

# Improving Data Quality in Primary Care: Modelling, Measurement, and the Design of Interventions

by

Justin St-Maurice

A thesis  
presented to the University of Waterloo  
in fulfillment of the  
thesis requirement for the degree of  
Doctor of Philosophy  
in  
Systems Design Engineering

Waterloo, Ontario, Canada, 2017

© Justin St-Maurice 2017



## Examining Committee Membership

The following served on the Examining Committee for this thesis. The decision of the Examining Committee is by majority vote.

External Examiner

DR. EMILY PATTERSON  
Associate Professor  
School of Health & Rehabilitation Sciences  
Ohio State University

Supervisor

DR. CATHERINE BURNS  
Professor  
Systems Design Engineering  
University of Waterloo

Internal Members

DR. CAROLYN MACGREGOR  
Associate Professor  
Systems Design Engineering  
University of Waterloo

DR. MOHAMED ALARAKHIA  
Adjunct Professor  
Systems Design Engineering  
University of Waterloo

Internal-external Member

DR. BRUCE BASKERVILLE  
Senior Scientist  
Applied Health Sciences  
University of Waterloo

## **Author's Declaration**

This thesis consists of material all of which I authored or co-authored: see Statement of Contributions included in the thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

## Statement of Contributions

### Papers included in this dissertation:

St-Maurice, J, & Burns, C. M. (in press). An Exploratory Case Study to Understand Primary Care Users and Their Data Quality Tradeoffs. *Journal of Data and Information Quality*.

This paper is incorporated in chapter 7 of this dissertation.

<i>Contributor</i>	<i>Statement of contribution</i>
<i>St-Maurice, J. (Candidate)</i>	<i>Conceptual design (95%) Data Collection &amp; Analysis (100%) Writing and editing (95%)</i>
<i>Burns, C.M.</i>	<i>Conceptual design (5%) Writing and editing (5%)</i>



## Abstract

In an era where governments around the world invest heavily in data collection and data management, poor-quality data is expensive and has many direct and indirect costs. While there are different types of data quality challenges, some of the more complex data quality problems depend on the design and production processes involved in generating data. Therefore, it is important to design systems that support better data quality. This involves understanding what quality means in a specific context, understanding how it can be measured, and identifying ways to encourage better data quality behaviours.

Healthcare is not immune to the challenges of data quality and can be classified as a complex socio-technical system by virtue of its characteristics. As such, the study of healthcare data quality and its improvement is well suited for the domain of systems design and human factors engineering. Cognitive Work Analysis (CWA) is especially well suited for this task, as it can be used to better understand the context and workflow of users in complex socio-technical domains. It is a conceptual framework that facilitates the analysis of factors that shape human-information interaction and has been used in healthcare for over 20 years. The approach is work-centred, rather than user-centred, and it analyses the constraints and goals that shape information behaviour in the work environment. I used CWA as a framework to help me analyse the problem of data quality in healthcare.

My research uses an instrumental case study approach to understand data quality in primary care. My goal was to answer three questions: In primary care, how are individual users influenced by their environment to input high-quality data? What techniques could be used to design systems that persuade users to enter higher-quality data? Is it possible to improve data quality in primary care by persuading users with the user interface of information systems in these complex socio-technical systems? The scope of work included modelling data quality, defining and measuring data quality in a primary care system, establishing design concepts that could improve data quality through persuasion, and testing the viability of some design concepts.

I began analysing this problem by creating an abstraction hierarchy of patient treatment with medical records. This model can be used to represent patient treatment from a primary care perspective. The model helped explain the patient treatment ecosystem and how data is generated through patient encounters.

After creating my model to represent patient treatment, I incorporated it into two CWAs of data quality and data codification. The first model represented codification in the primary care ecosystem, whereas the second model represented codification in community hospitals. After developing abstraction hierarchies

for both domains, I analysed similar tasks from each system with control task analysis, strategies analysis, and worker competencies analysis. The tasks that I analysed related specifically to data codification: in primary care, I modelled the record encounter task performed by clinicians at a Family Health Team (FHT), and in the community hospital, I modelled the abstract task performed by health information management professionals. I used the same record encounter task at the FHT as a continuing focus of my case study.

I used both models of codification to perform a comparison. My goal was to identify the differences between the ecosystems and tasks that were present in primary care and the community hospital. Comparing CWA models is not a well-defined process in the literature, and I developed an approach to conduct this comparison based on seminal works. I used the approach to systematically compare each phase of my CWA models. I found that the analysis of both system domains in parallel enabled a richer understanding of each environment that may not have been achieved independently. In addition, I discovered that a rich environment exists around data codification processes, and this context influences and distinguishes the actions of users. While the tasks in both domains were seemingly similar, they took place with different priorities and required different competencies.

After building and comparing models, I investigated the summarizing task in primary care more closely by analysing data within a FHT's reporting database. The goal of this study was to understand data quality tradeoffs between timeliness, validity, completeness, and use in primary care users. Data quality measures and metrics were developed through interviews with a focus group of managers. After analysing data quality measures for 196,967 patient encounters, I created baselines, modelled each measure with logit binomial regression to show correlations, characterized tradeoffs, and investigated data quality interactions. Based on the analysis, I found a positive relationship between validity and completeness, and a negative relationship between timeliness and use. Use of data and reductions in entry delay were positively associated with completeness and validity. These results suggested that if users are not provided with sufficient time to record data as part of their regular workflow, they will prioritize their time to spend more time with patients. As a measurement of the effectiveness of a system, the negative correlation between use and timeliness points to a self-reinforcing data repository that provides users with little external value. These findings were consistent with the modelling work and also provided useful insight to study data quality improvements within the system.

I used my measures from the data analysis to select design priorities and behaviour changes that should, according to my ongoing case study, improve data quality. Then I developed several design concepts by combining CWA, a framework for behaviour change, and a design framework for persuasive systems. The design concepts adopted different persuasion principles to change specific behaviours.



To test the validity of my design concepts, I worked with a FHT to implement some of my proposed interventions during a field study. This involved the introduction of a non-invasive summary screen into the user workflow. After the summary screen had been deployed for eight weeks, I received secondary data from the FHT to analyse. First, I performed a pre-post measurement of several data quality measures by doing a simple paired t-test. To further understand the results, I borrowed from healthcare quality improvement methodologies and used statistical process control charts to understand the overall context of the measures. The average delay per entry was reduced by 3.35 days, and the percentage of same-day entries increased by 10.3%. The number of records that were complete dropped by 4.8%. Changes to entry accuracy and report generation were not significant. Several additional insights could be extracted by looking at each the XmR chart for each variable and discussing the trends with the FHT. Feedback was also collected from users through an online survey.

Through the use of a case study spanning several years, I was able to reach the following conclusions: data codification and data quality are manufactured within complex socio-technical systems and users are heavily influenced by a variety of factors within their ecosystem; persuasive design, informed with data from a CWA, is an effective technique for creating ecologically relevant persuasive designs; and data quality in primary care can be improved through the use of these designs in the system's user interface. There are interesting opportunities to apply the results of my work to other jurisdictions. A strength of this work lies in its usefulness for international readers to draw comparisons between different systems and health care environments throughout the world.



## Acknowledgements

The last five years have been remarkable; they have changed the course of my life and career, and I've learned a lot about myself. In 2012, I finished my Masters, started working at Conestoga College, started my PhD program and received an NSERC Industrial Partnership Scholarship. In 2013, I completed my graduate coursework. In 2014, I finished my competency exams, became a full-time faculty member at the College and did my first Ironman in Wisconsin. In 2015, I started working with my industry contacts to collect data for my research. In 2016, I completed most my research and did my second Ironman in Quebec. In 2017, I am proudly finishing my dissertation. Yes, these are the personal achievements of a part time PhD student. However, it would be naive for me to suggest that I could have accomplished any of these goals without the support of my family, friends, and colleagues. Frankly, I am very lucky to be surrounded by great people. It is not possible for me to thank everyone here; without naming you all specifically, know that I am forever grateful for your friendship and encouragement.

I would like to give a sincere thanks to my advisor, Dr. Catherine Burns. Her patience and continuous support were invaluable. As a part-time student, I was successful because Cathy is a flexible, good-natured and reasonable human being. I truly appreciated how she was able to accommodate and encourage my 'sprints' of work when I was able to reserve random weeks of dedicated time; Cathy endured many spurts of e-mails from me with good humour. Cathy was always willing to provide me with quick feedback when I was on a roll, but was also willing to provide me space when I had to focus my attention elsewhere. I attribute the successful completion of my doctorate directly to my excellent working relationship with Cathy. She is well-deserving of her reputation and respect in the human factors community, and I could not have chosen a better mentor to guide me through this process. I look forward to future opportunities where I can collaborate with Cathy on new projects, and aspire to someday be as helpful a mentor to others as she has been to me.

I want to thank my Executive Dean, Marlene Raasok (now retired) and my Chair, Curt Monk. As a part-time student and full-time employee at Conestoga College, I needed flexibility and support to successfully complete my work. I am proud to work at the College and sincerely appreciate the encouragement, time and funding that has been invested towards helping me achieve my goals.

This work could not have taken place without a network of peers in the industry. I want to thank my friends from the vendor, academic, primary care and hospital community who were always excited to meet with me and support me with insights and ideas. Specifically, I would like to thank Justin Wolting and his team of students for their efforts over several years. I would also like to thank Shawn Vincent and

acknowledge the financial support provided by TELUS Health and the Natural Sciences and Engineering Research Council of Canada (NSERC) through my Industrial Partnership Scholarship in 2012.

In chapter 8, I created several mockups. I would like to give attribution to the artists of the icons I used and recognize Freepik, Vaadin, Dave Gandy, Madebyoliver, Pixel Buddha and Anton Saputro from [www.flaticon.com](http://www.flaticon.com). As well, I want to thank Dr. Trego and Cambridge Proofreading LLC for proofing my final manuscript. They corrected verb-tense grammar issues and identified typos. They also highlighted run-on sentences, and helped me consistently use Canadianized English and grammar conventions. I am thankful for their assistance in preparing a document that is generally free of distracting grammar issues.

Last but not least, I want to thank my wife for her support in this long pursuit. In addition to helping me with various edits and grammatical stuffs, she has listened to my babblings about modelling and data quality over the course of a Masters and a PhD, and seems to love me anyways.

*To my friends and colleagues.  
Your desire to engage and explore  
with me is very much appreciated.*

*To my family.  
Thank you for normalizing  
academia and making me feel  
like all the cool kids are doing it.*

*To my wife.  
I love you a lot.*



## Table of Contents

<b>Examining Committee Membership</b> .....	<b>iii</b>
<b>Author’s Declaration</b> .....	<b>iv</b>
<b>Statement of Contributions</b> .....	<b>v</b>
<b>Abstract</b> .....	<b>vii</b>
<b>Acknowledgements</b> .....	<b>xi</b>
<b>Dedications</b> .....	<b>xi</b>
<b>Tables of Content</b> .....	<b>xv</b>
<b>List of Figures</b> .....	<b>xi</b>
<b>List of Tables</b> .....	<b>xxi</b>
<b>List of Acronyms</b> .....	<b>xxiv</b>
<b>Part A Introduction and Background</b> .....	<b>1</b>
<b>Chapter 1 Introduction</b> .....	<b>3</b>
1.1 Scope of Work.....	6
1.2 Research Questions and Objectives.....	7
1.3 Overview of Research Methods .....	7
1.3.1 Case Study .....	7
1.3.2 Modelling.....	9
1.3.3 Theory Development .....	10
1.4 Structure of the Dissertation.....	10
<b>Chapter 2 Primary Care in Ontario</b> .....	<b>13</b>
2.1 Primary Care and the Biopsychosocial Model .....	13
2.2 Governance Structures.....	14
2.3 Medical Records, Registries, and Billing .....	15
2.3.1 Medical Records .....	16
2.3.2 Billing and Reporting .....	17
2.4 Data Quality.....	18
2.4.1 Data Quality Dimensions in Primary Care .....	19
2.4.2 Previous Work .....	22
2.5 A Complex Socio-Technical System.....	24
<b>Chapter 3 Cognitive Work Analysis</b> .....	<b>25</b>
3.1 Overview .....	25
3.1.1 Work Domain Analysis .....	26
3.1.2 Control Task Analysis .....	26

3.1.3 Strategies Analysis .....	28
3.1.4 Social Organization and Co-operation Analyses .....	28
3.1.5 Worker Competencies Analysis .....	28
3.2 Ecological Interface Design.....	29
3.3 Applications of Cognitive Work Analysis.....	29
3.4 Use of Cognitive Work Analysis to Study Data Quality .....	30
<b>Part B Systems Analysis and Modelling.....</b>	<b>31</b>
<b>Chapter 4 An Abstraction Hierarchy of Patient Treatment .....</b>	<b>33</b>
4.1 An Abstraction Hierarchy of Patient Care.....	33
4.1.1 Model Objective and Scope.....	34
4.1.2 Intended Uses .....	35
4.2 Model Development .....	35
4.2.1 Model Context .....	35
4.2.2 Information Gathering .....	35
4.2.3 Abstraction Hierarchy Development .....	36
4.3 Modelling Results.....	38
4.3.1 Functional Purpose and Treatment Goals.....	38
4.3.2 Abstract Functions and Treatment Constraints.....	41
4.3.3 Generalized Treatment Processes .....	44
4.3.4 Physical Treatment Elements and Attributes.....	44
4.4 Discussion.....	45
4.4.1 Comparisons to Other Models.....	45
4.4.2 Design Implications .....	45
4.4.3 Limitations.....	45
4.4.4 Future Work.....	46
4.5 Chapter Summary .....	47
4.5.1 Key Findings.....	47
4.5.2 Connections to Case Study and Research Questions.....	47
4.5.3 Connections to Other Chapters .....	47
<b>Chapter 5 Data Codification in Healthcare .....</b>	<b>49</b>
5.1 Two Healthcare Settings.....	49
5.1.1 The Family Health Team .....	50
5.1.2 The Community Hospital .....	51
5.2 Methods .....	52



5.2.1 Boundary and Scope .....	52
5.2.2 Information Gathering .....	53
5.2.3 Cognitive Work Analysis .....	53
5.3 Results .....	54
5.3.1 Work Domain Analysis .....	54
5.3.2 Control Task Analysis .....	54
5.3.3 Strategies Analysis .....	55
5.3.4 Social Organization and Cooperation Analysis .....	55
5.3.5 Worker Competencies Analysis .....	55
5.4 Discussion.....	55
5.5 Chapter Summary .....	56
5.5.1 Key Findings.....	56
5.5.2 Connections to Case Study and Research Questions.....	56
5.5.3 Connections to Other Chapters.....	57
<b>Chapter 6 Comparing Cognitive Work Analysis Models .....</b>	<b>59</b>
6.1 Introduction .....	59
6.2 Comparisons with Cognitive Work Analysis .....	62
6.2.1 Why CWA for comparisons?.....	62
6.2.2 How to Compare CWA Outputs? .....	62
6.3 Comparison of Two Codification Systems.....	66
6.3.1 Comparison Method .....	66
6.3.2 Results .....	68
6.3.3 Discussion.....	73
6.3.4 Future Work and Limitations.....	77
6.4 Suggestions and Recommendations .....	78
6.5 Chapter Summary .....	79
6.5.1 Key Findings.....	79
6.5.2 Connections to Case Study and Research Questions.....	79
6.5.3 Connections to Other Chapters.....	80
<b>Part C Measurement, Design, and Improvement.....</b>	<b>81</b>
<b>Chapter 7 Measurement of Data Quality in Primary Care.....</b>	<b>83</b>
7.1 An Exploratory Case Study to Understand Primary Care Users and Their Data Quality Tradeoffs.....	83
7.1.1 Study Purpose .....	84

7.1.2 Methods .....	85
7.1.3 Results .....	91
7.1.4 Discussion.....	96
7.1.5 Conclusions .....	106
7.2 Chapter Summary .....	107
7.2.1 Key Findings.....	107
7.2.2 Connections to Case Study and Research Questions.....	108
7.2.3 Connections to Other Chapters .....	108
<b>Chapter 8 Design Concepts for Data Quality Improvement .....</b>	<b>109</b>
8.1 Persuasive Design.....	109
8.1.1 Using Social Persuasion with Clinicians .....	109
8.1.2 Persuasive Technology .....	109
8.1.3 Fogg’s Behavioural Change Model .....	110
8.1.4 Persuasive System Design Model.....	111
8.1.5 Literature Gaps .....	112
8.2 Defining the Persuasion Context with Cognitive Work Analysis .....	112
8.2.1 Who is the target? .....	114
8.2.2 What is the change? .....	114
8.2.3 Why is the task completed? .....	115
8.2.4 When will the change occur?.....	115
8.2.5 Answering ‘How’ by Identifying the Persuasion Strategy .....	117
8.3 Persuasive Designs to Improve Data Quality .....	117
8.3.1 Goal 1: Reduce Entry Delay .....	117
8.3.2 Goal 2: Increase Completeness and Accuracy, and Enhance Believability.....	120
8.3.3 Goal 3: Increase Data Engagement and Use.....	123
8.4 Discussion.....	125
8.4.1 An Ecological Approach to Persuasive Design .....	125
8.4.2 Limitations.....	125
8.4.3 What about Ecological Interface Design? .....	126
8.5 Chapter Summary .....	126
8.5.1 Key Findings.....	126
8.5.2 Connections to Case Study and Research Questions.....	126
8.5.3 Connections to Other Chapters .....	127
<b>Chapter 9 Improving Data Quality in Primary Care .....</b>	<b>129</b>

