Each admission is given a unique journey number that would remain the same until discharge. Each movement of the patient from one ward to another ward is recorded with a timestamp, so at any point the “start time” in a ward and the “end time” in a ward is known together with the ward name being occupied. Each journey is also linked to the doctor treating the patient. The wards are a subset of Units and the Units are a subset of Division. Each journey is also given a status of Inliers, Outliers or Inliers/Outliers. An individual patient could have multiple admissions at various time frames and each of this admission will be allocated with new unique journey number. Timestamp for Admission is the combination the “Date” field and the “Admission Time” field. Timestamp for Discharge is the combination of “Date” field and “Discharge Time” field. The table below shows the relevant fields taken from the Patient Journey Database pertinent to process mining (see Figure 1).

Journey ID	Patient Unique ID	Date	Admission Time	Discharge Time	Ward	Unit	Age in Years	Doctor
2	100	03-Mar-04	16:49	23:59	"CIC"	"CIC"	59	"Dr. AB"
2	100	04-Mar-04	0:00	11:29	"CIC"	"CIC"	59	"Dr. AB"
2	100	04-Mar-04	11:29	15:45	"TL"	"CIC"	59	"Dr. AB"
3	209	01-May-06	22:08	23:05	"FMC"	"RESP"	87	"Dr. CD"
3	209	01-May-06	23:05	23:59	"SA"	"RESP"	87	"Dr. CD"
3	209	02-May-06	0:00	20:17	"SA"	"RESP"	87	"Dr. CD"

Figure 1: Snippet of data used for process mining

One of the notable criteria of this set of data is the ability to expand or link the patient journey with any other data from another database of interest. For example, the journeys could be linked to a separate database within the hospital that might only collect disease, drug or cost related information. The individual patients are not identifiable at any point.

The repository of data starts from the 1st of April 2003. New or latest data can be easily added to the dataset if deemed necessary for a particular analysis. The patient journey data will be clustered or grouped using appropriate parameters best for the particular analysis or knowledge discovery at hand.

4 Proposed Methodology – A Process Oriented Analysis Approach

4.1 Process Mining

Process mining in healthcare is still in its infancy. Mans, Schonenberg et al. (2008) used process mining techniques to better understand different clinical pathways taken by various groups of patients and used the technique to identify bottlenecks. Rebugue and Ferreira (2011) concluded that although process mining techniques have been proven in some instances as being successful in mining health data, there are still room for improvements to identify the right algorithm to handle noise in the data, complexity of data and the ad hoc nature of health data.

Process is embedded in every aspect and at every level of one's daily routines. The output of a process is a result that could be either favourable or not so favourable. Exploring the various activities within a process contributes to deeper knowledge, understanding and discoveries of the intricacies of what actually happens within a process that finally produces the result.

Health care industries are data rich as a result of embracing the notion of paperless system in a big scale. This has introduced some positive challenge to researchers in discovering knowledge from the use of such data. The notion of efficient patient care providing patient-centred approach has seen the emergence of various Health Information Systems as stated by Vezyridis, Timmons et al. (2011). Electronic Patient Management or Tracking Systems have all become not only common but essential systems for any hospital. These information systems store invaluable information that can be used for knowledge discoveries. Process mining enables the discovery of knowledge regarding a process. Process mining uses event or process logs to extract information regarding a process as it has taken place (van der Aalst and Weijters, 2004). These process / event logs do not have to necessarily originate from a Workflow Management System. A process log could be derived from a dataset that contains an order of events which could be used to assemble a process model that portrays the activity of the subject matter. This concept forms the basis of the methodology in this study.

van der Aalst, Reijers et al. (2007) stated that process mining aims to construct a process model from observed behaviour from a process perspective, organisational perspective or a case perspective; and the most significant output of process mining is the discovery of the main

process flow. Rozinat, Wynn et al. (2009) stated that process mining normally creates a static model that could be used by the users of the systems to reflect on the process.

In the context of this study, the aim of process mining is to discover the various paths taken by inpatients moving from ED to other ward/wards. This study is also looking to discover the most frequent path taken and the associated parameters concerned with this path. Another aim is to gather the "network" information of doctors and ward where the information presented will indicate a doctor's ward pattern. Braitberg (2007) argues that improving operational efficiency based on average bed occupancy is too weak to predict a complex hospital system and the dynamic nature of patient flow. Hence by using process mining the aim is to measure operation efficiency by calculating the throughput time for each ward, doctor and a cluster or group of patients. Another aim is to analyse inpatient journey or flow and to identify the streams that often cause delays for patients to flow outside of ED.

4.2 Data Analysis

The aim of data analysis in this study is to analyse inpatient journeys to reveal macro characteristics of different cluster of patients. This study focuses on LOS of Outliers and Inliers. Outlier is an inpatient who is admitted to Wards other than the Home Ward. An Inlier is an inpatient who is admitted to a Home Ward. Home Ward is a ward that is equipped with appropriate medical team and specialised equipment to treat the patient's primary disease at the time of admission.

Each patient journey is classified using Inliers, Outliers or Inliers/Outliers statuses. Inliers are patient journeys with 100% of the time at hospital spent in an Inlier ward. Outliers are patient journeys with 100% of the time at hospital spent in an Outlier ward. Inliers/Outliers consist of patient journeys where part of the hospital time is spent in an Inlier ward and the other part in an Outlier ward.

The outcome of this analysis will reveal if LOS is influenced by statuses. Process mining techniques will be used on this cluster of patient journeys for flow bottleneck identifications.

4.3 Collaboration with Clinicians

Regular contact with the clinicians enabled the portrayal of the actual undertakings at the hospital, which enhanced the quality of understanding as well as the relevance of the derived process mining and data analysis results. Clinicians' insight was also invaluable in explaining the processes within the hospital's context and the ways to interpret the result of the analysis and knowledge discovery.

The clinicians play a major role in identifying patient journey clusters or pool for a particular analysis and process mining that would give the best representation of activities taking place within a particular cluster of patients.

5 Preliminary Work

5.1 ProM (Process Mining) Toolkit

The Patient Journey Database is a well-maintained database containing historical records of the various movements of a patient during an admission episode. Information contained in this database is useful not only to construct a process model but also to discover hidden knowledge regarding a Patient's Journey. The following section demonstrates some of the information revealed from undertaking process mining on this database using ProM.

Records from the Patient Journey Database were pre-processed and formatted into MXML format so the data could be read by the ProM toolkit. MXML is an extension of Extensible Markup Language (XML). The paragraphs below show some of the output of ProM upon processing the input data from the Patient Journey Database. For the purpose of ensuring that the pre-processing exercise is done accurately and the output of ProM is in accordance to what is aimed to be achieved, a small subset of data has been used for the ease of manual confirmation of the results. Five patient journeys have been included in this analysis.

Basic Log Statistics reveal statistical information from the data set. Each unique ward's execution or processing information is represented. The statistical information that can be obtained is on minimum time a patient spent in the ward, maximum time a patient spent in a ward, arithmetic mean, standard deviation, geometric mean, sum and number of times the ward has been used.

Pattern Analysis output for the 5 patient journeys is shown in Figure 2. The pattern reveals 2 distinct flow or movement from ward to ward. It could be concluded from the output below, that ward "FMC" or ED, is frequently used. The aim is to derive common patterns used by a cluster of patients so this information could be used for better capacity planning.

Pattern Analysis Result	
Table view	
Name	Activities
Pattern_0	"6D"-"ANG"-"FMC"
Pattern_1	"3D"-"CCU"-"TL"-"ANG"-"FMC"

Figure 2: Pattern Analysis – Patient Journey Flow

Figure 3 depicts the frequency of usage for each Ward that was involved in the 5 patient journeys. A quick glance at the diagram shows that one particular ward is being used exponentially more than the other wards. To benefit from this result, it is important to have a pre-defined cluster or group of patients to be studied to discover the connections between patient streams or clusters of patients and the frequency of ward usage.

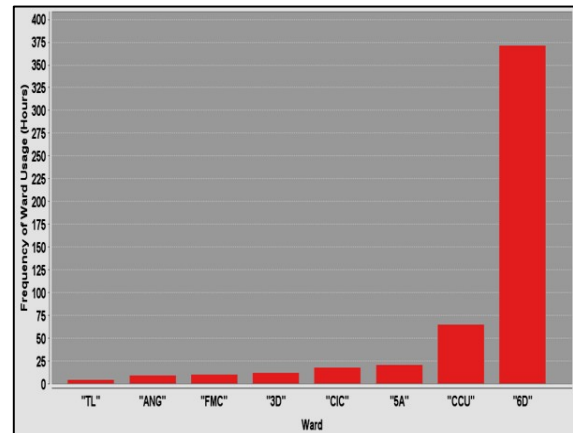


Figure 3: Frequency of Ward Usage

Individual Journeys could also be further analysed to depict the LOS (see Figure 4). This information combined with diagnostic information could be used to reveal commonalities in overall LOS and type of disease, which would contribute to better capacity planning.

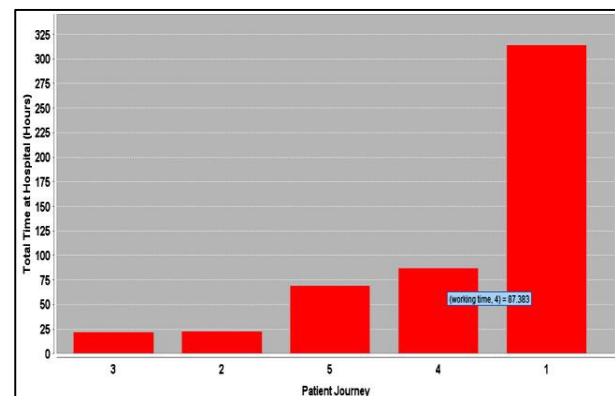


Figure 4: Journey Length of Stay

Performance Sequence Analysis facilitates the assessment of performance of the flow of journey categorised as patterns and performance of each ward involved in the pattern. Mean throughput time for each block under a pattern could be discovered. This analysis will aid in discovery of process patterns that could potentially cause issues to the system. For example identifying wards with high throughput would enable investigation into the cause of such behaviour. Similar analysis could be carried out on doctors, which will show the transfer of work between 2 doctors, throughput time as well as the frequency of certain behaviour or pattern relating to a doctor. Figure 5 is a sequence diagram showing the movements of patient between wards for the 5 journeys included in this analysis. The labelled boxes on the top of the diagram are the names of "Wards" involved in these journeys. The numbered scale on the left is the time taken for the journeys to move from one "Ward" to another "Ward". Measurement unit is in minutes. The coloured lines indicate 5 different journeys. It is also apparent from the diagram, that 2 journeys follow the same path where the movement is from ward "FMC" to ward "6D" to ward "ANG". The time taken for each movement could be obtained by hovering mouse on top of the horizontal lines between "Wards".

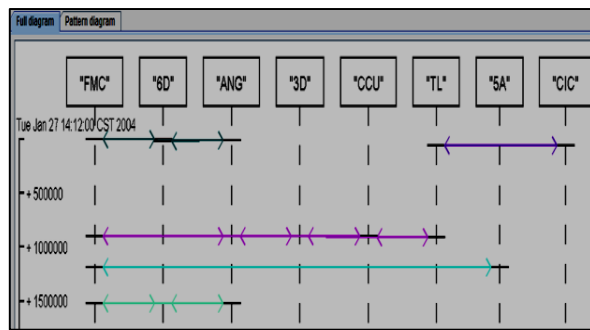


Figure 5: Sequence Diagram

The sequence diagram is further processed into a Pattern Diagram. Journeys with same path or pattern are grouped together. For example, in this set of data, 2 journeys have moved from ward "FMC" to ward "6D" and then to ward "ANG". These 2 journeys are grouped together forming 1 pattern and the rest of the journeys following 3 varied paths forming the other patterns. Figure 6 shows part of the Pattern Diagram for the five journeys. The pattern and analysis results shown are for the 2 journeys with the same path.

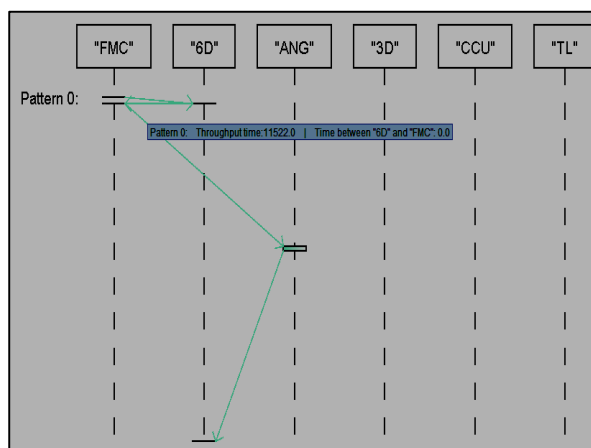


Figure 6: Pattern Diagram

Further information for the Pattern shown in Figure 6 is provided in Figure 7. The mean throughput time for each pattern is derived. Measurement unit is in minutes. The mean throughput time for the patterns could also be obtained with mouse hover on the horizontal lines between "Wards". The frequency of occurrence for each pattern is also listed. Patterns with higher frequency indicate that such a movement pattern in patient journey is common and the behaviours of such a movement or patient journey could be further analysed. For example, this information could be used to investigate the correlation of patterns and bottlenecks.

Pattern 0:

Frequency: 2

	Throughput time
avg	11522.0
min	4170.0
max	18874.0
stdev	10397.29811

Figure 7: Patter Diagram Information

Figure 8 depicts the patient flow derived from the 5 journeys. Each of the square boxes represents a Ward. The number below each Ward indicates the frequency of the Ward usage. The integers next to the arrows represent the number of journeys that have used that path. The decimal numbers next to the arrows represents the dependency relationship between the 2 wards involved. A decimal number close to 1 indicates a strong dependency relationship between the 2 wards. Strong dependency relationship shows the flow in that path is likely to happen.

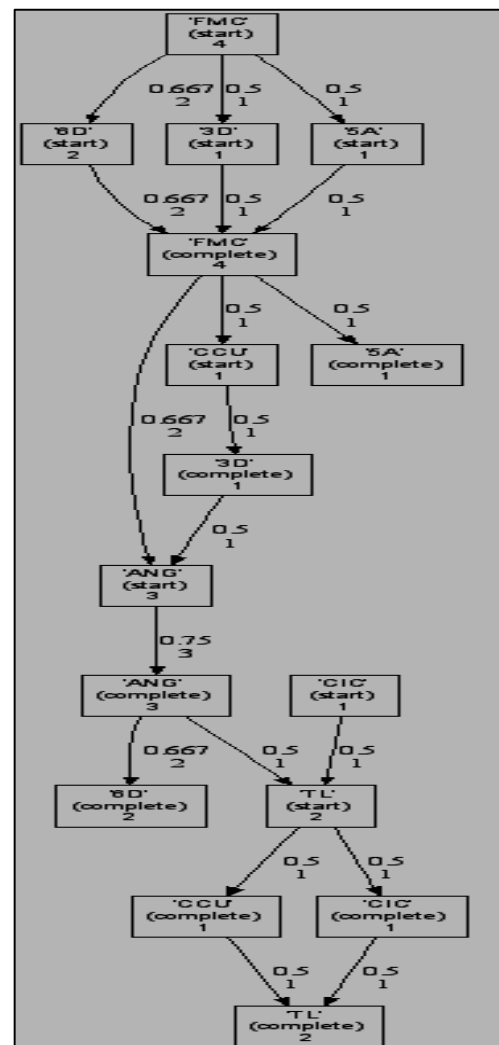


Figure 8: Patient Journey Control Flow Discovery

Figure 8 shows the complete journey derived for the 5 patient's journey, which reinforces the fact that the flow

discovery is complex and we need to classify and cluster the information properly before attempting to discover the flow pattern through process mining. The output of a flow pattern, when there are many more journeys involved, would become extremely complex and often meaningless without sound knowledge of the best way to derive a pool of data to be analysed. Therefore it is vital to invest time in deriving an appropriate cluster of patient journeys and pool size which will ultimately reveal useful knowledge and information that could be used for decision making.