9.1 Field Study.....	129
9.2 Experimental Design .....	130
9.2.1 Pre-Post Test.....	131
9.2.2 Statistical Process Control .....	131
9.3 Methods .....	133
9.3.1 Ethics .....	133
9.3.2 User Interface Design .....	133
9.3.3 Data Quality Measures .....	134
9.3.4 Deployment .....	135
9.3.5 Data Collection .....	135
9.3.6 Pre-Post Analysis.....	136
9.3.7 XmR Charts .....	136
9.3.8 User Survey .....	136
9.4 Results .....	137
9.4.1 User Interface .....	137
9.4.2 Deployment .....	137
9.4.3 Data Collection .....	137
9.4.4 Pre-Post Analysis.....	138
9.4.5 XmR Control Charts.....	139
9.4.6 User Feedback .....	142
9.5 Discussion.....	142
9.5.1 User Interface .....	142
9.5.2 Time Interval for XmR .....	144
9.5.3 Data Analysis.....	145
9.5.4 User Feedback .....	149
9.5.5 Primary Care User Hypotheses.....	152
9.5.6 Limitations.....	152
9.5.7 Future Work.....	153
9.6 Chapter Summary .....	155
9.6.1 Key Findings.....	155
9.6.2 Connections to Case Study and Research Questions.....	156
9.6.3 Connections to Other Chapters.....	156
<b>Part D Conclusion .....</b>	<b>159</b>
<b>Chapter 10 Conclusions .....</b>	<b>161</b>

10.1 Case Study and Research Questions.....	161
10.1.1 How are individual users influenced by their environment to input high-quality data?	161
10.1.2 What techniques could be used to design systems that persuade users to enter higher-quality data?.....	161
10.1.3 Is it possible to improve data quality in primary care by persuading users with the user interface? .....	161
10.1.4 Other Findings .....	161
10.2 Summary of Contributions .....	162
10.2.1 Contributions to Human Factors Engineering .....	162
10.2.2 Contributions to Health Informatics and Data Quality .....	163
10.3 Generalization of Findings .....	163
10.4 Future Research.....	164
10.4.1 Other Areas of Healthcare .....	164
10.4.2 Better Persuasion for Better Data .....	164
10.4.3 Comparison of Additional Systems .....	165
<b>References.....</b>	<b>167</b>
<b>Appendix A Abstraction Hierarchies of Patient Treatment (AH 1.A &amp; 1.B).....</b>	<b>185</b>
<b>Appendix B Work Domain Analysis of Data Codification.....</b>	<b>187</b>
Abstraction Hierarchy of Data Codification in a Family Health Team (AH 2.1) .....	187
AH 2.1 - Functional Purpose (FP).....	188
AH 2.1 - Abstract Function (AF).....	189
AH 2.1 - Generalized Function (GF).....	190
AH 2.1 - Physical Function (PF) .....	191
AH 2.1 - Physical Form (PFo).....	192
Abstraction Hierarchy of Data Codification in a Community Hospital (AH 3.1).....	193
AH 3.1 - Functional Purpose (FP).....	194
AH 3.1 - Abstract Function (AF).....	195
AH 3.1 - Generalized Function (GF).....	196
AH 3.1 - Physical Function (PF) .....	197
AH 3.1 - Physical Form (PFo).....	198
<b>Appendix C Control Task Analysis Decision Ladders and Summary Tables.....</b>	<b>199</b>
Summarize Encounter in Registry – Baseline (Control Task 1.1).....	199
Summarize Encounter in Registry – Novice User (Control Task 1.1.1) .....	202
Summarize Encounter in Registry - Intermediate User (Control Task 1.1.2) .....	205

Summarize Encounter in Registry – Expert User (Control Task 1.1.3) .....	209
Abstract Encounter to Registry - Baseline (Control Task 2.1) .....	212
Abstract Encounter to Registry – Novice User (Control Task 2.1.1) .....	215
Abstract Encounter to Registry – Expert User (Control Task 2.1.2) .....	219
<b>Appendix D Strategies Analysis Flow Maps and Summary Tables .....</b>	<b>222</b>
Summarize Encounter in Registry (Control Task 1.1) .....	222
Abstract Encounter to Registry (Control Task 2.1) .....	226
<b>Appendix E Worker Competencies Analysis Tables .....</b>	<b>228</b>
<b>Appendix F Qualitative Results of Data Quality Measure Interviews.....</b>	<b>232</b>
<b>Appendix G Registry System Screenshots.....</b>	<b>234</b>
<b>Appendix H Additional Quantitative Data.....</b>	<b>235</b>
<b>Appendix I Data Quality Design Worksheets .....</b>	<b>237</b>
<b>Appendix J Data Quality Summary Screenshot .....</b>	<b>239</b>
<b>Appendix K Survey Questions.....</b>	<b>240</b>
<b>Appendix L XmR Charts by Month .....</b>	<b>242</b>
<b>Appendix M Survey Results .....</b>	<b>246</b>

## List of Figures

Figure 1. Overview of research focus by chapter. ....	9
Figure 2. Summary of Cognitive Work Analysis applications and uses.....	30
Figure 3. General patient process functions.....	37
Figure 4. General information process functions.....	38
Figure 5. Codification processes at a Family Health Team. ....	50
Figure 6. Codification processes at a community hospital. ....	52
Figure 7. Flagged differences on Abstraction Hierarchy.....	54
Figure 8. Two paths on a common Decision Ladder (St-Maurice & Burns, 2015). ....	65
Figure 9. Opportunity at a junction (St-Maurice & Burns, 2015).....	67
Figure 10. Venn diagram of differences in the WDAs. ....	68
Figure 11. Junction at System State in ‘Record Encounter’ task.....	70
Figure 12. Junction at Alert State in ‘Record Encounter’ task. ....	70
Figure 13. A use of the praise dialogue principle to reduce delay.....	119
Figure 14. A suggestion dialogue to reduce delay.....	119
Figure 15. A simple use of the social facilitation principle to reduce entry delay.....	120
Figure 16. A scoring mechanism to encourage self-monitoring.....	122
Figure 17. A confirmation screen to develop credibility.....	122
Figure 18. A comparison of an arbitrary indicator to improve engagement.....	124
Figure 19. Social facilitation of visit volume to improve engagement.....	125
Figure 20. New User Experience Workflow.....	130
Figure 21. XmR for Timeliness Measure (Recording Delay, Raw). ....	139
Figure 22. XmR for Timeliness Measure (Recording Delay, Adjusted). ....	139
Figure 23. XmR for Timeliness Measure (Percent Same-Day Entry). ....	140
Figure 24. XmR for Completeness Measure.....	140
Figure 25. XmR for Validity Measure.....	141
Figure 26. XmR of Use Measure.....	141

## List of Tables

Table 1. Comparison of Work Domains. ....	69
Table 2. Comparison of Strategy Characteristics.....	71
Table 3. Comparison of Competencies. ....	72
Table 4. Open Ended Questions Used with Focus Group.....	88
Table 5. Data Quality Measures.....	91
Table 6. Data Summary by Program Area.....	93
Table 7. Measure Averages.....	93
Table 8. Simple Correlations (Pearson’s $r$ ).....	94
Table 9. Data Quality Model Results.....	94
Table 10. Coefficients for Completeness Measure. ....	95
Table 11. Coefficients for Timeliness Measure. ....	95
Table 12. Coefficients for Use Measure. ....	96
Table 13. Summary of Random Effect Coefficients.....	101
Table 14. Linking Frameworks for an Ecological Approach to Persuasive Design. ....	113
Table 15. Persuasion Context to Reduce Entry Delay. ....	118
Table 16. Persuasion Context to Increase Completeness, Accuracy and Enhance Believability. ....	120
Table 17. Persuasion Context to Enhance Use and Engagement.....	123
Table 18. Field Study Measures.....	135
Table 19. Pre vs. Post Results with Paired T-Test.....	138

## List of Acronyms

AFHTO: Association of Family Health Teams of Ontario  
ABE: Audit Based Education  
AHP: Allied Health Professional  
BCSS: Behavior Change Support Systems  
CCWA: Comparative Cognitive Work Analysis  
CHA: Canada Health Act  
CHIMP: Certified Health Information Management Professional  
CIHI: Canadian Institute for Health Information  
ConTA: Control Task Analysis  
CTA: Cognitive Task Analysis  
C2TA: Comparative Cognitive Task Analysis  
CPSO: College of Physicians and Surgeons of Ontario  
CWA: Cognitive Work Analysis  
CSA: Canadian Standards Association  
DL: Decision Ladder  
EID: Ecological Interface Design  
EMR: Electronic Medical Record  
EHR: Electronic Health Record  
EPEP: EMR Practice Enhancement Program  
FBM: Fogg Behaviour Model  
FHG: Family Health Group  
FHN: Family Health Network  
FHO: Family Health Organization  
FHT: Family Health Team  
HIM: Health Information Management  
HIMPs: Health Information Management Professionals  
HQO: Healthcare Quality Ontario  
HTA: Hierarchal Task Analysis  
IFM: Information Flow Map  
LHIN: Local Health Integration Network  
MOHLTC: Ministry of Health and Long Term Care  
NP: Nurse Practitioner



OHIP: Ontario Health Insurance Plan  
PD: Persuasive Design  
PDT: Persuasive Design Techniques  
PHIPA: Personal Health Information Protection Act (2004)  
PSD: Persuasive Systems Design  
QIDSS: Quality Improvement Decision Support Specialist  
RCT: Randomized control trial  
RN: Registered Nurse  
RPN: Registered Practical Nurse  
SOCA: Social Organization and Co-operation Analyses  
SRK: Skills, Rules, Knowledge taxonomy  
SME: Subject Matter Expert  
StrA: Strategies Analysis  
SPC: Statistical Process Control  
SUS: System Usability Scale  
UI: User Interface  
UX: User Experience  
WCA: Workers Competency Analysis  
WDA: Work Domain Analysis



*Messieurs et madames, soyez fiers de voter intelligence tant qu'il vous plaira; mais défier-vous beaucoup de l'autre, sur-tout quand vous êtes ensemble.*

*[Ladies and gentlemen, be proud of your intelligence as much as you'd like; but challenge yourself with others, especially when you're together]*

**Xavier de Maistre, 1794**

*And what is the soul, if not an abstract function with a functional purpose?*

**Dr. Catherine Burns, Personal Communication, October 2, 2016**



## **Part A**

### **Introduction and Background**

Part A of my dissertation is an introductory section that includes chapter 1 to chapter 3. These chapters introduce my dissertation and provide background information.



# Chapter 1

## Introduction

The consequences of poor-quality data are expensive, particularly in an era when governments around the world invest heavily in data collection and data management. In 2013, the United States Government budgeted \$52.6 billion dollars for National Intelligence Programs (Gellman & Miller, 2013). In Ontario, several billion dollars were invested to develop an electronic health record (EHR) for Ontarians (CBC News, 2009). In the United Kingdom, \$20 billion dollars were invested in a national EHR strategy. Despite these significant funding initiatives, President Obama could not obtain the reports he needed in a timely manner (Gellman & Miller, 2013); the auditor in Ontario concluded there was poor value for money (CBC News, 2009); and the Cameron government in the UK completely cancelled the eHealth program due to a lack of realizable benefits (Webster, 2012). Obviously, the economic and social impacts of poor-quality data are measured in the billions of dollars (Maguire, 2007; Strong, Lee, & Wang, 1997). Despite these challenges, the 21<sup>st</sup> century gold rush is still underway (St-Maurice, 2011), and leaders assume that the future benefits arising from investments in data are worthwhile pursuits; the implicit assumption in government is that the benefits of available data eventually will outweigh the costs.

Data quality challenges consist of different types. Some direct and indirect costs, such as operational inefficiencies, are caused by ‘simple’ data issues, such as incorrect name matching, information duplication, and interoperability problems (Batini & Scannapieco, 2006). Other problems are more complex and depend on the design and production processes involved in generating data. In these instances, it is important to design for better quality by understanding what quality how it is measured (Wand & Wang, 1996).

Healthcare is classified as a complex socio-technical system by virtue of its characteristics (Vicente, 1999) and is not immune to data quality challenges. As primary care becomes technologically complex with the increased use of Electronic Medical Record (EMR) systems (Gagnon et al., 2010), primary care is no exception to this paradigm. In addition, the field of primary care informatics faces several unique and interesting challenges, such as seeing patients with combinations of problems, providing continuous care, and incorporating a biopsychosocial scope of care (de Lusignan, 2003).

Primary care is also moving into patient-centred care models that adopt Shared Decision Making (SDM) with patients and their families (Legare & Witteman, 2013). Data requirements and the strategies needed to capture data efficiently are increasingly challenging in primary care.

For example, correctly labelling and coding ‘easy’ diseases such as diabetes (which has relatively clean diagnostic criteria) is challenging in primary care. Mislabeling the disease can result in impacts to treatment options, risk management, and psychological and financial issues for the provider and patient. These data problems have also had an impact on the validity of quality-of-care evaluations and research projects (Stone et al., 2010). As free text narratives can suffer from regionalisms, colloquialisms, and significant typos (St-Maurice, Kuo, & Gooch, 2013), capturing structured data through forms or dropdown menus is an easy approach to generating usable, high quality data but generally requires more clinician time (Sulmasy et al., 2017).

Capturing high quality data in primary care through the use of structured fields is not a technical challenge and can be easily introduced into software solutions. However, capturing data through structured data fields is still a challenging problem. From the clinician’s perspective, there are cultural barriers to entering structured data (de Lusignan et al., 2003). As well, there are ongoing debates about the ethics of prompting clinicians to take time to structure their data for secondary purposes, at the potential expense of direct patient care (Sulmasy et al., 2017). Physicians are often unmotivated to structure and code their data unless they are personally interested in helping with external research activities (Butz, Brick, Rinehart-Thompson, Brodник, & Agnew, 2016).

Generally, the challenge of changing data entry behaviours to capture better data is not technical in nature, and is better attributed to clinician social structures, culture, and work priorities. By contextualizing data quality as a complex social-technical problem, the breadth of potential solutions grows significantly. Instead of focusing on purely technical approaches to the problem (such as IBM-Watson), leveraging social processes to change data quality behaviours could be much more effective.

In the literature, there are many examples of encouraging change in healthcare through persuasive processes. In one instance, data quality was improved through the implementation of a competitive audit program in the United Kingdom. As an instrument, Audit Based Education (ABE) is an educational element consisting of local meetings where comparative data is presented to representatives of individual physician practices. In this approach, summary data was provided to clinical leads for comparison (de Lusignan, Belsey, Hague, Dhoul, & van Vlymen, 2006). ABE and its iterative comparisons between individual practices ultimately engaged physicians and produced higher-quality data. Although it is not clear if this approach improved care, or simply improved the quality of data regarding current practice, the approach has been shown to improve recorded cholesterol management, and was described as the most successful change agent toward clinician attitudes regarding data quality (de Lusignan, 2005). More



recently, researchers in the Netherlands developed a data quality feedback tool that evaluated differences in EHR data quality. The software package was part of a larger intervention and compared data quality measures between practices. The tool generated data quality feedback, and recording quality improved significantly (van der Bij, Khan, ten Veen, de Bakker, & Verheij, 2016).

In Ontario, a similar approach to the ABE approach was used in Kingston and achieved similar outcomes. In this example, a simple metric was employed to show the percentage of each physician's eligible patient roster that had received the annual flu shot. Each physician took charge to achieve the highest percentage vis-à-vis their colleagues. This implementation of ABE motivated the clinic and resulted in achieving a second place standing overall in the province for achieving flu shot targets (D. Barber, personal communication, October 10, 2012). It is interesting to note that the physician leaders in this study characterized the approach as a competition between colleagues, and not as a formal implementation of ABE. However, the approach matched de Lusignan's (2005) description and use of ABE.

The concept of creating a quality improvement collaborative is another example of introducing effective change by leveraging social processes (Patterson, Schweikhart, Anders, Brungs, & Render, 2007). As an example, researchers observed a positive change in behaviour when analysing the effects of an integrated decision support system. The system measured the likelihood of erroneous orders by scoring decisions based on an algorithm and presented this information to users on the system interface. Orders that did not appear to follow best practice were flagged, and users were given an aggregate scorecard. Initially the scores were low (e.g., many orders did not follow best practice), but scores gradually increased through subsequent reporting and feedback (Weilburg et al., 2009). Anecdotally, it was noted that part of this positive change was due to physicians reacting to seeing each other's scores; they were described as competing to ensure they were not scoring below the average (P. Nagy, personal communication, October 30, 2012). The results of the system showed significant improvements over time. Although the example is anecdotal, the improvements in the metrics were a result of the system's display of comparative data and the adoption of a quality improvement collaborative program.

Introducing change in primary care to improve data quality is unlikely to be successful with a techno-centric approach. Leveraging social processes has been shown to be effective at changing practice, and there are examples of improving data quality through this paradigm. From the standpoint of an engineer aspiring to improve the problem of data quality in primary care, engineering methods that incorporate the analysis of both social and technical aspects of a problem seem appropriate. Thus, the challenge of data quality and its improvement is well suited for the domain of systems design and human factors engineering. In addition, the previous examples demonstrate, either intentionally or accidentally, the use of persuasion to change attitudes or behaviours of users through social influence. Although the

previous studies do not describe their work as persuasive approaches per se, they exhibit many of its characteristics. Therefore, it appears that data quality could be improved by social processes, and that these processes could be designed into technology through persuasive design (PD).

## **1.1 Scope of Work**

Data and data quality in primary care have many facets. Data can be entered into systems as free text, unstructured narrative data, or structured data. When data are unstructured, it allows a clinician to articulate thoughts and observations freely based on the patient encounter. Information entered into charts as free text appeals to many clinicians, but has the disadvantage of being difficult to search and use consistently.