5.2 Inliers vs. Outliers LOS Analysis

Inliers and Outliers are one way to cluster patients. Clinicians have the perception that outlier patients end up staying longer in the hospital compared to Inliers, as they might be treated in a ward that is not the speciality ward for the patient's condition. However, this remains a hypothesis. This analysis attempts to establish the correlation between LOS of patients and their Inlier/Outlier status. The first cluster of patients included in this analysis contains patients admitted to wards belonging to the division of General Medicine as these journeys portray the diversity of movement.

The challenge experienced so far has been to find the correct cluster of data or patient journey within the division of General Medicine that would appropriately represent the characteristics needed to give a holistic representation taking into consideration of the various factors that influence the status of a patient journey at a point in time.

The results of the very first analysis from the cluster selected were contrary to the hypothesis, i.e., the LOS for Outliers is not necessarily longer than Inliers. Upon consultation with the Clinicians and looking at the data closely, it was established that LOS study from grouping the journeys using status field alone was not an accurate representation to derive LOS for each status group. Another filter was applied to ensure that all the journeys considered were exclusively within the General Medicine division. This cluster of data has been confirmed by the Clinicians to be the best representation of the data for the analysis.

The next stage in this analysis is to separate the time spent at ward "FMC" or ED from the overall LOS. It has to be noted here, that these are inpatient records and in theory ward "FMC" or ED time should be zero however, this is not the case for most of the records, which indicates that many inpatients are spending time at ED when they should already be placed in a ward.

This brings another question to be clarified before LOS results could be finalised. *"The question is whether time spent in ward 'FMC' should be classified as Inliers time or Outliers time?"* Following on from this clarification, LOS might be able to be calculated. Before this decision could be made, the cluster or the pool criteria has to be checked with the Clinicians.

Inpatient LOS has become one of the many ways used to measure performance of a hospital. Thomas, Guire et al. (1997) states that patient mean LOS has been used to measure quality of care and hospital efficiency in terms of resource usage. Thomas, Guire et al. (1997) further

asserts that lower than normal LOS could indicate that hospitals are discharging patients early possibly sacrificing quality of care. Hospitals react differently to the continuous rising cost of health care. One way is to reduce the average inpatient LOS and unfortunately some hospitals reduce the number of beds in the hospital as a direct response to increasing cost of healthcare.

6 Discussion

The process undertaken has proven to be a viable approach in analysing the inpatient journey. The main challenge has been in defining the boundaries of the parameters needed and defining the parameters in accordance to the actual practice rather than analysing or undertaking process mining based on the field value only.

According to the Clinician's view, the possibility of an Inlier staying longer is very viable as often when there are lack of beds the patients who are less sick are transferred to Outlier wards and patients who need more specialised care wait in ED before ending in an Inlier ward and consequently contributing to a longer Inlier LOS. The other scenario could be that a sicker patient staying longer at the hospital might eventually end up in an Inlier ward after being at various Outlier wards.

The collaboration with Clinicians has been an invaluable experience in this process. Hospitals are already undertaking various statistical data analysis for various reporting purposes to conform to KPIs and the approach taken in this study is to further break down the information and data to discover hidden knowledge.

7 Conclusion & Future Work

A smooth patient journey is an important aspect of a patient's experience in the hospital. It could be as important as the actual treatment given to the patient, as a smooth journey will aid the healing process of a patient with physiological tranquillity.

After the finalisation of the pool or cluster of patient journeys as an output of the data analysis, process mining techniques will be used for knowledge discovery of the inpatient journey. LOS of Inliers versus Outliers will be analysed using the process-oriented approach outlined here to investigate if Outliers have or have not followed and optimum flow. Patient journey control flow will be analysed using various process mining algorithms that are available within ProM, which at the same time will lead to the identification of the more effective or appropriate algorithm to use for this kind of analysis.

Process mining results in conjunction with the usage of the proposed methods are perceived to offer added benefit to the already successful implementation of "lean thinking" and possibly enhance the improvement in areas where "lean thinking" approach alone is inadequate to reveal insight to access block.

8 References

- Ben-Tovim, D. I., J. E. Bassham, et al. (2008). "Redesigning care at the Flinders Medical Centre: clinical process redesign using "lean thinking". *Medical Journal of Australia* **188**(6): S27-S27-31.

- Ben-Tovim, D. I., M. L. Dougherty, et al. (2008). "Patient journeys: the process of clinical redesign." Medical Journal of Australia **188**(6): S14-S14-17.
- Braitberg, G. (2007). "Emergency department overcrowding: dying to get in?" Medical Journal of Australia **187**(11/12): 624-624-625.
- Ceglowski, R., L. Churilov, et al. (2007). "Combining Data Mining and Discrete Event Simulation for a value-added view of a hospital emergency department." The Journal of the Operational Research Society **58**(2): 246-246-254.
- Fatovich, D. M. (2002). "Emergency medicine." British Medical Journal **324**(7343): 958-958-962.
- Lobach, M. D. P. M. S. D. F. and P. W. E. Hammond (1997). "Computerized Decision Support Based on a Clinical Practice Guideline Improves Compliance with Care Standards." The American Journal of Medicine **102**(1): 89-98.
- Mans, R., H. Schonenberg, et al. (2008). Process mining techniques: An application to stroke care.
- O'Connell, T. J., J. E. Bassham, et al. (2008). "Clinical process redesign for unplanned arrivals in hospitals." Medical Journal of Australia **188**(6): S18-S18-22.
- Rebuge, Á. and D. R. Ferreira (2011). "Business process analysis in healthcare environments: A methodology based on process mining." Information Systems In Press, Corrected Proof.
- Richard C, D. (2011). "Can Lean Thinking Transform American Health Care?" Annals of Emergency Medicine **57**(3): 279-281.
- Rozinat, A., M. T. Wynn, et al. (2009). "Workflow simulation for operational decision support." Data & Knowledge Engineering **68**(9): 834-850.
- Thomas, J. W., K. E. Guire, et al. (1997). "Is patient length of stay related to quality of care?" Journal of Healthcare Management **42**(4): 489-489-507.
- van der Aalst, W. M. P., H. A. Reijers, et al. (2007). "Business process mining: An industrial application." Information Systems **32**(5): 713-732.
- van der Aalst, W. M. P. and A. J. M. M. Weijters (2004). "Process mining: a research agenda." Computers in Industry **53**(3): 231-244.
- Vezyridis, P., S. Timmons, et al. (2011). "Going paperless at the emergency department: A socio-technical study of an information system for patient tracking." International Journal of Medical Informatics **80**(7): 455-465.

Effects of Physician Collaboration Network on Hospital Outcomes

Shahadat Uddin and Liaquat Hossain

Project Management Graduate Program, The University of Sydney, Australia

shahadat.uddin@sydney.edu.au; liaquat.hossain@sydney.edu.au

Abstract

Previous studies have documented the effect of collaboration among physicians on the effectiveness in delivering health services and in producing better patient outcomes. However, there is no systematic empirical study suggesting the underlying relationship between the collaboration network of physicians and its effect on hospital outcomes (i.e., *hospitalization cost* and *readmission rate*). In this study, we first propose a way to capture collaboration network among physicians from their visiting information to patients. Then we explore the effect of different attributes (i.e., *degree centrality*, *betweenness centrality*, and *network density*) of physician collaboration network (PCN) on hospital outcomes. Our results show that *degree centrality* (i.e., level of involvement) and *network density* (i.e., level of connectedness) of PCN are negatively correlated with *hospitalization cost* and *readmission rate*. In contrast, *betweenness centrality* (i.e., capacity to control the flow of information) is found positively correlated with *hospitalization cost* and *readmission rate*. In their respective hospitals, healthcare managers or administrators may follow our research findings to reduce cost and improve quality (i.e., lower readmission rate).

Keywords: Physician collaboration network; readmission rate; hospitalization cost; network attribute.

1 Introduction

Collaboration, the most important aspect of team care (Cunningham and Dillon, 1997), is an intricate concept with multiple attributes including sharing of planning, making decisions, solving problems, setting goals, assuming responsibilities, working together cooperatively, and communicating and coordinating openly (Luukkonen et al., 1993, Beaver and Rosen, 1978). It generally refers to individual and organizational approach of '*working together*' to address problems and deliver outcomes that are not easily or effectively achieved by working alone (Hoffman, 1998). According to De Vreede and Briggs (1995), collaboration is a recurring process where two or more people or organizations work together towards common goals.

Collaboration enables individuals and organizations to work together more effectively and efficiently. Collaborative practice is now central to the way we work, deliver services, and produce innovations. Collaborative relationships among individuals are highly celebrated in organizations because the synergies realized by combining multi-dimensional efforts and diverse expertise produce benefits greater than those achieved through individual effort (Knoben and Oerlemans, 2006, Uddin and Hossain, 2011b).

Collaboration is characterized by highly interdependent relationships among participating individuals or organizations (Erdoes et al., 1973). The behavior and role of participants in a collaborative environment are greatly influenced by the current level of fulfilment of the desired outcomes (Huxham and Vangen, 2000). Participants may have to radically alter the way they think, behave, and operate to cope up with the demands and challenges of a highly dynamic collaborative environment. For example, physicians may set up an initial treatment plan (e.g. medical check-up routine, drug intervention plan, and any possible surgery plan) for a hospitalized chronic patient suffering from asthma and diabetes. If that patient has a sudden '*massive heart attack*' then physicians have to change the initial treatment plan for effective and faster disease recovery.

As a facilitator to accelerated group performance, the importance of collaboration has been identified by many scholars in diverse research areas such as virtual research and development (R&D) organizations (Ahuja et al., 2003), scientific network among authors (Huang et al., 2008), evaluation of creative performance (Liu et al., 2005), and performance analysis of physical task and foreign markets (Newman, 2001, Newman, 2004). In the context of healthcare service providers or hospitals, collaboration among different healthcare professionals is recognized as a catalyst to improved patient outcomes such as less hospital length of stay and hospitalization cost (Cowan et al., 2006, Tschannen and Kalisch, 2009, Uddin and Hossain, 2011a), less death rate (Knaus et al., 1986), and higher satisfaction (Baggs et al., 1999, Lindeke and Sieckert, 2005). In healthcare settings, collaboration allows input from multiple professions (e.g. nurse, physicians) which could produce decisions leading to better patient outcomes because those decisions are based on more complete information.

The context of this study is the healthcare service providers or hospitals. It can be conceptualized that physicians collaborate with each other and with other hospital staff (e.g. nurses) in order to provide effective services to hospitalized patients. Based on the patient

condition and unavailability of their colleagues, physicians might seek advices or suggestions from other physicians working in different workplaces. Eventually, this type of medical practice culture inside healthcare service providers or hospitals develops a professional collaboration network among physicians, which we termed as '*Physician Collaboration Network, PCN*'.

In recent years, there has been an increased trend on clinical measures of quality, such as mortality and morbidity, to study collaboration and coordination in healthcare organizations (Kalkanis et al., 2003, Sylvia et al., 2008). However, to quantify the patient perception of quality is not an easy job, and could result different responses from different patients for similar services. Moreover, not all hospital admissions are life-threatening. Some of them have very low or zero chance of death, such as a hospital admission for a broken hand. In this study, we analyze how *hospitalization cost* and *readmission rate* are affected by different network attributes of PCN. Bavelas (1950) first, subsequently by many researchers (Guetzkow and Simon, 1955, Shaw, 1956), show that network attributes of any collaboration network (e.g., PCN) have impacts on its outcomes (e.g., readmission rate). The following two questions motivate this research:

- (i) How does the different network structure of PCN affect *hospitalization cost* and *readmission rate*?
- (ii) What structural properties of PCN are related to *hospitalization cost* and *readmission rate*?

2 Review of Collaboration Research within Healthcare Context

There are numerous studies in current literature exploring the effect of collaboration among healthcare professionals on patient outcomes and hospital performance. Most of these studies examine the impact of nurse-physician collaboration on patient outcomes. One classic study, led by Knaus and his team, identifies a significant relationship between the degree of nurse-physician collaboration and patient mortality in intensive care units (Knaus et al., 1986). They study treatment and outcome in 5030 intensive care unit patients, and find that hospitals where nurse-physician collaboration is present report mortality rate of 41%, which is lower than the predicted number of patient deaths. Conversely, hospitals that are noted for poor communication among healthcare professionals exceed their predicted number of patient deaths by 58%. In a two group quasi-experiment on 1207 general medicine patients ($n = 581$ in the experimental group and $n = 626$ in the control group), Cowan, Shapiro et al. (2006) notice average hospital length of stay is significantly lower for patients in the experimental group than the control group (5 vs. 6 days, $P < .0001$) which contributes a '*backfill profit*' of US\$1591 per patient to hospital. There are other studies that also highlight the importance of collaboration among healthcare professionals for better patient outcomes. In particular, greater collaboration between nurse and physician has been found to expedite hospital performance which is measured in terms of increased patient and professional satisfaction (Melin and Persson, 1996, Luukkonen et al., 1992), improved quality of care (Kramer and

Schmalenberg, 2003, Luukkonen et al., 1992), lower nursing turnover (Melin and Persson, 1996), and lower job stress (Melin and Persson, 1996, Luukkonen et al., 1992).

Sommers, Marton et al. (2007) examine the impact of an interdisciplinary and collaborative practice intervention involving a primary care physician, a nurse, and a social worker for community-dwelling seniors with chronic illnesses. They conduct a controlled cohort study of 543 patients in 18 private office practices of primary care physicians. The intervention group receives care from their primary care physician working with a registered nurse and a social worker, while the control group receives care as usual from primary care physician. They notice that the intervention group produced better result to readmission rate and average office visits to all physicians. Moreover, the patients in the intervention group report an increase in social activities compared with control group's decrease. There are other studies emphasizing collaboration for effective patient outcome across professional boundaries within hospitals. By analyzing data collected from 105 interviews (with 40 physician, 32 case managers, 23 physician office staff, 8 administrators, and 2 case assistants), Netting and Williams (2005) argue that there is a growing need to collaborate and communicate across professional lines rather than make assumptions about who can do what for better patient outcomes, professional satisfaction, and hospital performance.

From the review of collaboration literature in healthcare context, it is evident that some studies have given emphasis on nurse-physician collaboration, while some others have given importance to physician collaboration with all healthcare professionals (e.g. social worker, hospital administrator, and case managers) for better patient outcomes and hospital performance. However, studies to date have not considered a network-level analysis of physician-physician collaboration to understand its effect on hospital performance. In this study, we analyze *physician collaborative network (PCN)* to understand hospital performance.