While it is theoretically possible to extract rich information from unstructured sources of data with Natural Language Processing (NLP) tools, the results often are difficult to use. For example, the concept of poverty (e.g., to be poor) in free text could be a useful social indicator. However, the term can be used in documentation as an adjective (e.g., poor sleep, poor eyesight). The intended meaning of the sentence is not a poor person's sleep or eye-sight, but a lower-quality eyesight or sleep quality. Other examples involved the term 'kicking', which was found used in several colloquial ways: 'patient was kicked out of the mental health support group', 'patient was kicked in the face at the bar', 'tried to leave a message for the patient but her answering machine did not kick in', and 'pacemaker will kick in if heart rate falls below 60'. Practically, effective NLP relies on effective underlying dictionaries that currently are not capable of understanding the broadness of primary care progress notes or free text data sources (St-Maurice et al., 2013).

Structured data often becomes captured by users via forms or by associating common data elements to free text entries. In electronic environments, entering data into forms includes using a combination of check marks, labelled text boxes, or dropdowns to input discrete data. These electronic forms sometimes are known as e-Forms, or custom forms. This data is easier to search and use for secondary purposes and does not provide users with an opportunity to enter nuanced or colloquial notes. The scope of my research includes the analysis of processes and considerations involved with codifying data and entering structured data through forms. The scope of work includes analysing clinicians who, in primary care, are tasked with these data entry responsibilities. Coded data is an ideal point of focus for this research because many instances of poorly coded data quality occur in primary care and there are many opportunities for improvement (de Lusignan et al., 2010), coded data is the primary source of secondary use of records in primary care, and coded data lends itself well to systematic analysis as part of a study. Codifying data in this context includes associating a standard reference terminology, such as ICD-10-CA, to data elements.

The quality of unstructured text progress notes or problem lists within the medical record is not part of my current research program. I am also excluding the quality of lab data, biometric measures, system generated data, and the content system messages, such as HL7 messages. Furthermore, I am not including the challenges associated with information exchanges between systems.

## **1.2 Research Questions and Objectives**

I have three research questions:

1. In primary care, how are individual users influenced by their environment to input high-quality data?
2. What techniques could be used to design systems that persuade users to enter higher-quality data?
3. Is it possible to improve data quality in primary care by persuading users with the user interface?

Based on these research questions, my research objectives are to:

1. Try to model data quality as the output of a complex socio-technical problem.
2. Compare the primary care data ecosystem and other data ecosystems in healthcare and identify policies, procedures, structures, and workflows that could improve data quality in primary care.
3. Develop an approach to define data quality in context and determine how it can be measured.
4. Understand primary care users, their tradeoffs, and their approaches to codifying data.
5. Develop persuasive approaches to help improve the quality of data in primary care.
6. Test design interventions and demonstrate that the design concepts developed through modelling are viable approaches to improve data quality.

## **1.3 Overview of Research Methods**

I worked through my research questions by starting with a high-level perspective of my work domain, and diving deeper between each study to better understand factors that might impact data quality and structured data entry. I studied my context using Cognitive Work Analysis (CWA) and developed approaches and tools to help answer my research questions.

### **1.3.1 Case Study**

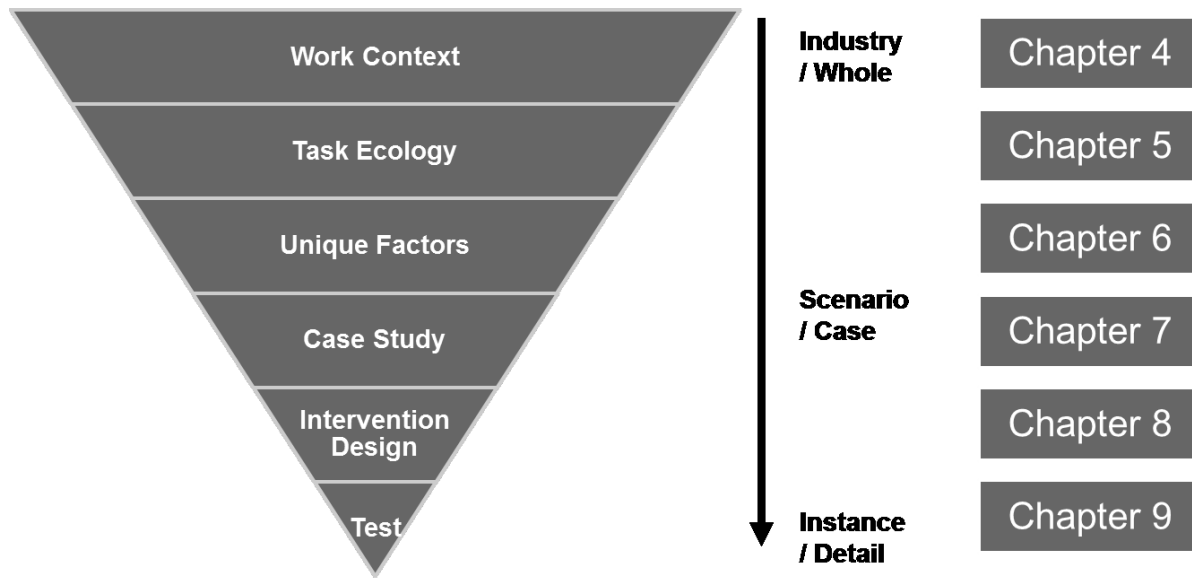
A case study design should be considered in the following instances: when either the focus of a study is to answer ‘how’ and ‘why’ questions, or the behaviour of the participants cannot be changed, or when you want to cover contextual conditions because you believe they are relevant to the phenomenon under study, or when the boundaries are not clear between the phenomenon and context (Baxter & Jack, 2008). Several types of case studies, including explanatory, exploratory, descriptive, intrinsic, and instrumental (Baxter & Jack, 2008), can be employed. Case studies are widely used in data quality

research (Madnick, Wang, Lee, & Zhu, 2009). Case studies are empirical methods that use a combination of qualitative and quantitative evidence to examine phenomena in their real-world contexts. The in-depth inquiry of a single instance can lead to a deeper understanding of the subject, generate useful information, and generate hypotheses (Flyvbjerg, 2006).

Having a combination of quantitative data in primary care and access to several Subject Matter Experts (SMEs), much of my work is threaded together as an instrumental case study approach. Instrumental case studies are used as instruments to accomplish a primary goal and provide insights into an issue. An instrumental case study plays a supportive role and can facilitate the understanding of something else. The case often is examined in-depth, and the study includes an analysis of context and ordinary activities and helps the researcher pursue an external interest (Baxter & Jack, 2008). In my research, I am trying to understand how users are influenced by their environment. However, my primary interest is to identify strategies to improve data quality. Thus, an instrumental case study is appropriate in this context. I extended the instrumental case study concept by including a quasi-experimental pre-post study design in the same organization as the case study.

My case study involved the study of a data codification system at a Family Health Team (FHT). I began by understanding the general patient treatment domain. Then, I took a closer look at the data codification context and environment, and tried to understand the factors that impact users. I compared different data codification domains to understand what made users in primary care unique, and similar, compared to hospital users. After understanding the task's context and environment, I analysed a specific case of data and data quality, and established some measurements and benchmarks. I used this information to identify specific behaviour changes that related to data quality. Finally, I tested an intervention on a user interface with a FHT and achieved a high resolution on the impact of PD elements on data quality. An overview of my work, from high level to case specific to a specific intervention, is shown in Figure 1. This defines the 'umbrella' of my work and research context.

Figure 1. Overview of research focus by chapter.



All my research activities enhanced my knowledge of real-world data quality phenomena in primary care and helped me to collect useful information for answering my research questions. I refer to my research objectives and show how my results related to the overarching case study throughout my dissertation.

### 1.3.2 Modelling

To better understand data in primary care, my goal was to study the flows, processes, and caveats that play important roles in improving or degrading its quality. In this sense, I am studying users, their environment, and the relationship between the two, and I am conducting an ecological study of primary care users. To effectively structure my approach and document my results, I needed an appropriate framework. Although systems design can be accomplished using many approaches, CWA is especially suited to better understand the context and workflow of users in complex socio-technical domains.

CWA is a conceptual framework that facilitates the analysis of factors that shape human-information interaction and has been used in healthcare for over 20 years (Jiancaro, Jamieson, & Mihailidis, 2013). The approach is work-centred, rather than user-centred, and it analyses the constraints and goals that shape information behaviour in the work environment. CWA examines the environmental, organizational, social, activity, and individuals in the workflow. As an output, CWA provides concepts and templates to facilitate an analysis of complex phenomena and supports the improved design of information systems (Fidel & Pejtersen, 2004). Traditionally, the results of a CWA are applied to design through the use of Ecological Interface Design (EID). EID aims to design interfaces that do not contribute to the difficulty of the task, but reveal the relationships important to successful performance in the work

domain (Vicente & Rasmussen, 1992). Many examples of using EID successfully in a healthcare setting can be given (Momtahan & Burns, 2004). Current work in CWA includes broadening the areas in which it is applied, and enriching the framework and analytical approach (Jiancaro et al., 2013).

As part of a case study, CWA is an excellent framework that supports the idea of uncovering contextual conditions relevant to the phenomena under study (Baxter & Jack, 2008). CWA provides a good way of documenting the findings of a case study and relating qualitative findings back to systems design.

### **1.3.3 Theory Development**

My theory work, including developing comparison approaches and linking CWA to design principles, was done out of necessity to better understand my real-world system. I had specific use cases and analytical goals that presently were not addressed in the literature.

For example, upon completing my CWA models, I wanted to compare my two sets of models to understand some of the subtle (but important) differences between data codification and structured data capture in primary care and community hospitals. Such an approach does not exist, so I needed to engage in theory work and determine an appropriate way to systematically compare Abstraction Hierarchies (AHs), Decision Ladders (DLs), Information Flow Maps (IFMs), and Skill-Rule-Knowledge (SRK) taxonomies. I interpreted the differences as insights and identified characteristics that could be transferred to improve data quality in primary care. These new tools helped me to understand the contextual conditions of data quality, and determine the ecological factors that affect users in different, but similar, domains.

Another situation in which I required a new approach occurred when I was trying to define a persuasion context for PD. To date, no work has linked the information from a CWA to PD principles. To address this need, I created an approach to link different phases of CWAs to different aspects of the persuasion context. This allowed me to identify PD principles that could target specific behaviour changes that were aimed to improve data quality.

## **1.4 Structure of the Dissertation**

My dissertation is a combination of papers under review and original work I plan to submit for publication. My dissertation is broken into four logical pieces: an introduction and general background; my modelling work; my field work; and a conclusion.

Part A of my dissertation is an introductory section that includes this chapter, chapter 2, and chapter 3. While each chapter will provide a literature review of relevant topics, I provide a general background of primary care in Ontario in chapter 2. I describe terms and concepts that should be

understood to appreciate the primary care ecosystem in Ontario and the dissertation as a whole. This includes defining primary care and the biopsychosocial model, reviewing primary care governance structures in Ontario, discussing medical records and data primary care, and discussing primary care data quality. Chapter 3 provides an overview of CWA for readers who are unfamiliar with this approach to systems analysis.

Part B of my dissertation spans chapters 4 to 6 and focuses on domain analysis and my modelling work. Chapter 4 details my analysis of the patient treatment domain and my development of an Abstraction Hierarchy (AH) to model treatment with medical records. Chapter 4 was adapted with minor edits and submitted to the *Journal of Medical Internet Research (JMIR) Human Factors* and has been accepted for publication. Chapter 5 builds on the treatment AH and describes the development of two CWAs related to data codification. Chapter 6 describes an approach for comparing two CWA models and offers insights based on the comparison of my CWA of codifications in the hospital and primary care. The content from chapters 3, 5, and 6 was summarized and submitted to the journal, *Theoretical Issues in Ergonomic Science*, and is under review at the time of this writing.

Part C of my dissertation spans chapters 7 to 9 and focuses on data quality measurement, interface design and improvement. In chapter 7, I explore data quality measurements for the data reporting system for a FHT. I discuss an approach to define data quality measures and build data quality models through logistic regression. This chapter was accepted for publication as an original paper in the *Journal of Data and Information Quality*. In chapter 8, I describe several socio-technical design concepts by referring to my models and using PD. I argue that these PD interventions can improve data quality in primary care. Finally, in chapter 9, I describe a field study where I deployed my design concepts at a FHT. I describe experimental results that measure the effectiveness of my intervention and describe the impacts of my design on the data quality of a real system. The results presented in this chapter were submitted for publication in the *International Journal of Medical Informatics (IJMI)*.

The final part of my dissertation is a conclusion in chapter 10. In this chapter, I combine my findings, provide a summary of my contributions, and articulate exciting opportunities for future work.





## **Chapter 2**

### **Primary Care in Ontario**

In this chapter, I provide readers with some background and context to my work by describing important elements of the primary care context and coding environment in Ontario. This includes defining biopsychosocial concepts in primary healthcare; reviewing primary care in Ontario by discussing its governance structures, funding, and institutions; discussing data generation and record keeping in primary care; and discussing primary care data quality. This chapter defines and clarifies various terms, actors, institutions, and relationships that are germane to the remainder of my dissertation.

#### **2.1 Primary Care and the Biopsychosocial Model**

Primary care is an important component of the health ecosystem. It has complex processes, procedures, and technological caveats. The field is strategically important to the evolving health care industry, as primary care data can be more inclusive of a patient's health record compared to medical records in other environments; the data may extend from the patient's birth until their death. In an electronic form, primary care data can be used systematically for audit, quality improvement, health service planning, epidemiological study, research, and to measure the quality of care (de Lusignan, Stephens, & Majeed, 2004; de Lusignan & van Weel, 2006). Primary care data also has been used in novel ways, such as predicting and assessing inappropriate emergency room use (St-Maurice et al., 2013), making quality-of-life predictions (Pakhomov et al., 2008), identifying cardiovascular disease prevention strategies (Volker, Davey, Cochrane, Williams, & Clancy, 2014), and supporting operations governance within clinics (Pearce, de Lusignan, Phillips, Hall, & Travaglia, 2013). Primary care data has many theoretical and practical uses.

Primary care is often described as the first-contact health service for a patient who is sick by staff trained as primary care professionals (de Lusignan, 2003). This definition, however, is overly simplistic. A better definition suggests that primary care is the specialty that sees any patient with any combination of problems. Primary care is characterized by first-contact care, longitudinality, comprehensive services, and coordination (Starfield, 1997). Primary care is characterized by de Lusignan (2003) as having: heuristic decision making for patients with vague symptoms and unstructured problems; a more holistic, biopsychosocial model grounded by a long-term relationship between individuals, families, and their

primary care providers; and its own scientific body of knowledge whose application enhances practice delivered using a patient-centred consulting style. This approach is formalized with care models such as SDM.

Biomedical models leave little room within their framework for the social, psychological, and behavioural dimensions of illness. In primary care, biopsychosocial models are employed instead of biomedical models (de Lusignan, 2003). This distinction is relevant from a complexity perspective: using systems-thinking in primary care is much more complex, and modelling tasks of care go beyond organ systems and must incorporate a patient's environment, attitudes, and beliefs.

## **2.2 Governance Structures**

Healthcare in Canada is a provincial responsibility. The Canadian healthcare system as it functions today was established by the Canada Health Act (CHA) of 1984. The law requires that healthcare in each province be publicly administered, provide comprehensive health coverage, provide universal access, be portable within the country, and be equally accessible to all citizens. The law specifically prohibits extra-billing, and providers are required to operate entirely by payments from the provincial insurance system (Thompson, 2015).

Since the accession of the CHA into law, several accords have been negotiated between the federal government and the provinces. These accords have established additional funding, but they also have established priorities and reform goals for healthcare throughout the country. The accord negotiated in 2003 reserved specific funding for primary care reforms and the implementation of a national electronic medical record strategy.

Prior to the reforms of the early 2000s, primary care operated on a fee-for-service basis. In this model, physicians would operate independently in their own practices, and bill the government for each service provided to patients. A fee schedule was established to govern the list of services and the value of their compensation.

As reforms began, new capitation models were introduced to primary care. Doctors were able to form groups as Family Health Networks (FHNs), Family Health Groups (FHGs), or Family Health Organizations (FHOs). As groups, doctors were compensated for having patient rosters instead of primarily being paid on a fee-for-service basis. If a doctor has 1,200 patients in their care, they would be paid a fee for each patient regardless of whether or not these patients had been seen. The capitation model facilitates budgeting for the provincial government and enables cost-controls. Doctors, with capitation, are encouraged to roster more patients and identify strategies to ensure they are healthy, in contrast with a fee-for-service model in which doctors are funded when patients are sick. In Ontario's capitation model, a fractional fee-for-service component remains. Thus, the FHO, FHN, and FHG models are a hybrid of fee-

for-service and capitation. The differences between FHNs, FHGs, and FHOs relate to the establishment of after-hours clinics and slight differences between the schedules of billable services and fee-for-service components.

FHTs in Ontario are part of the same health reforms from the early 2000s (Kralj & Kantarevic, 2012). Ontario has approximately 250 FHTs that vary in size and scope (Grant, 2015). Some of the larger FHTs hire a diverse range of allied health professionals (AHPs<sup>1</sup>), including nurse practitioners (NPs), registered nurses (RNs), registered practical nurses (RPNs<sup>2</sup>), pharmacists, mental health workers, dietitians, social workers, kinesiologists, and health promoters (Grant, 2015). Family physicians must belong to FHNs or FHOs to start a FHT or be part of a FHT. Once physicians are members of FHTs, they are entitled to use the services of the FHT by bringing AHPs into their practice to support their patients.

The concept behind FHTs is to enable physicians to ‘roster’ (e.g. capitate) additional patients by freeing their time and putting complex patients under the care of AHPs. For example, instead of providing a mental health consult over the course of an hour, a physician could ask a mental health counsellor to provide that service and see as many as six patients instead. The investment in FHTs by the province is intended to make primary care more accessible, improve the management of chronic disease, and improve long-term health outcomes within the population.