3 Physician Collaboration Network and Research Model Development

PCNs evolve over time at each hospital. In this study, we assume that collaborations among physicians emerge when they visit common hospitalized patients. Therefore, when physicians visit common patients within the same hospital or healthcare organization PCN emerges among them. Figure 1 illustrates an example of such a PCN construction. In hospital H1, patient Pa1 is seen by Ph1, Ph2 and Ph4 physicians, and physician Ph2, Ph3 and Ph4 visit patient Pa2. This is depicted in the *patient-physician* network as in Figure 1(a). The corresponding PCN for this *patient-physician* network is demonstrated in Figure 1(b). In this PCN, there are network connections with weight 1 between Ph1 and Ph4, between Ph1 and Ph2, and between Ph2 and Ph3 because they visit only one common patient. The weight of the link between Ph2 and Ph4 is 2 as they have two common patients.

In constructing PCN from *patient-physician* network, for each patient actor we first find out all physician actors

who are connected with that patient actor. Then, we add 1 in the corresponding *physician-physician* adjacency matrix X . We follow this process for all patient actors of a *patient-physician* network.

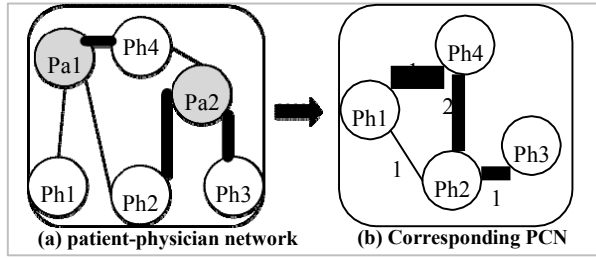


Figure 1: (a) patient-physician network; and (b) corresponding PCN

The principal goal of this study is to investigate the impact of the attributes of PCN on *hospitalization cost* and *hospital readmission rate*. With that purpose, we propose a research model (see Figure 2) where network measures of PCN are considered as independent variables, and *hospitalization cost* and *hospital readmission rate* are taken as dependent variables. The selection of network measures of PCN is guided by two network theories: (i) Bavelas' Centralization Theory (Bavelas, 1950); and (ii) Freeman's Centrality Theory (Freeman, 1978). These two theories can explain structural influences of collaboration and communication network on group performance.

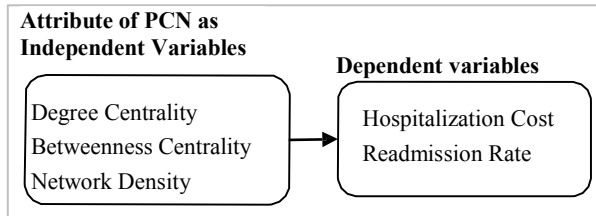


Figure 2: Proposed research model of this study.

3.1 Variables of Proposed Research Model

The approach for constructing PCN as illustrated in Figure 1 defines a link (with weight 1) between two physicians when they visit one common patient. If two physicians have two common patients then the weight of the link between them will be 2 and so on. This means the PCNs evolved over time within hospitals are weighted networks. Thus, we have to consider approaches that are suitable for weighted networks to quantify network variables of our proposed research model.

Degree Centrality

Degree centrality is defined by the number of direct links that a particular node has in the network (Newman et al., 2002). It can be defined both at actor-level and at network-level. As we analyze PCNs having different sizes in this research, the network-level quantification for degree centrality is considered. The equation for network-level *degree centrality* of a weighted graph with n actors:

$$\text{Degree_Centrality} = \frac{\sum_{i=1}^n \sum_{j=i+1}^n w_{ij}}{n*(n-1)} \dots\dots\dots (1)$$

Where, w_{ij} is the weight of the link between actor i and j

Betweenness Centrality

Like degree centrality, *betweenness centrality* can be measured from the perspective of both node-level and network-level perspective where the first one is used to measure the later one. In node-level, *betweenness centrality* views an actor as being in a favoured position to the extent that the actor falls on the shortest paths between other pairs of actors in the network. That is, nodes that occur on many shortest paths between other pair of nodes have higher *betweenness centrality* than those they do not (Freeman, 1978). The node-level betweenness for a node n_{ib} (i.e. $C_B(i_n)$) can be represented by the following equation (Wasserman and Faust, 2003):

$$C_B(n_i) = \sum_{j < k} \frac{g_{jk}(n_i)}{g_{jk}} \dots\dots\dots (2)$$

Where, $i \neq j \neq k$; $g_{jk}(n_i)$ represents the number of shortest paths linking the two actors that contain actor i ; and g_{jk} is the number of shortest paths linking actor j and k .

To compare betweenness across different networks, network-level betweenness is used. It represents the average difference between the relative centrality of the most central node ($C_B(p^*)$), and that of all other nodes (Freeman, 1977):

$$C_B = \frac{\sum_{i=1}^n [C_B(p^*) - C_B(p_i)]}{n^3 - 4n^2 + 5n - 2} \dots\dots\dots (3)$$

Where, $C_B(p^*)$ is the largest realized actor betweenness index for the set of n actors, and $C_B(p_i)$ is the *node-level betweenness* for actor i .

Network Density

The *density* measure represents the proportion of existing links in a network relative to the total number of possible ties among all the network actors (Wasserman and Faust, 2003). The density value for a network is 1 only when all the nodes of that network are connected with each other. Conversely, for a completely sparse network, the density value is 0. In an undirected network of size n , theoretically there are $n*(n-1)/2$ possible links among its n nodes. Thus, mathematically, density can be defined as:

$$\text{Density} = \frac{2 * \sum_{i=1}^n \sum_{j=i+1}^n w_{ij}}{n*(n-1)} \dots\dots\dots (4)$$

Where, w_{ij} is the weight of the link between actor i and j

Dependent Variables

In this research, we consider *hospitalization cost* and *readmission rate* as dependent variables. PCNs evolve for each particular type of disease during the course of providing healthcare services to patients within each hospital. Some of these patients might need readmission for the same disease. We use this information as one of the independent variables. To measure *hospitalization cost*, which is utilized as another dependent variable in our proposed model, of a particular PCN we consider the hospitalization cost of only those patients who are belong to that PCN, and then calculate the average of hospitalization cost of all those patients.

4 Methodology

4.1 Research Dataset

In this research, to test our proposed hypotheses we consider health insurance claim data from a non-profit Australian health insurance organization, hospital contribution fund (HCF). It includes members claim data from January 2005 to February 2009. This dataset contains information about three different categories of claims: ancillary claim, medical claim, and hospital claim. *Ancillary claims* are auxiliary claims for medical services, e.g., dental, optical, physiotherapy, dietician, and pharmaceutical. All claims coming from specialist physicians except of the ancillary type are medical claims. The claims for the services provided to hospitalized patients are considered as hospital claims. In general, patients have medical claims, hospital claims, and very few ancillary claims for their admissions to hospitals.

In our dataset, there are about 14.87 million ancillary, 8.98 millions medical, and 3.1 millions hospital claims that HCF received from 2507 hospitals for the health services provided to its 0.44 million members. As people have hospital admissions for a wide range of illness and patients with a particular disease need to be seen by particular specialist physicians, different types of PCNs (e.g., a PCN for knee-surgery patient, and a PCN for heart-attack patient) evolve inside a hospital for hospitalized patients. In this study, we consider PCNs only for THR patients from 85 different hospitals where at least 5 THR patients get admitted during the data collection period. In these 85 hospitals, there are 2229 patients get admitted during our data collection period. These patients lodge in total 1383 ancillary, 65871 medical, and 23369 hospital claims. None of the patients of our research dataset died during the hospitalized period.

4.2 Construction of PCN

From the 'medical claims' details of HCF dataset, we can trace how many physicians visit a particular hospitalized patient because physicians make a 'medical claim' to HCF for every single visit to hospitalized patients. Based on this information and by applying the PCN development concept of Figure 1, we figure out the structure of PCN of each hospital for THR patients.