### **2.3 Medical Records, Registries, and Billing**

Before the primary reforms of the early 2000s, physicians worked independently and were responsible for managing their own medical records. Few physicians were using electronic systems. After the accord from 2003, investments occurred to encourage EMR<sup>3</sup> adoption in primary care. This marked a significant culture shift: simultaneously, physicians were expected to operate within a capitation model, were encouraged to work in groups, and were expected to change from paper to electronic record systems. To further complicate the issue of medical records, the Government of Ontario also adopted the Personal

---

<sup>1</sup> AHPs is not a typical acronym/title used in the United States. PT, OT, RT, SLP, etc. are often used instead to capture different types of clinicians acting as AHPs.

<sup>2</sup> In other jurisdiction, RNs & RPNs may have different titles. In the United States, these providers would be called Licensed Practical Nurses (LPNs).

<sup>3</sup> The use of the acronym EMR versus EHR varies from country to country. In Canada and Europe, the term EMR is meant to describe a clinician centric record system in primary care. In the United States, the term EHR is typically used to describe the same type of information system.

Health Information Protection Act (PHIPA<sup>4</sup>) in 2004, which refined record custodianship laws, patient access rules, and record management requirements. Since the early 2000s, primary care medical records and the collection of data has been subjected to a significant number of transformative reforms.

### **2.3.1 Medical Records**

In primary care, individual physicians who are in ‘custody’ of medical records are considered health information custodians under PHIPA. Under this law, these physicians are responsible for managing the 10 fair information practices of the Canadian Standards Association (CSA), including patient consent, accountability, public notice, safeguarding information, and providing individual access. Primary care physicians are wholly responsible for their information management practices and are ultimately responsible for managing the content of their records and implementing minimum standards. Generally, the College of Physicians and Surgeons of Ontario (CPSO) defines minimum requirements with a medical record policy (College of Physicians and Surgeons of Ontario, 2012b). The record’s policy defines minimal legal requirements and does not mandate specific data points or record codification.

As owners of the medical record, physicians in primary care operate differently than physicians in the hospital environment; in hospitals, medical records belong to the hospital corporation, and the hospital is responsible for implementing the CSA’s privacy principles. Hospitals hire Health Information Management Professionals (HIMPs<sup>5</sup>) who can ensure that records have standardized content, manage the data codification process for each patient visit, and analyse data and trends. Clearly, primary care physicians are challenged to effectively manage records themselves without reducing the time they typically would spend with patients. These resource and time constraints eventually created a need to streamline processes and access to information through EMRs and to acquire information management supports.

As part of the primary care reforms of the early 2000s, the Ministry of Health and Long Term Care (MOHLTC) in Ontario funded an EMR adoption program through OntarioMD. Since its implementation, 11,650 physicians received over \$25,000 in funding over three years to purchase and adopt EMRs (OntarioMD, 2015). As a result, a majority of primary care physicians in Ontario have an EMR system. Until 2015, OntarioMD’s original program had several extensions and renewed funding for

---

<sup>4</sup> For the purposes of understanding the context of privacy within the lens of this dissertation, PHIPA can be considered substantially similar to the United States’ Health Insurance Portability and Accountability Act (HIPAA) of 1996.

<sup>5</sup> Different acronyms can be used to describe this role. In some areas, such as the United States, the acronyms of HIM or MIM (Master of Information Management) is used. In the context of my dissertation, any of these acronyms capture the same intent and professional role within a hospital.

physicians. More recently, OntarioMD has launched an EMR Practice Enhancement Program (EPEP). EPEP offers “an analysis of EMR use, practice workflow and data quality to help physicians move beyond data capture and use information to improve patient care and practice efficiency” (OntarioMD, n.d.). EPEP is based on an EMR Maturity Model and is an evidence-based framework for measuring, understanding, and enhancing EMR use. It is designed to measure the effective use and value of an EMR by encouraging a series of strategies and behaviours.

FHOs, FHNs, and FHTs changed the governance structure in primary care. Whereas providing support to 14,000+ individual physicians throughout the province would have been exceptionally impractical and expensive, providing information management support to primary care clinics through the administrative capacity of 250 FHTs was relatively reasonable. After advocacy and lobbying by the Association of Family Health Teams of Ontario (AFHTO), FHTs and primary care practices were given access to a new type of resource, Quality Improvement Decision Support Specialists (QIDSS<sup>6</sup>).

The role of QIDSS within a FHT is intended to support and coach primary care clinics toward the effective use of EMRs. This includes building the capacity within primary care to collect and codify data, analyse trends, identify priorities from data, and generate care quality reports. The data and reports are aimed to drive improvements in primary care. QIDSS also would have the capacity to support physician participation in registry initiatives by preparing data for submission.

### **2.3.2 Billing and Reporting**

Medical records enable several uses of primary care data.

#### **2.3.2.1 Billing**

Per the CHA of 1984, physicians providing healthcare services in Canada may only be paid by the provincial healthcare system. Physicians are compensated for providing medically necessary assessments and procedures. In Ontario, physicians bill the Ontario Health Insurance Plan (OHIP) and are provided with payment for their services. Ethically and legally, bills provided to the government must be supported by clinical documentation within the medical record.

#### **2.3.2.2 Healthcare Quality Improvement**

Primary care data has been recognized as a source of information that could be used to identify challenges and solutions in the broader healthcare system. Several organizations are interested in

---

<sup>6</sup> The QIDSS role is fairly unique to Ontario and its implementation of team-based care through the FHT strategy. It does not necessarily have a corollary in other jurisdictions, such as the United States.

collecting data from primary care, including Local Health Integration Networks (LHINs), Health Quality Ontario (HQP), and the Canadian Institute for Health Information (CIHI).

Much data is contained within medical records but is not accessible or usable for secondary purposes without addressing data quality issues. Recognizing the potential value of primary care data, HQO and LHINs work with QIDSS to identify priorities, adopt new processes to collect new data points, and report useful information that helps measure health care quality improvements.

### 2.3.2.3 Family Health Team Activities

The Province of Ontario provides funding for FHTs to hire staff that provide services to patients within physician offices. As a condition of funding, FHTs must report their activities to the MOHLTC. This information is available within medical records, but extracting the information in a format that aligns with the MOHLTC requirements can be challenging. For example, some EMRs may not have the ability to categorize initial encounters versus follow ups for individual clinicians, and attempting to automatically export this information is not practical. For larger FHTs, manually capturing this information also is impractical. Thus, some FHTs have developed their own registries to facilitate reporting to the MOHLTC. These registries generally are accessed by the clinicians providing services to patients and would capture tabular information in addition to the information within medical records. The secondary entry adopts a forms-based paradigm to data collection.

FHT registries are interesting sources of data and can be used for secondary purposes by FHTs. Instead of merely using the registry to report statistics to the MOHLTC, the system can be used to establish performance benchmarks, identify patient populations, track employee time, and manage vacation bookings. Larger FHTs can span over 20 physical locations, have greater organizational complexity, and are more likely to use a registry approach to meet their MOHLTC reporting requirements.

## 2.4 Data Quality<sup>7</sup>

When using primary care data, it is important to understand the differences between a data element's true meaning and its representation within a record (Wand & Wang, 1996). These differences are important, and can be described through the study of data quality. The issue of data quality in primary care is a field of study in of itself, and has been studied extensively (e.g. Brouwer et al. 2006; Brown et al. 2012; Collins and Janssens 2012; de Lusignan 2005; de Lusignan et al. 2006; de Lusignan et al. 2010;

---

<sup>7</sup> Section 2.4 and the first paragraph of section 2.1 were part of a manuscript accepted for publication in the *Journal of Data and Information Quality* (St-Maurice & Burns, in press). To provide better structure and readability to the dissertation, that content is in this chapter, while the remainder of the publication is shown in chapter 7.

Hogan and Wagner 1997; Thiru et al. 2003; Majeed et al. 2008; de Lusignan et al. 2003; Stone et al. 2010; St-Maurice 2011; Pearce et al. 2013).

There are many ways to describe data quality. Data quality dimensions are usually contextual to the user and their ecosystem. Dimensions may include concepts such as accuracy, consistency, reliability, timeliness, relevance, completeness, currency, consistency, flexibility, precision, format, interpretability and usefulness (Wand & Wang, 1996; Wang, Strong, & Guarascio, 1996). There are many studies that have discussed and studied specific dimensions in depth: accuracy (Arts, 2002; Blake & Mangiameli, 2011; Fisher, Lauria, & Matheus, 2009), completeness (Arts, 2002; Bettencourt-Silva et al., 2015; Blake & Mangiameli, 2011; Larsen et al., 2009), consistency (Blake & Mangiameli, 2011), validity (Bray & Parkin, 2009; Larsen et al., 2009), and timeliness (Bettencourt-Silva et al., 2015; Blake & Mangiameli, 2011; Bray & Parkin, 2009; Larsen et al., 2009) have been studied and measured in various ways. Some studies have also shown interactions between different data quality dimensions (Blake & Mangiameli, 2011; Lee, Strong, Kahn, & Wang, 2002).

In primary care, there are different types of data. Unstructured data fields allow clinicians to document complex biopsychosocial concepts in free text and do not force clinicians to enter data in a specific way. An important drawback to free text is that it introduces curious acronyms and abbreviations which are difficult to interpret with even advanced natural language processing techniques (St-Maurice et al., 2013). As well, there is a dichotomy between what clinicians prefer and what managers require: clinicians feel data codification is not sufficiently expressive and managers dislike the use of free text because it is hard to analyse (de Lusignan et al., 2003).

To make strategic health planning decisions and analyse population data, analysts need data coded into a common structured format. The benefits of data codification, through forms or other computerized mechanisms, is important as it can simplify the submission of data to medical registries for secondary analysis. However, merely providing structure to data does not preempt data quality issues, as data in registries is known to suffer from data quality deficiencies (Arts, 2002).

#### **2.4.1 Data Quality Dimensions in Primary Care**

Primary care data is highly contextual and many data quality dimensions are relevant within its domain (St-Maurice & Burns, 2014). *Accuracy*, *validity*, *completeness*, and *timeliness* are of particular interest when using primary care data for secondary purposes, and are often assessed together to pragmatically assess overall data quality (Bray & Parkin, 2009). *Data use*, *usability* and *usefulness* are also helpful attributes for characterizing more contextual aspects of data quality.

#### 2.4.1.1 Accuracy, Correctness, and Validity

While accuracy is a relatively common term to use in the study of data quality, there are no exact definitions (Wand & Wang, 1996). Generally, labelling data as inaccurate implies that the real-world and its digital representation are misaligned. Thus, inaccuracy can be interpreted as a result of mapping data incorrectly from the real-world state into its information representation (Wand & Wang, 1996). The term ‘accuracy’ can be vague and should be avoided unless precisely defined (Thiru et al., 2003).

In the literature, the terms validity, correctness and accuracy are sometimes used interchangeably or as synonyms (Bray & Parkin, 2009; Larsen et al., 2009; Liaw et al., 2013). When describing the data quality of healthcare registry data, validity is often used to describe soundness and integrity of data, or as a measure of faults (Larsen et al., 2009). Validity can also be defined as the proportion of cases in a dataset which truly have the attribute value (Bray & Parkin, 2009). In other areas of the literature, the term ‘correctness’ can be used to describe the same general idea and can be measured through the proportion of correct data (Todoran, Lecornu, Khenchaf, & Caillec, 2015).

In some ontologies, correctness includes the concepts of accuracy and completeness and there is a hierarchical relationship between these concepts (Liaw et al., 2013). Some frameworks include measurement errors and validity checks as aspects of an ‘accuracy’ dimension (Canadian Institute for Health Information, n.d.). Generally, the concepts of accuracy, correctness, validity, soundness, integrity and consistency have some nuanced differences, and at times relationships, depending on the area of application, but capture similar facets of a common data quality paradigm.

In primary care, accuracy refers to the extent to which data conforms to the correct value or truth. Validity refers to whether or not there is integrity and logic to the dataset. In practice, accuracy and validity are important concepts: ‘easy’ tasks such as identifying a list of patients with a common disease, like diabetes, is challenging in primary care. In a normal practice, as many as 80 diabetes patients may be misclassified, miscoded or misdiagnosed (de Lusignan et al., 2010; Stone et al., 2010). Aiming to ensure data is both valid (logical) and accurate is a priority in primary care. From the perspective of secondary use of data, such as an epidemiological study within a region, inaccurate data or invalid data could significantly impact results and change the allocation of resources. Without accurate high-quality data, it is not possible to create a centralized database to support policy or epidemiological research (Collins & Janssens, 2012).

#### 2.4.1.2 Completeness

Data can be considered complete if all necessary values are included (Wand & Wang, 1996). Completeness is the ability of an information system to represent every meaningful state of the real-world



system. In primary care, completeness refers to the quantity of data elements that have been used to describe the breadth and depth of the clinical encounter.

In primary care completeness is a balance between clinical efficiency and documentation necessity. Brevity can be sufficient for clinical care, but does not usually enable effective secondary use (St-Maurice & Burns, 2014).

In primary care, there are often uses of abbreviations and acronyms (St-Maurice et al., 2013) and clinicians choose not to categorize their entries with classification systems, such as ICD-9. In other situations, clinicians leave extremely brief notes in their records that are only useful for their own use (St-Maurice & Burns, 2014). Leaving records incomplete (e.g. without searchable terms or with too much brevity) prevents important information from being found and used for secondary purposes. To enable effective secondary uses, records need to be complete as possible.

From a design perspective, completeness can be managed by forcing users to complete mandatory fields. This solution can come at a cost of accuracy. This is exemplified by a curious majority of Albanian nationals within a clinical dataset in the United States; when users were forced to select a nationality, the majority opted to pick the first item in the drop down list that was sorted alphabetically (Hirsch, 2012). Hence, there are tradeoffs between designing for accuracy and completeness and these data quality challenges are closely related.

#### 2.4.1.3 Timeliness

Timeliness has been defined in terms of whether the data is out of date and available for use on time (Wand & Wang, 1996). Timeliness is affected by three factors: how fast the information system state is updated after the real-world system changes (e.g. currency and speed of input), the rate of change of the real-world system (volatility), and the time the data is actually used (Wand & Wang, 1996).

In primary care, timeliness is affected by currency issues. Some data is only valid near the time it was recorded. For example, the ‘ex-smoker’ status that can change very rapidly, whereas other data, such as a diagnosis for Type 1 diabetes, is usually valid for a patient’s entire life (de Lusignan et al., 2010). Since the currency of data is very relevant to its interpretation, understanding the delay in data entry is important. The measure of the delay between the clinical encounter and the recording of data, which is hopefully minimal, is an important aspect of the timeliness dimensions and can influence the extent to which data is complete and accurate (Bray & Parkin, 2009).

An interesting secondary use of primary care data would be outbreak detection for public health agencies. This type of use would be significantly impacted by a delay in data entry. Understanding and measuring entry delay is an important aspect of data quality.

#### 2.4.1.4 *Usefulness, Usability, and Use*

Data usefulness is a highly contextual dimension of data quality as it depends on the data user (Haug & Stentoft Arlbjørn, 2011). At the same time, usefulness and usability of data are important aspects of quality, as high-quality data must be fit for the use by data consumers or its collection is irrelevant (Strong et al., 1997).

In primary care, data must serve a purpose. If the purpose and the usefulness of data is unclear for users, other data quality dimensions may be impacted. For example, if users do not understand the purpose of their data, or do not engage in its use, data can be arbitrarily inputted into the database. This is exemplified by our previous example of high Albanian demographics (Hirsch, 2012); users were not concerned with the use of their data and did not use the demographic data and therefore recorded arbitrary data. If users required specific demographic data, such as nationality, to provide care or generate reports, they would likely enter data more conscientiously. In this sense, usefulness, usability and use of data are very important dimensions to consider in the primary care environment. There is a balance between encouraging use, enforcing completeness and ensuring accuracy.

#### **2.4.2 Previous Work**

There have been many studies aimed at understanding data quality in primary care. Previous studies have drawn attention to the impact of data quality on patient care and have identified the challenges associated with incorrect or incomplete coding and classification. There are many recent examples: in diabetes, the prevalence of miscoding, misclassification and misdiagnosis impacts 8% of records (de Lusignan et al., 2010; Stone et al., 2010); in cancer patients, 30% to 40% of primary care records lack an identifiable cancer diagnosis (Sollie, Roskam, Sijmons, Numans, & Helsper, 2016; Sollie, Sijmons, Helsper, & Numans, 2017); with thyroid disorders, many diagnoses within records were undocumented, imprecise or unspecific (Münch et al., 2016). Other work has drawn attention to the impact of poor quality on secondary uses, such as clinical governance activities (Pearce et al., 2013). Some studies have taken novel approaches to improving data quality in primary care, such as deploying a data quality feedback tool (van der Bij et al., 2016). Generally, there remain significant opportunities for improving data quality in primary care.

Barriers to data quality in primary care include user skill gaps, task time, and professional and organizational priorities (de Lusignan, 2005). Unlike hospital environments, the crux of the problem with data quality in primary care is that clinicians are asked to code data. As clinicians have a primary interest in providing high-quality patient care, the creation of high-quality data for secondary purposes is not a prevalent concern; there is a lack of understanding about the value of data for patient management and service planning (Collins & Janssens, 2012). One important conclusion of previous studies is that the

completeness and accuracy of data entry relies mainly on the enthusiasm of clinicians (Majeed et al., 2008). In many cases, clinicians do not fully accept the merit of maintaining data, which has a direct consequence for data quality.

Arguably, data quality in primary care is a complex socio-technical issue that is impacted by workflows, technology, culture and context. As such, there has historically been a lack of empirical knowledge regarding how to approach the improvement of data quality in primary care (Brouwer et al., 2006).

As a first step to understanding and solving data quality challenges in primary care, there is a need to characterize users, their priorities and their workflows relating to data. In practice, this could be accomplished by developing models that define and correlate data quality dimensions for primary care users. For example, Blake and Mangiameli (2011) studied the interactions between accuracy, completeness, timeliness and consistency. A similar approach to understanding data quality in primary care would be helpful. Understanding multi-dimensional data quality interactions could help understand and enhance interventions aimed at improving data quality.

When using primary care data for secondary purposes (e.g. not for direct patient care), it is important to understand the differences between its true meaning and its representation within a dataset (Wand & Wang, 1996). These differences are important, and can be described through the study of data quality. The issue of data quality in primary care has been studied extensively (Brouwer et al., 2006; Brown et al., 2012; Collins & Janssens, 2012; de Lusignan, 2005; de Lusignan et al., 2010, 2006, 2003; Hogan & Wagner, 1997; Majeed et al., 2008; Pearce et al., 2013; St-Maurice, 2011; Stone et al., 2010; Thiru et al., 2003).

There are many ways to describe data quality. Data quality dimensions are usually contextual to the user and their ecosystem. Dimensions may include concepts such as accuracy, consistency, reliability, timeliness, relevance, completeness, currency, consistency, flexibility, precision, format, interpretability and usefulness (Wand & Wang, 1996; Wang et al., 1996). There are many studies that have discussed and studied specific dimensions in depth: accuracy (Arts, 2002; Blake & Mangiameli, 2011; Fisher et al., 2009), completeness (Arts, 2002; Bettencourt-Silva et al., 2015; Blake & Mangiameli, 2011; Larsen et al., 2009), consistency (Blake & Mangiameli, 2011), validity (Bray & Parkin, 2009; Larsen et al., 2009), and timeliness (Bettencourt-Silva et al., 2015; Blake & Mangiameli, 2011; Bray & Parkin, 2009; Larsen et al., 2009) have been studied in various ways. As well, some studies have shown interactions between different dimensions, such as a relationship between data complexity and data quality (Blake & Mangiameli, 2011; Lee et al., 2002).

In primary care, there are different types of data. Unstructured data fields allow clinicians to document the nuances of complex biopsychosocial treatment and does not force clinicians to enter data

into specific data fields. While this may be the preferred way for clinicians to practice medicine, it introduces opportunities for curious acronyms and abbreviations which are difficult to interpret with even advanced natural language processing techniques (St-Maurice et al., 2013). To make strategic health planning decisions and analyse population data, analysts and managers need data coded into a common format, which is not viable with unstructured data. The dichotomy between what clinicians prefer and what managers require can be characterized as cross-cultural, as clinicians feel data codification is not sufficiently expressive and managers dislike the use of unstructured data (de Lusignan et al., 2003). The benefits of data codification, through forms or other computerized mechanisms, is important as it can feed medical registries. Structured data in registries, however, also suffers from data quality problems (Arts, 2002).

## **2.5 A Complex Socio-Technical System**

Healthcare is considered a complex socio-technical system (Jiancaro et al., 2013). Primary care is no exception (Gagnon et al., 2010). Furthermore, primary care in Ontario has been undergoing major changes. These changes have included developing physician groups, implementing EMRs, and providing clinics with clinical and administrative reports. Technical challenges in primary care are rooted in history, structures, culture, and changes. Characterizing any challenges in primary care as purely technical would be a significant oversight, and the sector should be understood and modelled as a complex socio-technical system.

Clearly, data quality in primary care is a complex socio-technical issue. Obvious problems with data exist, and these problems affect patient care and establish a need to create effective interventions as well as to characterize users to better understand their priorities and workflows regarding data.

## Chapter 3

# Cognitive Work Analysis

I used CWA to model my domains. In this chapter, I provide readers with some background regarding CWA to help them understand my work in chapters 4 to 6.

### 3.1 Overview

The field of human factors and environmental ergonomics has a variety of tools to help understand complex system ecologies, develop strategies to support users, and incorporate technology into workflows (Saleem et al., 2009). Various approaches have been used to develop models describing patient health and clinical decision making (Ashoori, Burns, D'Entremont, & Momtahan, 2014; Burns, Enomoto, & Momtahan, 2009; Carayon et al., 2014; Chow & Vicente, 2002; Chu, Hsu, Cardenas, & Taira, 1998; Hajdukiewicz, Doyle, Milgram, Vicente, & Burns, 1998; Hajdukiewicz, Vicente, Doyle, Milgram, & Burns, 2001; Nemeth, 2006; Rezai & Burns, 2014; Watson & Sanderson, 1998). Specifically, CWA and Work Domain Analysis (WDA) (Vicente, 1999) can help describe user decision making and competency requirements within complex socio-technical systems. A detailed CWA can assist in designing more-effective systems through the use of EID.

CWA is a conceptual framework that facilitates the analysis of the environment at various levels of detail and assesses how the environment impacts and shapes the human-information interaction (Vicente, 1999). CWA is a systematic framework that can be used to examine the work activities of participants in workflows and processes using environmental, organizational, and social lenses (Vicente, 1999). CWA provides concepts and templates that aid in the analysis of complex phenomena and supports better design in complex socio-technical systems (Fidel & Pejtersen, 2004).

CWA is a formative framework and describes how the system *could* behave, rather than how the system *should* or *does* perform (McIlroy & Stanton, 2015). The approach is work-centred, rather than user-centred, and it analyses the constraints and goals that shape behaviour in work environments (Jiancaro et al., 2013). CWA is broken down into five stages of analysis: WDA, Control Task Analysis (ConTA), strategies analysis (StrA), social organization and co-operation analysis (SOCA), and worker competencies analysis (WCA) (Vicente, 1999).

Although the CWA framework includes five phases, not all stages are necessarily required in each analysis; the CWA method serves as a toolkit for practitioners, and the analysis phases chosen should match the inherent system constraints (McIlroy & Stanton, 2015).

### 3.1.1 Work Domain Analysis

As the first phase of CWA, WDA describes the system's work domain in its entirety to display the functional means-ends relationships between system functions and components. Through an AH, the work domain is analysed and broken down into various levels of abstraction. The Abstraction Decomposition Space (ADS) can be used in conjunction with the AH to break down the analysis into different levels of resolution (Vicente, 1999). Each type of model details investigations of the work domain and highlights underlying functional structures that must be respected for a system to achieve its intended focus; it is the documentation of the functional ecology (Kilgore, St-Cyr, & Jamieson, 2009). WDA allows a practitioner to understand what needs to be done by the system, and helps understand *why* the system exists.

WDA can describe how structures, abstract values, and constraints affect the normal functions of a system (Bisantz & Mazaeva, 2008). CWA has been used previously in healthcare (Ashoori et al., 2014; Burns et al., 2009; J. A. Effken, Loeb, Kang, & Lin, 2008; Jiancaro et al., 2013; Lim, Anderson, & Buckle, 2015; Momtahan & Burns, 2004; Rezai & Burns, 2014; Watson & Sanderson, 2007).

The AH is a modelling tool that describes the results of a WDA (Vicente, 1999). The AH is intended to be a full depiction of the necessary constraints that must be considered for the system to achieve its purpose, while describing the system's underlining ecology and limitations (Kilgore et al., 2009). Using AHs can help bridge the psychology-culture-medicine gap in healthcare. These hierarchies can be used to develop representations of patient care that are aligned with biomedical knowledge, support medical problem solving, and act as a frame of reference (Hajdukiewicz et al., 2001). As a structured approach to WDA, the AH includes a layer to describe the system's functional purposes, abstract functions, generalized functions, physical functions, and physical forms. Lines are shown between each layer to show means-end or how-why relationships (Kilgore et al., 2009).

### 3.1.2 Control Task Analysis

ConTA is the second phase of the CWA framework and is a formative phase that describes how tasks can be undertaken. ConTA allows a practitioner to understand *what* must be executed and/or processed by the system and controller and describes these goals regardless of who is supposed to conduct the activity, or how. Because CWA is a formative framework, the ConTA highlights when activities *can* be carried out, and how they are *likely* to be carried out (McIlroy & Stanton, 2015).

ConTA uses Rasmussen's DLs (Rasmussen, 1974). DLs capture the flow of information processing associated with individual control tasks, and enable the use of shortcuts to describe expert and novice decision-making patterns (Vicente, 1999). DLs are represented by a series of data-processing (rectangles) and knowledge states (circles) that represent information-processing patterns. Consecutive states of knowledge are separated by mental activities to transform one state of knowledge to the next. ConTA requires information about the state of the work domain, and outputs the required actions on the work domain (Vicente, 1999).

The relationship between work domain and task analysis can be conceived as a partnership. At one end, a WDA identifies information requirements that are event- and time-independent, providing a basis for supporting worker adaptation to novelty and change. On the other end, task analysis identifies information requirements that are event- and time-dependent, thus providing an efficient basis for supporting worker performance to anticipated situations. In other words, WDA is ecological and helps understand which tasks may be performed by the user, whereas task analysis describes those actions (Hajdukiewicz and Vicente 2004).

ConTA allows a practitioner to take an expert user into consideration, and allows a human actor to engage in a constructive learning process (Vicente, 1999). ConTA anticipates user learning and adaptation as a constructive learning process. In this context, experts do not retrieve a pre-planned solution from memory when accomplishing a task, but actively generate contextually tailored sequences of cognitive activities that are appropriate in a particular situation (Vicente, 1999). ConTA aims to model not only novice cognitive tasks, but also aims to map these expert sequences.

Theoretically, all users gain experience and eventually become experts. Instead of trying to determine what would be required for a novice to accomplish a task, we can instead try to determine the preconditions that have to be satisfied for expert performance. A cornerstone of ConTA is to model processing behaviour toward designing computer-based information systems that deliberately induce and support expert action, which thereby leads to gains in cognitive efficiency (Vicente, 1999). DLs are tools designed to achieve the objective of discovering and documenting expert cognitive sequences. Through the use of shortcuts, a practitioner can articulate novice and expert paths; in this way, the DL serves as a tool to illuminate ways to induce expert action.

Expertise in a DL can be expressed through the creation of 'shunting' effects that bypass higher, abstract levels of thinking. As described by Rasmussen (1974), a trained operator will only occasionally have to move through all the steps of the basic sequence. Shunting results from subconscious or intuitive data processing; "mental activity may take place at a symbolic level, resulting directly in a knowledge state later in the sequence", or "an association based upon previous experience may result in a leap directly from one state of knowledge to another one in the sequence". These shunts evolve in the

constructive learning process associated with training and are represented by arrows from data processing activities (e.g. boxes) to more abstract states of knowledge (e.g. circles).

In addition to shunts, DLs allow for the description of ‘leaps’, which represent direct transitions from two separate knowledge states (e.g. going from a circle to a circle). Leaps represent intuitive data processing performed by the subconscious and describe associations that bypass more-complex, higher-level reasoning. Leaps rely on the large capacity of holistic perception and recognition, as well as the large data processing capacity of the subconscious functions of generalization and modelling (Rasmussen, 1974).

### **3.1.3 Strategies Analysis**

The third phase of CWA, which is StrA, builds on the control tasks from ConTA by describing the alternate ways each control task may be performed. StrA describes the variety of approaches and ways that tasks can be completed. StrA allows a practitioner to understand *how* the system executes its functions.

The results of StrA are generally represented through the use of IFMs (Kilgore et al., 2009). IFMs are modelling tools from the CWA framework that can be used to describe how a task can be accomplished (Vicente, 1999). IFMs should be idealized categories of task procedures, and they should be context-specific. They are represented by processing tasks (squares) and mental states (circles). When IFMs are used for SOCA, the circles and squares are coloured to represent the team members performing each processing activity, and denote who has a particular mental state.

### **3.1.4 Social Organization and Co-operation Analyses**

SOCA, the fourth phase of CWA, builds on the analysis performed in StrA by determining the actors performing tasks. The purpose of this analysis is to address and document organizational constraints imposed by job roles and definitions (Vicente, 1999). As an example, one way of approaching this phase is to adapt the IFMs developed in StrA by colour coding to indicate either technical or human actors involved in the task completion. SOCA allows a practitioner to understand *who* completes the system functions (Jiancaro et al., 2013). Another way of looking at this phase is through Team CWA (Ashoori et al., 2014).

### **3.1.5 Worker Competencies Analysis**

The final stage of CWA, WCA, uses Rasmussen’s (1983) SRK taxonomy and describes the cognitive requirements needed to fulfil the system’s functions (McIlroy & Stanton, 2015). The analysis describes the system actor and their required skill-based behaviour, rule-based behaviour, and knowledge-



based behaviour. WCA allows a practitioner to understand the means and conditions (*'by what means'*) that are required to execute the system functions.

### **3.2 Ecological Interface Design**

EID is a design framework that has been used in a large number of socio-technical systems (McIlroy & Stanton, 2015) and typically is paired with CWA. For example, a model of anaesthesia (Watson & Sanderson, 1998) was used to support the development of auditory alarm systems (Watson & Sanderson, 2007), and a model of the cardiovascular system (Hajdukiewicz et al., 1998) was employed in the design of clinical displays (J. Effken, Loeb, Johnson, Johnson, & Reyna, 2001). EID is constructed on two stages of the CWA framework: the AH and the SRK taxonomy. EID uses WDA to identify system variables for each level of the AH model, and the SRK taxonomy helps ensure that user needs and competencies are respected (Burns & Hajdukiewicz, 2004). The goal of EID is to design an interface that does not force mental processing at higher levels than the demands of a task require. EID supports each level of the three levels of cognitive control in the SRK and aims to design interfaces that do not contribute to the difficulty of the task, while supporting the entire range of activities that users might face (Vicente & Rasmussen, 1992). EID is about shifting and changing cognitive control. CWA uses the AH in its first phase, and uses the SRK in its final phase. Thus, EID and CWA are said to be intimately linked (McIlroy & Stanton, 2015); EID is the natural design implementation of analyses conducted with CWA. Conceptually, the AHs and WCAs that are developed in later chapters can be used to support design through the use of EID.

In its original conceptualization and description, EID only used the WDA and WCA phases of the CWA analysis. The phases of ConTA, StrA, and SOCA do not address ecology, per se, and do not fit the original intent of the framework. However, ConTA, StrA, and SOCA models are important for a full CWA-based design approach.

### **3.3 Applications of Cognitive Work Analysis**

CWA has proven to be a versatile and broadly used framework. It has been used for a variety of purposes, including designing automation (e.g. Mazaeva and Bisantz 2007), modelling and designing systems (e.g. Hajdukiewicz 1998; Bisantz et al. 2003; St-Maurice and Burns 2015), assessing training requirements (e.g. Naikar, Sanderson, and Lintern 1999), analysing information requirements (e.g. (Ahlstrom, 2005), supporting procurement (e.g. Lintern and Naikar 2000; Naikar and Sanderson 2001), and designing effective training for teams (e.g. Naikar et al. 2003; Naikar and Saunders 2003; Naikar, Sanderson, and Lintern 1999). The domains of CWA have included air traffic control and aviation (e.g. Ahlstrom, 2005; Beevis, Vicente, & Dinadis, 1998; Naikar et al., 2003), power generation (e.g. Sanderson

et al. 2005; Lehane, Toleman, and Benecke 2000), naval (e.g. Bisantz et al. 2003; Burns, Bisantz, and Roth 2004), manufacturing (e.g. Higgins 1998; Upton and Doherty 2008), military command and control (e.g. Jenkins et al. 2008; Lintern, Miller, and Baker 2002), petrochemical (e.g. Jamieson and Vicente 2001), and transportation (e.g. (Birrell, Young, Jenkins, & Stanton, 2012; Jansson, Olsson, & Erlandsson, 2006; Salmon, Regan, Lenné, Stanton, & Young, 2007). CWA has also been broadly used in healthcare (e.g. Jiancaro, Jamieson, and Mihailidis 2013; Effken et al. 2008; Ashoori et al. 2014; Rezai and Burns 2014), including the analysis of cardiac care teletriage (Burns et al., 2009), trauma centre redesign (Sarcevic, Lesk, Marsic, & Burd, 2010), and nurse decision support (Momtahan & Burns, 2004).

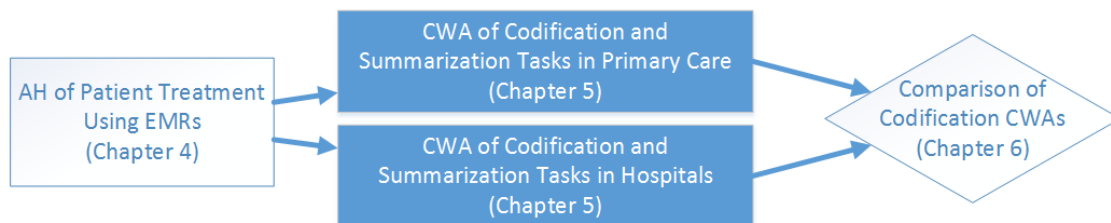
### 3.4 Use of Cognitive Work Analysis to Study Data Quality

CWA is a helpful framework for analysing complex socio-technical systems. I used this framework to analyse data codification in primary care. My approach to using CWA started by creating an AH to represent patient treatment in healthcare. This work is presented in chapter 4. The need for developing this AH arose during my work in developing models for codification. It became important to understand where data was coming from and how it was used in clinical practice before trying to understand the information management aspects of data codification.

As a next step, I developed CWA models in primary care and hospital settings. I wanted to understand what was taking place in both environments to better understand the depth and breadth of the cognitive and ecological aspects of coding. I focused on a data codification task in each domain for my control tasks. The results of this work is presented in chapter 5.

After building CWA models for each environment, I wanted to draw a comparison. However, attempting to compare CWA models is uncommon. Thus, I developed an approach to compare CWA models and used this approach to compare and contrast the systems I studied. I gathered several useful insights from the comparison. This work is presented in chapter 6. This work includes several contributions to CWA. To my knowledge, this is the first CWA of data coding and records management. Second, I am making a contribution by proposing and demonstrating a method for conducting a detailed systematic look at comparing models to extract new insights. A summary of my work with CWA regarding data quality is shown in Figure 2.

Figure 2. Summary of Cognitive Work Analysis applications and uses.