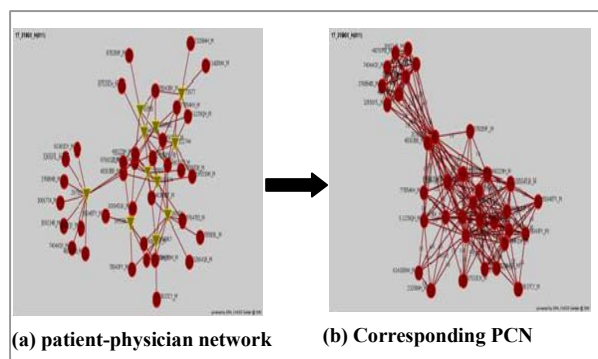


Figure 3: Construction of PCN from research dataset. The red circle represents physician and the yellow triangle represents patient.

4.3 Method of Analysis

Using UCINET-6 (Borgatti et al., 2002), we first calculate the four network measures of our proposed model from the PCNs that are constructed from our research dataset. Second, we apply correlation test to assess relations of four network measures with *hospitalization cost* and *readmission rate*. As it is revealed that histograms for these four network measures do not follow a complete normal distribution curve, we choose a non-parametric (i.e. Spearman test) correlation test.

5 Results

By inferring the Spearman coefficient values of Table 1 and *t*-test results from Table 2, we reach to the following statements:

(i) Degree centrality of PCN has negative correlation with both Hospitalization Cost and Readmission Rate.

We find that *degree centrality* of PCN is negatively correlated with *hospitalization cost* ($\rho = -0.212$, $p < 0.05$ at 2-tailed) and *readmission rate* ($\rho = -0.366$, $p < 0.01$ at 2-tailed). A decrease in *degree centrality* produces a downturn for both *hospitalization cost* and *readmission rate*. That means high *degree centrality* among physicians makes healthcare organizations possible to provide healthcare services to hospitalized patient with efficient cost structure (i.e., low cost) and better quality (i.e., low readmission rate). As *degree centrality* represents actors' involvement in a given network higher level of physicians' participation should be encouraged in hospital settings.

(ii) Betweenness centrality of PCN has positive correlation with both Hospitalization Cost and Readmission Rate.

The Spearman correlation test shows that *betweenness centrality* of PCN is positively correlated with *hospitalization cost* ($\rho = 0.264$, $p < 0.01$ at 2-tailed) and *readmission rate* ($\rho = 0.460$, $p < 0.01$ at 2-tailed). As it is always expected to have low *hospitalization cost* and *readmission rate*, this result indicates that high *betweenness centrality* is not conducive for healthcare service providers or hospitals. PCNs that do not have stronger level of capacity to control the flow of information should be highly encouraged in any hospital

setting as *betweenness centrality* indicates the capacity to control flow of information.

(iii) Density of PCN has negative correlation with both Hospitalization Cost and Readmission Rate.

Results indicate that *density* of PCN is negatively correlated with *hospitalization cost* ($\rho = -0.228$, $p < 0.05$ at 2-tailed) and *readmission rate* ($\rho = -0.282$, $p < 0.05$ at 2-tailed). Both *hospitalization cost* and *readmission rate* of a hospital change inversely with the change in the *density* of the PCN of that hospital. Because *density* indicates the level of connectedness among actors in a given network PCNs having higher level connectedness among physicians should be celebrated in a hospital setting.

Table 1: Spearman correlation coefficient among our research variables

	Hospitalization Cost	Readmission Rate
Degree Centrality	-0.212*	-0.366**
Betweenness Centrality	0.264**	0.460**
Network Density	-0.228*	-0.282*
**. Correlation is significant at the 0.01 level (2-tailed)		
*. Correlation is significant at the 0.05 level (2-tailed)		

6 Discussion

For any PCN, it is desirable to have low *hospitalization cost* and *readmission rate*. Therefore, negative correlations between any pair of independent and dependent variables of our research model represent a way to control these two dependent measures in positive direction, and vice versa.

We discover that *degree* and *density* (first and third findings) of PCN have negative correlation with *hospitalization cost* and *readmission rate*. Having professional relation with more colleagues (i.e., higher *degree*) makes it easier for an individual physician to share medical knowledge effectively. The importance of sharing of physician knowledge is also identified by present healthcare literature. Ryu et al. (2003), for example, find that sharing knowledge of physicians within hospitals is very critical to the success and survival in competitive environments for hospital organizations. According to our first finding, healthcare managers or administrators should encourage the involvement of more physicians for the treatment of hospitalized patients in order to reduce *hospitalization cost* and *readmission rate*. The measure *density* is the proportion of possible links that are actually present in the network. In a dense PCN, increased number of links exists among physicians. This means higher *degree* also ensures increased *density*. A PCN could have high *degree centrality* when its member physicians work in small well connected groups though this type of PCN does not confirm a higher *density*. Thus, our third finding suggests that physician should not work in small groups; instead they need to work with more of their colleagues. This will increase both *degree* and *density*, which eventually enable faster sharing of known

knowledge (Ryu et al., 2003), and reduce *hospitalization cost* and *readmission rate*.

On the other hand, we notice that *betweenness* (from second finding) of PCN has positive correlation with both *hospitalization cost* and *readmission rate*. From the perspective of PCN structure, at network-level, high *betweenness centrality* indicates big differences in the node-level *betweenness centralities* between the most central node and that of all other nodes. That means denominator of equation 3 will be high if the network-level *betweenness* is high. Thus, our result regarding *betweenness centrality* indicates that big differences among the node-level *betweenness centrality* scores in PCN are not favourable to the dependent variables (i.e., *hospitalization cost* and *readmission rate*) of our proposed model. A big fluctuation among *betweenness centrality* scores in any network reflects an unequal participation, and uneven collaboration and communication control of its member nodes (Freeman, 1978). In this type of network, only a small number of actors play major collaboration and communication role. Therefore, in their corresponding hospitals, healthcare managers or administrators have to make sure the equal participation of physicians in PCNs.

PCN, *hospitalization cost* and *readmission rate* are the three key measures of this study. First, we construct PCN from the information of physicians' visits to patients during their hospitalization period. We assume that collaboration emerges between two physicians when they visit a common patient. It is standard professional practice around the world that when physicians visit patients they give advice or suggestions to patients based on patient health condition and previous medication history deposited in the patient *log book*. All previous advice or suggestions by any physician to a patient have been taken into consideration during any subsequent physician visit to that patient. This kind of practice culture in healthcare organizations or hospitals establishes the validity and reliability of the construction process of PCN, and the generalizability of our research findings. Second, we consider *hospitalization cost* and *readmission rate* as dependent variables in our proposed model. The use of *readmission rate* as an outcome measure has well acceptance in healthcare literature (Chen et al., 2010, Ross et al., 2010). Though *hospitalization cost* is not an well accepted outcome measure in healthcare literature (Chen et al., 2010), coupled with *readmission rate*, it could be a potential outcome measure.

7 Conclusion

This study is motivated by two research questions: "How does the different network structure of PCN affect *hospitalization cost* and *readmission rate*?" and "What structural properties of PCN are related to *hospitalization cost* and *readmission rate*?" In line with these research questions, we find that the PCNs which are characterized by higher *degree centrality*, lower *betweenness centrality*, and higher *network density* are conducive to *hospitalization cost* and *readmission rate*.

In this research, we propose a new way to capture the collaboration network among physicians within healthcare providers or hospitals. Further, we investigate

PCN using social network approach of network analysis to explore the impact of the attribute of PCN on *hospitalization cost* and *readmission rate*. To our knowledge this is the first study which takes the initiative to understand the impact of physicians' collaboration within hospitals by applying structural network methods and measures.