## **Part B**

### **Systems Analysis and Modelling**

This part of my dissertation features three chapters focused on the development and interpretation of models to improve data quality during data codification. I present an AH of patient treatment in chapter 4, and I develop several models of data codification with CWA in chapter 5. Then, I compare the results from the two analyses in chapter 6.

The chapters in Part B were turned into two publications. A manuscript titled ‘Using CWA to Compare Complex System Domains’ was submitted to *Theoretical Issues in Ergonomic Science* and is under review. A second manuscript titled ‘Modelling Patient Treatment with Medical Records: An Abstraction Hierarchy to Understand User Competencies and Needs’ was accepted in the *Journal of Medical Internet Research (JMIR) Human Factors*. Both manuscripts incorporate some of the modelling work described in this section.



## **Chapter 4**

### **An Abstraction Hierarchy of Patient Treatment**

In the previous chapters, I provided an introduction and context for my work. In this chapter, I begin modelling the coding environment by building a high-level understanding of the work domain. I also perform a WDA of patient treatment.

#### **4.1 An Abstraction Hierarchy of Patient Care**

Healthcare is considered a complex socio-technical system (Jiancaro et al., 2013). As well, a movement is underway in patient care to move away from paternalistic healthcare approaches (Charles, Gafni, & Whelan, 1999; Weston, 2001) and engage patients in their own care. For example, a trend is to adopt SDM (Elwyn et al., 2012; Legare & Witteman, 2013) to improve patient care through engagement. Similarly, new healthcare laws are promoting patient centred care as a priority paradigm shift (e.g. the Patient's First legislation in Ontario). As the healthcare delivery environment incorporates new constraints and develops new goals, clinicians have unique needs and require a rich set of competencies to practice medicine. As a complex socio-technical system, using the CWA framework can be an effective approach to understand and describe the complexities of care in this challenging world.

As an output of WDA, many AHs have been developed to describe patient health. Some of these AHs were developed through a WDA, whereas others were developed within the context of a fuller CWA exercise. Some of these abstractions treat patients as biomedical machines with physiological processes (Hajdukiewicz et al., 1998, 2001; Hall, Rudolph, & Cao, 2006; Miller, 2004; Watson & Sanderson, 1998). For example, some models represent the human body in its resting state during anaesthesia (Watson & Sanderson, 1998), decompose the human body into systems and organs (Burns et al., 2009), or describe the cardiovascular system as an independent system (Hajdukiewicz et al., 1998). The scope of these analyses are more biomedical in nature because they describe treatments and procedures, and are modelling medical treatment without including the consultation phase. Within the confines of emergency or surgical care to unconscious patients, patient values and wishes fall out of the scope of analysis, and the model can be scoped to a largely physical model of repairing damage or improving physiological

processes. In these contexts, ‘aberrations in physiological and biological regulatory processes’ are the ‘domain upon which clinicians work’ (Miller, 2004).

In other cases, patients are conscious and therefore capable participants in their own healthcare. Ashoori and Burns (Ashoori et al., 2014) modelled the patient-as-an-actor approach effectively during a study of a birthing unit. The CWA showed rich coordinative points, shared artefacts and adjusting structures, and described the patient as an active partner that engaged in her own health. In particular, the AH modelled the patient as a physical function of prescription, assessment, and consulting. In another example, Rezai and Burns (Rezai & Burns, 2014) modelled patient values, skills, support systems, and abilities in a home healthcare scenario with WDA and ConTA. The scope did not include the patient within clinical practice. Regardless, both examples demonstrate that CWA is capable of characterizing patients as emotionally complex, social creatures and that CWA can successfully describe patients as decision makers with rich sets of values and capabilities to support their own healthcare. Within CWA, WDA can describe many complex relationships that are both biomedical and patient-related.

Building a model of patient treatment is challenging, and is further complicated by the nuances of effectively treating patients with the assistance of EMRs. At present, no CWA models or AHs of patient treatment address this context and need.

#### **4.1.1 Model Objective and Scope**

The objectives of the AH is to capture the complexities, balances, and challenges regarding patient treatment from a clinician’s perspective. Such a model could be specific to an individual physician, practice, or speciality. In an effort to offer a breadth of utility, the goal of this AH is to capture generic and common healthcare processes and priorities, without worrying about specific or unusual use cases. The goal is to develop a model that could represent all types of clinicians involved in providing and triaging care, including, but not limited to, physicians, physiotherapists, nurse practitioners, dietitians, mental health workers, and pharmacists.

The model must go beyond the laymen’s and paternalistic impression of medicine as a purely biomedical process. Whereas physicians are experts in disease, patients are experts in their own experience of disease and in their preferences (Weston, 2001). One of the challenges of patient care is incorporating the patient’s values and preferences into decision making (Legare & Witteman, 2013). The abstraction must articulate the challenges of treating patients in a world of anti-vaccine campaigns, conflicting personal values, and complex determinants of health. In this sense, the model needs to describe the biopsychosocial constraints and nuances of patient treatment.

Finally, the model must capture the impact and role of EMRs in delivering healthcare. The model needs to describe the complex processes associated with using EMR records and how they interact with clinical practice.

#### **4.1.2 Intended Uses**

As an insightful model of patient treatment, the AH should serve several purposes. The goal of developing this AH was to support many use cases, including:

1. **Change Management:** The AH should provide greater context when trying to plan for the implementation of new systems, new processes, and new workflows.
2. **EMR Development:** As patient treatment complexities evolve, so must EMRs. Providing a better context and understanding of patient treatment could offer valuable insights to EMR developers. Developing a modern model of patient treatment addresses a gap and could lead to the design of improved EMR systems.
3. **Additional AHs:** Clear value is seen in understanding healthcare as a complex system. An AH of patient treatment could provide a basis for additional analysis. For example, understanding patient treatment would be a precursor to understanding the management of health information and data codification.

### **4.2 Model Development**

#### **4.2.1 Model Context**

The study was conducted through collaborations with SMEs in Ontario. This included managers and clinicians within a community hospital and at a FHT. Some of the concepts that were included are reflections of a single-payer system in Canada and reflect a Canadian perspective on social determinants of health (Mikkonen & Raphael, 2010). The ideas of for-profit healthcare and third-party insurance companies are not present.

#### **4.2.2 Information Gathering**

Collection of data to build and validate models for the CWA framework is not very well described in the literature. The scope of my work did not include making a contribution to this gap, and I adopted standard practices in the field to meet with SMEs and users to inform my work. The data collection was semi-structured and did not follow a rigid social science approach, such as grounded-theory. The development of the AH and the WDA was meant to capture important aspects of patient

treatment, but was not intended to be a perfect representation of all caveats in the domain. The WDA served as a guide for inquiry.

I gathered my qualitative data to conduct a WDA by contacting local thought leaders in the community. To understand how data is used for secondary purposes, I met with hospital directors, program managers, and executives. To understand a clinical perspective, I met with several doctors and observed users who were entering data. My observations and interviews lasted approximately an hour each, and I took detailed notes. I let participants discuss their ideas and thoughts openly, and asked follow up questions to direct the conversation to areas of interest.

After taking notes with each participant, I transcribed the notes to a computer and summarized the conversation and ideas. I shared my notes and summaries with the participants. The participants reviewed the notes and provided additional clarifications and notes about the important themes and ideas I had summarized.

Based on the results of each interview, I directed questions to subsequent participants to explore ideas that I had not fully understood, to get a second opinion on a topic, or to explore new areas. My goal was to validate my present models and expand on them with new information. When contradictions occurred, I asked more participants to clarify the topic, and when similar comments arose, I focused the interview time on new areas. I interviewed participants until I felt confident about the resulting AH. I met with some participants a second time to review the AH in detail and confirm my understanding of the concepts and ideas of our previous meeting. This enabled me to validate my results.

The development of the AH took place over the span of 12 months. A total of 10 iterations and versions of the AH were created before the development of the final version. In addition to working with SMEs, information for the WDA was collected by reviewing textbooks such as pathophysiology textbooks (Gould & Dyer, 2014) and health system textbooks (Thompson, 2015), best practice guidelines, professional standards (College of Physicians and Surgeons of Ontario, 2012a, 2012b), and existing literature (Biro et al., 2016; de Lusignan, 2003; de Lusignan & van Weel, 2006; de Lusignan et al., 2003; Elwyn et al., 2012; Gagnon et al., 2010; Jiancaro et al., 2013; Majeed, 2004; Saleem et al., 2009; St-Maurice et al., 2013). Insightful information and anecdotes also were gathered during previous research (St-Maurice & Burns, 2014).

### **4.2.3 Abstraction Hierarchy Development**

The development of AHs is challenging because modelling abstract concepts and ideas can be accomplished in many ways. Model developers need to engage and observe users and articulate thoughts and suggestions into the AH. Often, the literal suggestions and ideas from SMEs must be abstracted into



high-level concepts and ideas. AHs are intended to be helpful, but not perfect, and managing the scope and level of detail of the modelling exercise is a challenge in and of itself.

The first phase of a WDA is to determine the system boundary. A balance must be achieved in the analysis: a domain boundary that is too narrow will leave out connections and interactions that exist outside the boundary, whereas a broad boundary can distract the modelling effort as time is spent developing concepts that are not germane to the modelling objective (Burns et al., 2009). The scope of work was primarily clinical, and the boundary was restricted to activities that were within the clinician's control during a patient's use of services, even if they were indirect. The patient and their attributes were included in the scope of the analysis. Patient flows and activities outside of an encounter with a clinician were excluded from the scope (e.g. patient opting not to take medicines, choosing to perform exercises, adjusting diet, consulting with family, etc.).

Patient process flow diagrams (Figure 3) and information process flow diagrams (Figure 4) were developed with SMEs to describe the generalized activities of the clinic and its clinicians. As the process flow diagrams presented generalized overviews of common processes, not all components of the process diagrams were necessarily 'activated' during each patient encounter; they represented possible processes during a visit. In the case of patient flows, most functions within the clinic (involving a combination of triage, assessment, treatment, care transfer, and scheduling) were captured. In the case of information flows, most functions involving the EMR (involving a combination of summarization, sharing, updating, and interpretation) were also captured. These flow diagrams later were translated into the 'Generalized Function' layer of the AH.

Figure 3. General patient process functions.

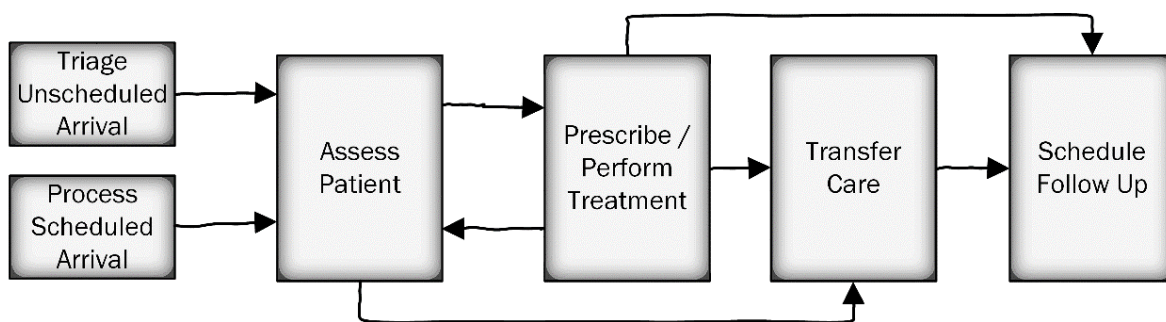
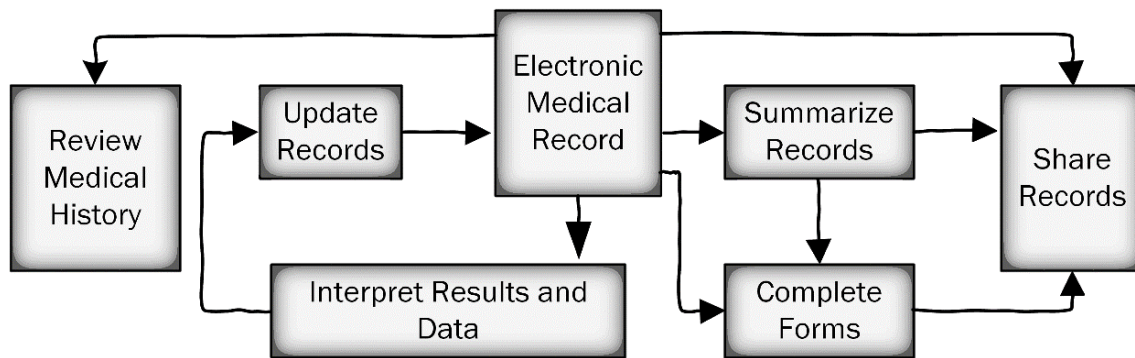


Figure 4. General information process functions.



Using my boundary definition, previous work regarding medical records (St-Maurice & Burns, 2014), and discussions with SMEs, five goals were developed to describe the purpose of patient treatment. These formed the ‘Functional Purpose’ layer of the AH. Whereas a purely biomedical treatment goal would be to ‘improve health’, concepts such as patient education and public safety were included in the scope of patient treatment.

After describing goals and process flows, a list of concepts that linked these layers was developed with help from SMEs. This included articulating abstract concepts such as values and balances, and showing how system goals were mediated to perform functions. These concepts were translated into the ‘Abstract Function’ layer of the AH.

The physical function layer of the AH represented concepts, objects, and actors that were needed to perform the patient and information processes. The physical form represented details and attributes of the objects and actors that were relevant to the system processes. For example, the social status and severity of symptoms were relevant attributes of the patient.

### 4.3 Modelling Results

The patient flows and information flows each were placed into separate views of the same AH (see Appendix A). Showing two views increased the readability of the hierarchy and allowed each type of process to be displayed separately. Other than the generalized functions and specifically noted omissions, all elements of the model are shown in each view.

#### 4.3.1 Functional Purpose and Treatment Goals

Five functional purposes were identified in the AH. The treatment purposes (e.g. goals) included concepts of patient education, financial compensation, health improvement, sustainable care, and public safety. These goals are linked to abstract functions which represent constraints to be respected in achieving each goal. In some situations, each goal is met during treatment. In some situations, one goal

may take priority over the other. For example, a patient with a communicable disease may need to be quarantined to ensure public safety at the expense of their individual wellness and freedom. However, the goal of the modelled system generally is to achieve all goals simultaneously outside of fringe cases. The details underlying these concepts were developed in consultation with SMEs.

#### 4.3.1.1 Functional Purpose: Educate Patient

As part of treating patients, clinicians aim to educate patients. This includes providing information about health conditions, treatments, and lifestyle. Educating patients is an important goal in their treatment since poor education or incorrect information can interfere with treatment and must be considered as a goal. For example, SMEs mentioned that some patients may not wish to be vaccinated based on individual patient beliefs about vaccines. In this context, the overall goal of treating a patient is a combination of education, improving their health, and ensuring public safety from communicable diseases.

As shown in the AH, during treatment, patient education is mediated by patient means and abilities (e.g. patients who cannot afford physio therapy might be educated about exercises instead of receiving a referral) and patient values (e.g. not being willing to accept a certain treatment).

#### 4.3.1.2 Functional Purpose: Receive Financial Compensation

In Ontario, fee-for-service payments are provided by the Government, a third party insurance provider, or the patients themselves. In other situations, such as clinicians who are part of a FHT or physicians who work at a Community Health Centre, clinicians are salaried and employed by the Government to provide healthcare services and treat patients. Sometimes physicians are compensated through a combination of patient-capitation (e.g. payment per patient per year), by the services provided and according to special bonuses for achieving specific care practices (Thompson, 2015).

While treatment could be modelled altruistically, payment to clinicians impact the treatment approach. As mentioned by SMEs, some doctors in the fee-for-service model adopt a ‘one visit, one problem’ approach to maximize potential remuneration. Since this decision is influenced by financial remuneration and not driven by best practice or health outcomes, this concept is important to capture in the AH and show as a treatment goal that impacts clinical processes through abstract functions.

As shown in the AH, compensation is mediated by a patient’s resources (e.g. ability to pay uncovered costs and fees), best practice guidelines (e.g. government bonuses for specific additional interventions, which are based on best practice guidelines), patient flow (e.g. volume and theoretical maximum billable time), system resources (e.g. the government budget), professional values and training (e.g. what services can be performed and opting to select strategies favouring maximum remuneration),

balancing risks versus benefits (e.g. determining if receiving compensation for the treatment is worth any potential risks or benefits to the patient), and professional standards (e.g. what actions are permitted, ethical, and appropriate).

#### 4.3.1.3 Functional Purpose: Improve Patient Health and Wellness

An obvious goal of patient treatment is to improve patient health and wellness. Patients who are not looking to improve their health or wellbeing (directly or indirectly) will not seek treatment. SMEs mentioned that sometimes patients seek assistance for social reasons and not for strictly medical reasons; the biopsychosocial nature of care accords services to patients who are isolated socially, or experiencing significant life challenges such as job loss or homelessness.

Improving patient health is mediated by patient resources (e.g. financial ability to pay clinicians when required, afford drugs, or have social supports to support care), the ability to actually see the patient (e.g. patient flow), patient values and beliefs (e.g. willingness to accept recommendations), best practice guidelines, health system constraints (e.g. scheduling constraints for referrals), physiology and psychology principles, professional values and training (e.g. what treatment can be performed), and balancing the costs and benefits of a treatment plan.

#### 4.3.1.4 Functional Purpose: Provide Sustainable Care

In Ontario, clinicians need to select appropriate tests and treatments that support a sustainable healthcare system. Clinicians also need to avoid unnecessary procedures that are of limited clinical value. For example, SMEs described patients who request ‘fad’ bloodwork, such as a vitamin test, without a clinical reason. Unlike fully privatized healthcare systems, clinicians need to make treatment choices that respect the public purse and support a sustainable healthcare system by ensuring diagnostics are medically necessary. This type of conflict resolution is challenging (Weston, 2001) and is important to include as a constraint in treatment. Patients are not always able to receive the tests and treatments that they want because of limited health resources.

Sustainable care is moderated by best practice guidelines, patient flow (e.g. volume and capacity), system constraints (e.g. budgetary limits), professional values (e.g. caring about the public purse), and professional standards of care (e.g. guidelines).

#### 4.3.1.5 Functional Purpose: Ensure Public Safety

Clinicians must place individual patient treatment into the context of public safety. Patients who are a danger to others, have communicable diseases, or could endanger their community in other ways (e.g. poor eyesight in a senior citizen who drives) require interventions that are not necessarily in the patient’s best interest. For example, SMEs discussed that taking a senior citizen’s driver’s license may protect

public safety, but may also result in social isolation and poor medical outcomes for the individual patient. Public safety is an important element to model in health care. Ensuring public safety is moderated by professional values and training, the balancing of risks versus outcomes, and professional standards. This also has an impact on many information flows, such as mandatory reporting requirements (College of Physicians and Surgeons of Ontario, 2012a).

#### **4.3.2 Abstract Functions and Treatment Constraints**

Abstract functions represent constraints that need to be respected during clinical processes (e.g. generalized functions) to achieve the system's treatment goals. These concepts were created in consultation with SMEs.

##### **4.3.2.1 Abstract Function: Balance Patient Means and Abilities**

Patient means (e.g. financial, social) and abilities (e.g. mental competency, self-care) need to be balanced and considered in their treatment. For example, SMEs mentioned that a physician will need to take a patient's ability to pay for drugs into consideration when issuing a prescription or recommending physiotherapy. Likewise, an elderly patient's access to peer groups and family would impact their ability to live at home or require homecare. This constraint influences treatment functions and plays a role in how medical records are processed (e.g. looking up patient details and social circumstances) and summarized for sharing (e.g. summarizing data for a referral).

##### **4.3.2.2 Abstract Function: Patient Values and Beliefs**

Patients have varying worldviews and values that need to be understood and balanced during treatment. For example, SMEs mentioned that some religions would object to blood transfusions, some cultures will not tolerate birth control, some peer groups adhere to false information about vaccines, and some female patients may be uncomfortable with a male doctor performing certain medical procedures. This abstract concept plays a role in patient assessment and treatment procedures. Patient beliefs also may play a role in how information is shared with other providers, based on patient perspective about privacy rules and regulations (Perera, Holbrook, Thabane, Foster, & Willison, 2011).

##### **4.3.2.3 Abstract Function: Best Practice Guidelines**

Best Practice Guidelines suggest health screenings, preventative tests, and appropriate actions for patients with specific characteristics (e.g. age, diagnosis, etc.). SMEs referred to guidelines that recommend specific treatment functions (e.g. recommending a test), or specify that a patient be transferred to another level of care (e.g. sending a patient to a stroke unit from the emergency room). Best practice guidelines have a significant impact on the review of medical history. The constraints on

treatment that are associated with best practice guidelines are represented through this abstract function, but guideline documents were not included in the Physical Function of the AH in order to manage project scope.

#### 4.3.2.4 Abstract Function: Patient Flow

Patient flow is a representation of patients entering, moving through, and exiting the treatment process. Patient flow represents limits related to patient volume and throughput. Patient volume is an important constraint on the system, as the flow of the patient through the clinic and the healthcare system must be taken into consideration, and is important for all generalized functions. Without capacity, treatment is not possible.

#### 4.3.2.5 Abstract Function: Balance System Resources and Constraints

As a single-payer, publicly-funded healthcare system, healthcare dollars and resources in Ontario must be taken into consideration during treatment. Not all drugs or treatments are available, and some procedures have significant waiting lists due to insufficient system resources (e.g. number of beds, number of surgeons, etc.). This abstract function describes a constraint in selecting treatment options for patients while achieving treatment goals.

#### 4.3.2.6 Abstract Function: Physiology and Psychology Principles

Human anatomy, physiology, and pathophysiology principles are important constraints to be considered during treatment. When patients are suffering from situations that are not strictly biomedical in nature (e.g. social distress, isolation, stress, etc.) psychological principles need to be taken into account. This abstract function helps describe constraints during triage, patient assessment, treatment, and transfer of care. From an information perspective, these principles are important when clinicians interpret results and data and update the medical record.

#### 4.3.2.7 Abstract Function: Professional Values and Training

Clinicians are not uniform in their decisions. As with patients, clinicians have worldviews, professional values, and priorities. For example, physicians may choose to see more patients in a day (e.g. volume) and provide care to a large number of patients, or may choose to see fewer patients for full assessments to provide higher-quality care. Worldviews also may impact ethical decisions, such as valuing the public purse. A professional's scope of practice, practice style, and set of priorities is based on training and personality characteristics. This abstract value system plays a role in assessing patients, performing treatments, and deciding when it is appropriate to transfer care. It also plays a role in a

clinician's interest in creating high-quality documentation that is above minimum standards. Patient and documentation processes are constrained by professional values and training.

#### 4.3.2.8 Abstract Function: Balance of Risks, Costs, and Outcomes

Whenever treatment is provided to a patient, there are risks, possible outcomes, and costs. If a clinician determines that the risk is high and the probability of a positive outcome is low, another treatment option may be selected. Similarly, a clinician may balance the healthcare costs of surgery for an arthritic patient versus a prescription, and make a treatment determination that is based on total costs, recovery periods, and quality of life. Risk balancing takes place in consultation with patients who describe their preferences and capabilities. In situations where patients pose a risk to public safety, a clinician must make an appropriate determination between risks and potential negative outcomes to the patient and public.

This abstract concept plays a role in assessments and treatments. Risks also are evaluated when choosing to transfer care. Information functions assist in determining risk.

#### 4.3.2.9 Abstract Function: Professional Standards

All clinicians are governed by professional associations and colleges. For example, physicians in Ontario are governed by the College of Physicians and Surgeons of Ontario (CPSO). The CPSO establishes specific conditions and training requirements for all physicians in Ontario. They have policies on medical records (College of Physicians and Surgeons of Ontario, 2012b) and provide guidelines regarding reporting information to third parties (College of Physicians and Surgeons of Ontario, 2012a). The concept of professional standards constrains patient assessment; prescription and treatment; transfer of care; and maintaining, reviewing, updating, and sharing medical records. To manage the scope of the domain analysis, the standards documents were not included in the scope of the model and are not included in the Physical Function of the AH.

#### 4.3.2.10 Abstract Function: Information Flow

Information Flow is a representation of information that enters the system, and is used and stored in an EMR. Information flow is important in managing care and impacts decision making and timing. If information is not available when needed, it will affect many aspects of treatment. As an abstract concept, information flow is important through all information functions in the Generalized Function layer of the model. Information flow impacts financial compensation (e.g. ability to bill and document encounters), patient health (e.g. improved care quality through information), and public safety (e.g. reporting mandatory information to appropriate authorities).

### **4.3.3 Generalized Treatment Processes**

The generalized functions represent the general processes in health care, as described in Figure 3 and Figure 4. Each generalized function was linked to abstract function constraints that had to be respected to achieve the system goals, and linked to the appropriate physical components of the processes.

### **4.3.4 Physical Treatment Elements and Attributes**

The physical functions layer of the AH represents concepts, objects, and actors that were needed to perform the processes modelled in the generalized functions. The physical form represents details and attributes of the objects and actors that are relevant to the system processes. Keeping in mind that the clinician is the system controller (and is not represented in the physical form), the relevant actors and objects in the AH include the patient, type of assessment, clinic staff, level of care, forms, and medical records.

#### **4.3.4.1 Physical Form: Patient**

The patient is obviously an important actor associated with all generalized functions. The patient's attributes that are relevant in treatment include patient's family and friends (e.g. presence of social supports to facilitate treatment), the patient's biopsychosocial status (e.g. social circumstances such as employment, stressors, etc.), the severity of the patient's symptoms or problems, and the complexity of the clinical case. The patient and their most important attributes are included in the model as they affect the entire treatment ecosystem.

#### **4.3.4.2 Physical Form: Assessment Type**

Different types of assessments are used. A physical exam would be detailed, whereas a 10-minute assessment would be problem-oriented. Other assessments may play the role of triage and refer a patient directly to the hospital (from primary care) or admit a patient (from the emergency room). Severity and complexity play roles in the type of assessment that will be used with the patient.

#### **4.3.4.3 Physical Form: Clinic Staff**

The clinical staff support many processes. Depending on the specializations and location of the care delivery, resources may be greater or fewer. Larger clinics with multiple clinicians will have a larger support staff with specific roles and responsibilities. Smaller clinics with an individual doctor may only have a single support resource who plays a generalist role. The type of staff and their abilities varies according to location of the practice.



#### 4.3.4.4 Physical Form: Forms

Many forms are employed to support the information processes during treatment. The location of the forms and their type (paper or electronic) are relevant attributes to the information flows and processes described in Figure 4.

#### 4.3.4.5 Physical Form: Medical Record

Medical records support all the information flow processes. The location of the status (e.g. availability), record type (e.g. paper or electronic), and location of the system are relevant attributes.

### 4.4 Discussion

#### 4.4.1 Comparisons to Other Models

My AH provides a high-level perspective of patient treatment within a biopsychosocial perspective and includes information procedures in separate views. My AH is different compared to existing models; as an AH, the model can articulate complex ideas within the patient treatment environment. It is a formative reference model. No existing AH describes how treatment takes place with clinician-controllers and modern patient. As a macro-level view of patient care, the model is similar to a model of medication administration in home care, which facilitated an in-depth understanding of medication safety problems and analysed medication errors (Lim et al., 2015).

#### 4.4.2 Design Implications

WDAs and AHs are consumed during design by using the EID approach (Burns & Hajdukiewicz, 2004; Momtahan & Burns, 2004). The AH can support system designers by properly articulating the ecosystem and clinician decision making in context. The model supports system thinking and can help articulate how changes may impact the ecosystem through linear and ripple effects. (Hignett, Carayon, Buckle, & Catchpole, 2013). Based on my analysis, the decision support requirements for health care are becoming increasingly complex. The challenge for system engineers will be to determine how electronic systems could support, and not hinder, the treatment process. In addition, the analysis is a reminder that technology-centric solutions and implementations that do not take the larger healthcare environment into consideration during the entire treatment process will likely fail to thrive. Creating a product that is compatible with the nuances that are described in the AH would be a competitive advantage.

#### 4.4.3 Limitations

My AH is intended to be helpful, but not perfect. The model is limited to a clinician's perspective and aims to provide a high-level overview of treatment. Obvious opportunities are present for a deeper

analysis of the work domain in special areas. For example, complex nuances to medication prescription and administration have been simplified and abstracted in my model as ‘Prescribe / Perform Treatment’. It would be possible to do a more detailed WDA on this specific issue. For example, Lim, Anderson, and Buckle developed a detailed analysis of medication administration in home care (Lim et al., 2015), and this could be performed within the scope of patient treatment. In this sense, my work is incomplete. In this same sense, the amount of modelling to be performed is infinite, and my hierarchy is a contextual overview that could serve as a blueprint for additional work.

#### **4.4.4 Future Work**

The current AH describes patient treatment and takes a biopsychosocial perspective over a biomedical one. Taking a patient centred perspective further, the AH could more formally incorporate aspects of SDM thinking. This would be compatible with the current work, as general qualities of treatment with SDM include deliberation with patients, an individualized approach, information exchange, involvement of multiple parties, finding middle ground, espousing mutual respect, developing patient education, encouraging patient participation, and following a process with stages (Makoul & Clayman, 2006). Adopting SDM is considered important because inadequate patient involvement in decision making can result in poor health outcomes (Thompson-Leduc, Turcotte, Labrecque, & Légaré, 2016). However, it is important to note that SDM is not always easy for clinicians to implement, and barriers exist to its use in patient care: in addition to requiring more time, it also might not apply to the patient’s characteristics or their clinical situation (Légaré et al., 2008). Thus, a goal would be to capture SDM and non SDM procedures, values, and concepts.

It would be interesting to compare SDM and non-SDM perspectives to patient care. Inviting SMEs to comment and develop a similar AH could lead to an interesting comparison of work, since the current work does include a SDM expert in its development. Such a comparison could help to describe the perceptions and realities of what shared decision making is and how it is (or is not) incorporated in routine clinical care. The idea of drawing comparisons is further discussed in chapter 6.

Another interesting perspective about SDM is that it is a shared process between at least two actors; colloquially, SDM has been described as a dance between providers and patients (Edwards & Elwyn, 2006). Thus, developing a full perspective of SDM will require at least one other AH describing patients as a controller. Work by Rezai & Burns (2014) could provide a good starting point for developing an AH from a patient perspective. Team perspectives to patient care modelled with SOCA (see Ashoori et al. [2014]) also could be helpful for understanding SDM in care teams comprised of family physicians, nurse practitioners, pharmacists, medical specialists, caregivers, and patients (Légaré et al., 2008; Yu et

al., 2014). Generally, further work on this AH and line of inquiry could lead to interesting contributions to SDM research.

## **4.5 Chapter Summary**

### **4.5.1 Key Findings**

I created an AH to represent the findings of a patient treatment WDA. The AH was developed through collaborations with SMEs; it represents a biopsychosocial perspective in patient treatment, and includes processes that are associated with electronic information management. The AH is broken down into two views which capture clinical processes and information management. The model can be used by system developers to improve systems in health care by better supporting complex decision making in context. The model also can be used as a reference to describe different healthcare systems in various environments, such as primary care clinic, or emergency room environments. There are interesting opportunities to further explore patient centred care and SDM by continuing to build on this AH.

### **4.5.2 Connections to Case Study and Research Questions**

This chapter helped define the constraints within the overarching work context, within which the remainder of my analysis takes place. Subsequent chapters will analyse elements of this large space in increasing levels of detail. This chapter provides context for my overall work, but does not directly answer any of my research questions.

### **4.5.3 Connections to Other Chapters**

I use the AH of patient treatment in the chapter 5 as part of my CWA of data codification.



## **Chapter 5**

### **Data Codification in Healthcare**

In the previous chapter, I developed an AH for patient treatment. In this chapter, I construct models with CWA to better understand data and data quality in primary care and incorporate my analysis of patient treatment. To encompass a broad perspective, I develop two sets of CWA models. The first CWA represents codification of data at a FHT in Ontario, and the second CWA represents data codification in a community hospital. The reason for creating two separate models is to draw comparisons between both environments and extract insight from each domain. The comparison of these models is done in chapter 6. I discuss using the models from this chapter, and their comparisons for design, in chapter 8.

#### **5.1 Two Healthcare Settings**

I created two sets of CWA models from two similar healthcare ecosystems in Ontario. The reason for studying two separate systems was to broaden my overall understanding of data codification in healthcare, and not rely exclusively on observations and information from primary care. This parallels the use of the Comparative Cognitive Task Analysis (C2TA) by Kirschenbaum, Trafton, and Pratt (2007). In their study, they argued that modelling multiple parallel domains resulted in more generalizable results and enabled a better and broader application of results. In this case, I am conducting two analyses using CWA to broaden my perspective of data quality. Later, I formally make comparisons between my modelling outputs.

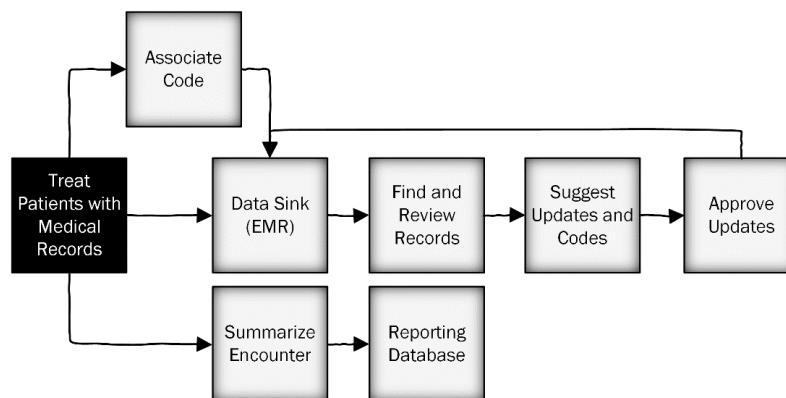
The two domains I studied were data codification at a FHT, and data codification at a hospital. Both domains aim to treat patients effectively, aspire to create useful data for analysis, and have reporting requirements to government agencies. Early on, I also noted the presence of interesting differences: the hospital uses a different type of medical record system compared to the FHT (hybrid vs. fully electronic); the organizations provide different types of care (primary care vs. urgent care) and are therefore governed differently; and data is coded by different types of employees (e.g. in the hospital, HIMPs are responsible for coding data, vs. clinicians who are responsible at the FHT). Although these differences were obvious, it was unclear how they would impact data codification. My goal was not to perform a full CWA analysis

of both systems, but rather to focus on tasks and processes that contributed to the accurate summarization of clinical information and its capture through structured input fields.

### 5.1.1 The Family Health Team

As a starting point for understanding and analysing data codification at the FHT, I developed a series of process flow diagrams to represent relevant codification tasks. This initial process mapping was accomplished as a result of interviews with SMEs and a preliminary research project concerning user impressions of data and data quality (see St-Maurice & Burns, 2014). The codification processes are shown in Figure 5. As depicted, data is created during patient treatment and placed into an EMR. Depending on user preference or ability, while treatment takes place, a clinician may associate a code and insert this information into the EMR system while creating medical records. In other situations, the records from the EMR are searched and reviewed by support staff or AHPs after a patient visit, and new codes or record enhancements are recommended to physicians. These updates must be approved by the record owner (typically, the physician associated with the patient). For clinicians working at FHTs, data needs to be summarized into a reporting database. AHPs typically perform this work themselves and generate the information for their record based on their work with patients.

Figure 5. Codification processes at a Family Health Team.