Reference

- AHUJA, M., GALLETTA, D. & CARLEY, K. (2003) Individual centrality and performance in virtual R&D groups: An empirical study. *Management Science*, 49, 21-38.
- BAGGS, J., SCHMITT, M., MUSHLIN, A., MITCHELL, P., ELDREDGE, D., OAKES, D. & HUTSON, A. (1999) Association between nurse-physician collaboration and patient outcomes in three intensive care units. *Critical Care Medicine*, 27, 1991.
- BAVELAS, A. (1950) Communication patterns in task-oriented groups. *Journal of the acoustical society of America*, 22, 725-730.
- BEAVER, D. B. & ROSEN, R. (1978) Studies in scientific collaboration. *Scientometrics*, 1, 65-84.
- BORGATTI, S., EVERETT, M. & FREEMAN, L. (2002) UCINET for Windows: Software for social network analysis. *Harvard Analytic Technologies*, 6.
- CHEN, L. M., JHA, A. K., GUTERMAN, S., RIDGWAY, A. B., ORAV, E. J. & EPSTEIN, A. M. (2010) Hospital cost of care, quality of care, and readmission rates: penny wise and pound foolish? *Archives of Internal Medicine*, 170, 340.
- COWAN, M., SHAPIRO, M., HAYS, R., AFIFI, A., VAZIRANI, S., WARD, C. & ETTNER, S. (2006) The effect of a multidisciplinary hospitalist/physician and advanced practice nurse collaboration on hospital costs. *Journal of Nursing Administration*, 36, 79.
- CUNNINGHAM, S. J. & DILLON, S. M. (1997) Authorship patterns in information systems. *Scientometrics*, 39, 19-27.
- ERDOES, P., GRAHAM, R. L., MONTGOMERY, P., ROTHCHILD, B., SPENCER, J. & STRAUS, E. G. (1973) Euclidean ramsey theorems. I* 1. *Journal of Combinatorial Theory, Series A*, 14, 341-363.
- FREEMAN, L. (1978) Centrality in social networks: Conceptual clarification. *Social Networks*, 1, 215-239.
- FREEMAN, L. C. (1977) A set of measures of centrality based on betweenness. *Sociometry*, 40, 35-41.
- GROSSMAN, J. W. & ION, P. D. F. (1995) On a portion of the well-known collaboration graph. *Congressus Numerantium*, 129-132.
- GUETZKOW, H. & SIMON, H. A. (1955) The impact of certain communication nets upon organization and performance in task-oriented groups. *Management Science*, 233-250.
- HOFFMAN, P. (1998) *The man who loved only numbers: The story of Paul Erdos and the search for mathematical truth*, Hyperion.
- HUANG, J., ZHUANG, Z., LI, J. & GILES, C. L. (2008) Collaboration over time: characterizing and modeling network evolution. *ACM*.
- HUXHAM, C. & VANGEN, S. (2000) Leadership in the shaping and implementation of collaboration agendas: How things happen in a (not quite) joined-up world. *The Academy of Management Journal*, 43, 1159-1175.
- KALKANIS, S., ESKANDAR, E., CARTER, B. & BARKER, F. (2003) Microvascular Decompression Surgery in the United States, 1996 to 2000: Mortality Rates, Morbidity Rates, and the Effects of Hospital and Surgeon Volumes. *Neurosurgery*, 52, 1251-1262.
- KNAUS, W. E., DRAPER, WAGNER, D. P. & ZIMMERMAN, J. E. (1986) An evaluation of outcome from intensive care in major medical centers. *Annals of Internal Medicine*, 104, 410.
- KNOBEN, J. & OERLEMANS, L. (2006) Proximity and inter-organizational collaboration: A literature review. *International Journal of Management Reviews*, 8, 71-89.
- KRAMER, M. & SCHMALENBERG, C. (2003) Magnet hospital nurses describe control over nursing practice. *Western Journal of Nursing Research*, 25, 434.
- LINDEKE, L. & SIECKERT, A. (2005) Nurse-physician workplace collaboration. *Online Journal of Issues in Nursing*, 10, 10.
- LIU, X., BOLLEN, J., NELSON, M. L. & VAN DE SOMPEL, H. (2005) Co-authorship networks in the digital library research community. *Information Processing & Management*, 41, 1462-1480.
- LUUKKONEN, T., PERSSON, O. & SIVERTSEN, G. (1992) Understanding patterns of international scientific collaboration. *Science, Technology & Human Values*, 17, 101.
- LUUKKONEN, T., TIJSSSEN, R. J. W., PERSSON, O. & SIVERTSEN, G. (1993) The measurement of international scientific collaboration. *Scientometrics*, 28, 15-36.
- MEHO, L. I. & YANG, K. (2007) Impact of data sources on citation counts and rankings of LIS faculty: Web of Science versus Scopus and Google Scholar. *Journal of the American Society for Information Science and Technology*, 58, 2105-2125.
- MELIN, G. & PERSSON, O. (1996) Studying research collaboration using co-authorships. *Scientometrics*, 36, 363-377.
- NEWMAN, M., FORREST, S. & BALTHROP, J. (2002) Email networks and the spread of computer viruses. *Physical Review E*, 66, 35101.
- NEWMAN, M. E. J. (2001) The structure of scientific collaboration networks. *Proceedings of the National Academy of Sciences of the United States of America*, 98, 404.
- NEWMAN, M. E. J. (2004) Coauthorship networks and patterns of scientific collaboration. *Proceedings of the National Academy of Sciences of the United States of America*, 101, 5200.
- ROSS, J. S., CHEN, J., LIN, Z., BUENO, H., CURTIS, J. P., KEENAN, P. S., NORMAND, S. L. T., SCHREINER, G., SPERTUS, J. A. & VIDÁN, M. T. (2010) Recent

- national trends in readmission rates after heart failure hospitalization. *Circulation: Heart Failure*, 3, 97.
- RYU, S., HO, S. & HAN, I. (2003) Knowledge sharing behavior of physicians in hospitals. *Expert Systems with Applications*, 25, 113-122.
- SHAW, M. E. (1956) Random versus Systematic Distribution of Information in Communication Nets1. *Journal of Personality*, 25, 59-69.
- SYLVIA, M., GRISWOLD, M., DUNBAR, L., BOYD, C., PARK, M. & BOULT, C. (2008) Guided care: cost and utilization outcomes in a pilot study. *Disease Management*, 11, 29-36.
- TSCHANNEN, D. & KALISCH, B. (2009) The effect of variations in nurse staffing on patient length of stay in the acute care setting. *Western Journal of Nursing Research*, 31, 153.
- UDDIN, M. S. & HOSSAIN, L. (2011a) Social Networks Enabled Coordination Model for Cost Management of Patient Hospital Admissions. *Journal for Healthcare Quality*, 33, 37-48.
- UDDIN, S. & HOSSAIN, L. (2011b) Disaster coordination preparedness of soft target organisations. *Disasters*, 35, 623-638.
- WASSERMAN, S. & FAUST, K. (2003) *Social network analysis: Methods and applications*, Cambridge Univ Pr.

Author Index

- Abawajy, Jemal, 3
Al Mazari, Ali, 41
- Bellucci, Emilia, 3
Butler-Henderson, Kerry, iii
- Davis-Desmond, Patrick, 13
Day, Karen, 31
Dhillon, Jaspaljeet Singh, 21
- Garsia, Roger J., 41
Gray, Kathleen, iii
Groza, Tudor, 49
Gu, Yulong, 31
- Hakendorf, Paul, 59
Hossain, Liaquat, 67
Hunter, Jane, 49
- Kasmiran, Khairul A., 41
- Lutteroth, Christof, 21
- Mollá, Diego, 13
- Muecke, Nial, 3
- Paul, Razan, 49
Perimal-Lewis, Lua, 59
Pollock, Malcolm, 31
- Qin, Shaowen, 59
- Ramos, Czarina, 21
- Stranieri, Andrew, 3
- Thompson, Campbell, 59
- Uddin, Shahadat, 67
- Venkatraman, Sitalakshmi, 3
- Wünsche, Burkhard C., 21
Warren, Jim, 31
White, Sue, 31
- Zomaya, Albert Y., 41

Recent Volumes in the CRPIT Series

ISSN 1445-1336

Listed below are some of the latest volumes published in the ACS Series *Conferences in Research and Practice in Information Technology*. The full text of most papers (in either PDF or Postscript format) is available at the series website <http://crpit.com>.

- Volume 113 - Computer Science 2011**
Edited by Mark Reynolds, The University of Western Australia, Australia. January 2011. 978-1-920682-93-4.
Contains the proceedings of the Thirty-Fourth Australasian Computer Science Conference (ACSC 2011), Perth, Australia, 17-20 January 2011.
- Volume 114 - Computing Education 2011**
Edited by John Hamer, University of Auckland, New Zealand and Michael de Raadt, University of Southern Queensland, Australia. January 2011. 978-1-920682-94-1.
Contains the proceedings of the Thirteenth Australasian Computing Education Conference (ACE 2011), Perth, Australia, 17-20 January 2011.
- Volume 115 - Database Technologies 2011**
Edited by Heng Tao Shen, The University of Queensland, Australia and Yanchun Zhang, Victoria University, Australia. January 2011. 978-1-920682-95-8.
Contains the proceedings of the Twenty-Second Australasian Database Conference (ADC 2011), Perth, Australia, 17-20 January 2011.
- Volume 116 - Information Security 2011**
Edited by Colin Boyd, Queensland University of Technology, Australia and Josef Pieprzyk, Macquarie University, Australia. January 2011. 978-1-920682-96-5.
Contains the proceedings of the Ninth Australasian Information Security Conference (AISC 2011), Perth, Australia, 17-20 January 2011.
- Volume 117 - User Interfaces 2011**
Edited by Christof Lutteroth, University of Auckland, New Zealand and Haifeng Shen, Flinders University, Australia. January 2011. 978-1-920682-97-2.
Contains the proceedings of the Twelfth Australasian User Interface Conference (AUIC2011), Perth, Australia, 17-20 January 2011.
- Volume 118 - Parallel and Distributed Computing 2011**
Edited by Jinjun Chen, Swinburne University of Technology, Australia and Rajiv Ranjan, University of New South Wales, Australia. January 2011. 978-1-920682-98-9.
Contains the proceedings of the Ninth Australasian Symposium on Parallel and Distributed Computing (AusPDC 2011), Perth, Australia, 17-20 January 2011.
- Volume 119 - Theory of Computing 2011**
Edited by Alex Potanin, Victoria University of Wellington, New Zealand and Taso Viglas, University of Sydney, Australia. January 2011. 978-1-920682-99-6.
Contains the proceedings of the Seventeenth Computing: The Australasian Theory Symposium (CATS 2011), Perth, Australia, 17-20 January 2011.
- Volume 120 - Health Informatics and Knowledge Management 2011**
Edited by Kerry Butler-Henderson, Curtin University, Australia and Tony Sahama, Queensland University of Technology, Australia. January 2011. 978-1-921770-00-5.
Contains the proceedings of the Fifth Australasian Workshop on Health Informatics and Knowledge Management (HIKM 2011), Perth, Australia, 17-20 January 2011.
- Volume 121 - Data Mining and Analytics 2011**
Edited by Peter Vamplew, University of Ballarat, Australia, Andrew Stranieri, University of Ballarat, Australia, Kok-Leong Ong, Deakin University, Australia, Peter Christen, Australian National University, Australia and Paul J. Kennedy, University of Technology, Sydney, Australia. December 2011. 978-1-921770-02-9.
Contains the proceedings of the Ninth Australasian Data Mining Conference (AusDM'11), Ballarat, Australia, 1-2 December 2011.
- Volume 122 - Computer Science 2012**
Edited by Mark Reynolds, The University of Western Australia, Australia and Bruce Thomas, University of South Australia. January 2012. 978-1-921770-03-6.
Contains the proceedings of the Thirty-Fifth Australasian Computer Science Conference (ACSC 2012), Melbourne, Australia, 30 January – 3 February 2012.
- Volume 123 - Computing Education 2012**
Edited by Michael de Raadt, Moodle Pty Ltd and Angela Carbone, Monash University, Australia. January 2012. 978-1-921770-04-3.
Contains the proceedings of the Fourteenth Australasian Computing Education Conference (ACE 2012), Melbourne, Australia, 30 January – 3 February 2012.
- Volume 124 - Database Technologies 2012**
Edited by Rui Zhang, The University of Melbourne, Australia and Yanchun Zhang, Victoria University, Australia. January 2012. 978-1-920682-95-8.
Contains the proceedings of the Twenty-Third Australasian Database Conference (ADC 2012), Melbourne, Australia, 30 January – 3 February 2012.
- Volume 125 - Information Security 2012**
Edited by Josef Pieprzyk, Macquarie University, Australia and Clark Thomborson, The University of Auckland, New Zealand. January 2012. 978-1-921770-06-7.
Contains the proceedings of the Tenth Australasian Information Security Conference (AISC 2012), Melbourne, Australia, 30 January – 3 February 2012.
- Volume 126 - User Interfaces 2012**
Edited by Haifeng Shen, Flinders University, Australia and Ross T. Smith, University of South Australia, Australia. January 2012. 978-1-921770-07-4.
Contains the proceedings of the Thirteenth Australasian User Interface Conference (AUIC2012), Melbourne, Australia, 30 January – 3 February 2012.
- Volume 127 - Parallel and Distributed Computing 2012**
Edited by Jinjun Chen, University of Technology, Sydney, Australia and Rajiv Ranjan, CSIRO ICT Centre, Australia. January 2012. 978-1-921770-08-1.
Contains the proceedings of the Tenth Australasian Symposium on Parallel and Distributed Computing (AusPDC 2012), Melbourne, Australia, 30 January – 3 February 2012.
- Volume 128 - Theory of Computing 2012**
Edited by Julián Mestre, University of Sydney, Australia. January 2012. 978-1-921770-09-8.
Contains the proceedings of the Eighteenth Computing: The Australasian Theory Symposium (CATS 2012), Melbourne, Australia, 30 January – 3 February 2012.
- Volume 129 - Health Informatics and Knowledge Management 2012**
Edited by Kerry Butler-Henderson, Curtin University, Australia and Kathleen Gray, University of Melbourne, Australia. January 2012. 978-1-921770-10-4.
Contains the proceedings of the Fifth Australasian Workshop on Health Informatics and Knowledge Management (HIKM 2012), Melbourne, Australia, 30 January – 3 February 2012.
- Volume 130 - Conceptual Modelling 2012**
Edited by Aditya Ghose, University of Wollongong, Australia and Flavio Ferrarotti, Victoria University of Wellington, New Zealand. January 2012. 978-1-921770-11-1.
Contains the proceedings of the Eighth Asia-Pacific Conference on Conceptual Modelling (APCCM 2012), Melbourne, Australia, 31 January – 3 February 2012.
- Volume 131 - Advances in Ontologies 2010**
Edited by Thomas Meyer, UKZN/CSIR Meraka Centre for Artificial Intelligence Research, South Africa, Mehmet Orgun, Macquarie University, Australia and Kerry Taylor, CSIRO ICT Centre, Australia. December 2010. 978-1-921770-00-5.
Contains the proceedings of the Sixth Australasian Ontology Workshop 2010 (AOW 2010), Adelaide, Australia, 7th December 2010.