Several important notes should be mentioned concerning FHTs and data codification:

1. Every regulated profession has policies recording clinical documentation and medical records. For example, in Ontario the College of Physicians and Surgeons of Ontario (CPSO) has a medical records policy that outlines specific requirements and guidelines (2012a). Data codification is not typically addressed in these policies. Thus, primary care clinicians are under no obligation to code data.
2. Primary care clinics operate independently and are, arguably, small businesses. Lacking a universal standard, clinicians and FHTs who choose to codify data are not bound by a collective standard.

Physicians and FHTs can make individual choices about terminology, coding, and decide which clinical processes merit codification.

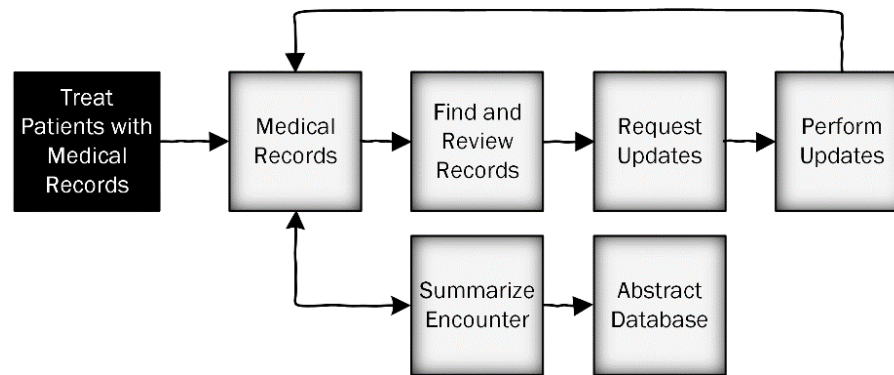
3. There is a difference between coding a diagnosis and making a diagnosis. For example, lab results for a patient may clearly indicate the patient suffers from diabetes. A physician is ultimately responsible for associating a diabetes diagnosis in the patient's medical record, and may choose to omit the diagnosis. This is a different issue than coding the diagnosis in an accurate way and supports secondary use.
4. FHTs have an obligation to report their activities to the government. The process to collect this data is based on capturing structured data, but varies from FHT to FHT.
5. Data codification is a clinical responsibility, and generally is not offloaded to clerical or administrative support staff. The responsibility to code, and acceptance of codes, typically lies with the patient's physician.

Colloquially, primary care is sometimes described as the 'wild wild west'. This is an allusion to lawlessness, or the lack of strict policies or procedures in primary care. From a data codification perspective, this characterization seems appropriate.

### **5.1.2 The Community Hospital**

As a starting point for understanding and analysing data codification at the hospital, I developed a series of process diagrams to represent relevant codification tasks. This initial process mapping was performed as a result of interviews with SMEs. As well, I am presently enrolled in a diploma program to become a Certified Health Information Management Professional (CHIMP). In addition to theoretical lessons, I have had several hospital practicums where I worked in the health information management (HIM) department of the hospital, and abstracted data. The hospital codification processes are shown in Figure 6. As depicted, the HIM department is responsible for codifying data, and employees of the department perform their duties by reviewing medical records. Since the HIM department is not associated with direct patient care, it must rely entirely on the information within the medical record to abstract and codify data. During quality assurance processes, the department finds and reviews records and may request updates from clinicians if any additional documentation is missing.

Figure 6. Codification processes at a community hospital.



Several important points about hospitals and data codification can be noted:

1. Hospitals have specific medical records policies, processes, and dedicated staff.
2. Hospitals need to provide information to government agencies for mandatory reporting purposes. New funding formulas in Ontario also rely on coded historic data. Consistent coding from hospital to hospital is assured by ethical guidelines, quality audits, and a management structure.
3. Hospitals own medical records and have a mandate to ‘force’ clinicians to update deficient records.
4. Data codification is not a clinical responsibility. Data codification and records management is managed by HIMPs from the Health Information Management (HIM) department. However, some user interface (UI) implementations of hospital systems ask clinicians to select items from lists and structure data elements, which is increasingly comparable to the work of primary care clinicians.

## 5.2 Methods

Two sets of CWA models were developed for two different healthcare ecosystems in Ontario.

### 5.2.1 Boundary and Scope

The first step in conducting a CWA is to determine a system boundary for the WDA and to define a scope for the task analysis. Without a well-defined focus, a CWA can be unnecessarily long, and the length and breadth of a CWA are not proportional to its overall value. Thus, articulating goals and defining a scope early in the modelling task is critical to making a good use of time and extracting maximum value. In my analysis, the WDA was scoped for understanding the work domain associated with secondary use of data. The boundary was restricted to activities and tasks that were part of a manager’s work scope, influence, and control. This included tasks and processes that related to coding data and managing records, and included aspects that a manager could understand, oversee, and control



through policies and procedures, even if indirect. If a user had two roles (e.g. clinician and manager), the WDA was conducted to cover their scope of work and needs as managers.

### **5.2.2 Information Gathering**

The development of the CWA models occurred over the span of 12 months and was done alongside the patient treatment AH from chapter 4. Similarly, data was collected from meeting with eight SMEs, iterating through various model concepts, gathering feedback, and reworking my CWA models as required. Information was also collected by reviewing textbooks such as pathophysiology textbooks (Gould & Dyer, 2014) and health system textbooks (Thompson, 2015), best practice guidelines, professional standards (College of Physicians and Surgeons of Ontario, 2012a, 2012b), and prior published works on the subject (Biro et al., 2016; de Lusignan, 2003; de Lusignan & van Weel, 2006; de Lusignan et al., 2003; Elwyn et al., 2012; Gagnon et al., 2010; Jiancaro et al., 2013; Majeed, 2004; Saleem et al., 2009; St-Maurice et al., 2013). Insightful information and anecdotes also were gathered during previous research (St-Maurice & Burns, 2014). Several iterations of each phase of the CWA exercise were produced before the final versions were completed, validated, and reproduced in Microsoft Visio.

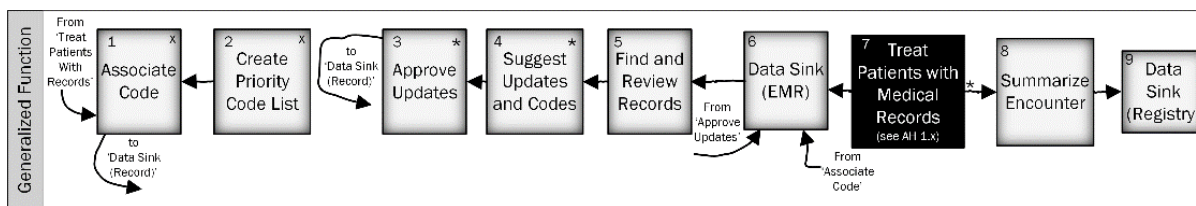
### **5.2.3 Cognitive Work Analysis**

#### **5.2.3.1 Work Domain and Environment**

The development of AHs is challenging because many ways can be employed to model abstract concepts and ideas. Practitioners must engage and observe users and articulate thoughts and suggestions into the useful models. Often, the literal suggestions and ideas from SMEs need to be generalized and abstracted into high-level concepts and ideas. The outputs of CWA are intended to be helpful, but not perfect. The models and the requirements that are extracted from them are intended to provide helpful insights about the ecosystem that could be applied by designers.

Using my boundary definition, previous work regarding medical records (St-Maurice & Burns, 2014), and discussions with SMEs, I developed an AH to describe data codification in primary care and within hospitals. The two AHs were developed in tandem. As each of the AHs was created, I kept track of any elements that were different with a ‘\*’, and denoted elements that were not present in the other AH with an ‘x’. All boxes from the AH were annotated with numbers to allow easy cross-referencing and comparisons. This is shown in Figure 7. Marking differences also facilitated later comparisons.

Figure 7. Flagged differences on Abstraction Hierarchy.



### 5.2.3.2 Task Analysis

After completing the AHs, several user tasks could be modelled and further studied. Because I was interested in how data is coded and structured for secondary uses, I selected two tasks that were directly related to my research program. The generalized function for ‘Summarize Encounter’ existed in both systems. At a FHT, the implementation of this function is the ‘Record Encounter in Registry’ task, which is performed by clinicians. At the community hospital, the implementation of this function was the ‘Abstract Encounter to Registry’ task, which is performed by HIMPs.

These two tasks were analysed using ConTA, StrA, and WCA. Since these tasks were generally performed by individuals, and not teams, a SOCA analysis was not completed. For the WCA phase, I linked the competency requirements back to the AHs that were created, in order to link the task analysis and work domain.

## 5.3 Results

### 5.3.1 Work Domain Analysis

Two AHs were created to represent the findings of my WDA. Each model incorporated AH 1.A and AH 1.B that were developed in chapter 4Chapter 3 (refer to Appendix A). An AH for primary care (AH 2.1) and an AH for a community hospital (AH 3.1) are presented in Appendix B. For each AH, tables describe each box in the AH and provide a context for each box in the Functional Purposes, Abstract Function, Generalized Function, Physical Function, and Physical Form layers. These tables also are provided in the appendices.

### 5.3.2 Control Task Analysis

The generalized function ‘Summarize Encounter’ task was analysed at a FHT and a community hospital. At the FHT, this task involved inputting data into an organizational registry by clinicians by using a web-based reporting tool. At the community hospital, this was accomplished through abstraction tasks by professional coders. Neither implementation of the ‘Summarize Encounter’ task involved an EMR to generate data, and both tasks included manually inputting data into another system. Based on my

interviews and observations, the findings of the ConTA are described as two sets of ConTA DLs and summary tables in Appendix C, as described in the next paragraph.

A base model of the control task at the FHT is shown in DL 1.1 and the accompanying summary table, DLSUM 1.1, which describe the conceptual basis of the control task. Separate models represent the novice user (DL 1.1.1, DLSUM 1.1.1), the intermediate user (DL 1.1.2, DLSUM 1.1.2), and the expert user (DL 1.1.3, DLSUM 1.1.3). A base model of the control task in the community hospital is shown in DL 2.1 and the accompanying summary table, DLSUM 2.1, which illustrate the conceptual basis of the control task in the hospital. Separate models represent the novice user (DL 2.1.1, DLSUM 2.1.1) and the expert user (DL 2.1.2, DLSUM 2.1.2).

### **5.3.3 Strategies Analysis**

Based on my interviews and observations, I produced two sets of strategies that were used to complete the summarization and abstracting tasks in each environment. Details are provided in StrA IFMs and summary tables in Appendix D.

IFMs describing seven strategies for the ‘Summarize Encounter’ task at a FHT are shown in IFM 1.1. Each strategy is illustrated in detail in the table IFMSUM 1.1, as well as triggers that were associated with each strategy. IFMs describing three strategies for the ‘Summarize Encounter’ task in a community hospital are shown in IFM 2.1. Each strategy is explained in detail in the table IFMSUM 2.1, as well as triggers that were associated with each strategy.

### **5.3.4 Social Organization and Cooperation Analysis**

While analysing the ‘Summarize Encounter’ task at the FHT and in the community hospital, I found that users were summarizing encounters individually. The task was not team based and did not include social interactions. A SOCA analysis was not appropriate.

### **5.3.5 Worker Competencies Analysis**

I continued my study of the ‘Summarize Encounter’ task at a FHT and in a community hospital by describing the skill-based behaviour, rule-based behaviour, and knowledge-based behaviour (e.g. the SRK taxonomy) required to complete the task in each environment. The SRKs taxonomies are presented in Appendix E. SRK 1.1 represents my findings from the FHT, and SRK 2.1 gives my findings from the hospital.

## **5.4 Discussion**

Many benefits accrued from building CWA models for two systems in parallel. Mainly, broadening my scope to investigate similar processes into two different environments doubled the size of

my pool of SMEs, and gave me twice as many processes to observe and consider. I strongly believe that my models are both better because they were built in parallel. Both provided advantages by encouraging me to ask different types of questions and gave me ideas of concepts to model. As a result of my parallel development, my models have provided more generalizable results and may enable a broader application of my findings. These benefits are similar to those described by Kirschenbaum, Trafton, and Pratt (2007), who performed Cognitive Task Analysis (CTA) on two similar domains.

Although broadening the investigation provided two rich CWAs, it is not likely separate, independent analyses would have resulted in the same outputs. In this sense, the act of generating the models in tandem may have biased or influenced each model, and I may have missed some important nuances or interesting artefacts as a consequence. Whereas I have a clearer picture of codification in healthcare, it may have been at the expense of higher-resolution models of hospital or primary care.

Regardless of these potential challenges and biases, building the CWA models demonstrated the usefulness of the CWA framework for analysing complex socio-technical systems. The tools that CWA provided were helpful and offered many useful templates.

The results of my CWAs in this chapter are further discussed in chapter 6.

## 5.5 Chapter Summary

### 5.5.1 Key Findings

**Two CWAs:** In this chapter I created two CWAs of codification in similar domains. My results are shown in the Appendices of my dissertation. These CWAs are a first use of the framework to study this type of problem.

**Pros and Cons:** While creating two CWAs in tandem resulted in a larger pool of SMEs and exposed me to a bigger range of modelling ideas, it is possible that a cross-contamination of ideas occurred. This cross-contamination could have resulted in the outputs of both CWAs being more general than specific to my individual domains.

### 5.5.2 Connections to Case Study and Research Questions

I have three research questions:

1. In primary care, how are individual users influenced by their environment to input high-quality data?
2. What techniques could be used to design systems that persuade users to enter higher-quality data?
3. Is it possible to improve data quality in primary care by persuading users with the user interface?

This chapter involved an analysis of the codification task environment and helped clarify how users are influenced by their environment. For the case study, the chapter continued my analysis of context.

Through the development of two CWA models, it became clear that the data codification problem is complex and features interesting socio-technical components. By employing CWA, I learned that users codifying data in healthcare are influenced by various organizational considerations, such as:

- Time management decisions and a user's overall scope of work
- Role assignment and training
- Record ownership
- Professional values and standards
- Funding models
- Quality improvement initiatives

Generally, I found that the technology used to codify data in each domain was similar (e.g. data is entered by putting data into labelled fields). I also found that complex social and organizational structures exist, and these structures are crucial contributors to the overall process. Coding and data quality will not be improved purely through technological lenses, and influencing factors exist within a larger complex environment.

### **5.5.3 Connections to Other Chapters**

The two sets of CWAs are compared in chapter 6. The models presented in Appendix A through Appendix E will be used to support design in chapter 8.



## Chapter 6

# Comparing Cognitive Work Analysis Models

In the previous chapter, I developed two CWAs of codification in similar domains. In this chapter, I continue by comparing the models I created in the previous chapter. The idea behind this comparison is to review hospital data management processes, compare them to processes in primary care, identify gaps and differences, and transfer knowledge.

Performing a comparison of two CWAs is a novel approach and can be used to extract unique insights from comparable systems. The approach, itself, is a contribution to the field of human factors, and I have submitted a manuscript to *Theoretical Issues in Ergonomic Science*, which is under review at the time of this writing.

### 6.1 Introduction

In ergonomics, situations occasionally arise in which a practitioner wants to transfer ideas and solutions between dissimilar system domains. Transferring knowledge between two complex socio-technical systems has the potential to be very rewarding, but quite challenging. For example, attempts are often made to take lessons from aviation and leverage them to make healthcare safer, as shown in the examples of Kapur et al. (2016) and Clay-Williams and Colligan (2015). While these comparisons often are enlightening and interesting, they are not conducted systematically. Opportunities to improve comparisons can be taken with a careful analysis of work and environmental constraints in the comparator domains.

When making comparisons, several potential comparators might generate value to system designers:

- *Comparing Similar Systems with Different Performance Characteristics*: Suppose two systems appear to share similar processes and physical characteristics, but have clear performance differences. A comparison of the two complex domains could reveal interesting behaviours, attitudes, values, and training that create different system characteristics that may contribute to good or poor performance. The identification of these characteristics would be useful for design.

- *Comparing Modelling Choices*: Suppose a new paradigm to treatment care exists; it progresses away from paternalistic care, and focuses on patient-centred care. Building separate models with SMEs who are experts in both approaches of care and comparing the results could provide useful insights about approaches and perspectives to patient care. This use-case was mentioned in chapter 4.
- *Comparing Current and Future System States*: The implementation of rich software solutions in complex socio-technical domains requires the careful consideration of behaviours, attitudes, values, and training to support future change. A comparison of a complex socio-technical system's current state vs. its intended future state could identify gaps and mitigate potential problems.
- *Comparing Systems in Conflict*: Consider two systems in conflict, such as two nation-states or two corporations. A detailed analysis and comparison of each system's environment, abstract values, and goals could help parties to identify common ground and areas of difference. This could be useful during various types of negotiations.

In practice, many techniques and approaches exist for modelling processes and UI designs (see Limbourg and Vanderdonck 2004). The idea of making comparisons between these types of models is not novel, but only a few tools, such as the Concurrent Task Trees (CTT) (see Paternò, Mori, and Galiberti 2001), have been created to support task modelling with comparison features. These tools are useful for designing user interfaces with teams and support knowledge consolidation. Unified Modelling Language (UML) version control also can be thought of as a series of model differences and model unions (Alanen & Porres, 2003). Engaging in version control strategies using difference algorithms can be thought of as a comparison approach (see Kelter, Wehren, and Niere 2005). Version control and CTT are relatively straightforward comparison approaches, but are based on different use cases and address different analytical needs.

Transferring knowledge between complex socio-technical systems requires more rigour than simply comparing processes. For example, comparing the performance of health systems from an international perspective requires a deep understanding of each system's rich ecosystem, including culture, governance structures, infrastructure, worker capabilities, and values (Papanicolas & Smith, 2013). The dimensions for comparison extend far beyond process analysis. These dimensions heavily influence the context and execution of processes, and without taking these, the user, and their environmental characteristics into consideration, modelling and comparisons would be incomplete, and knowledge transfer limited. Thus, comparing performance between complex systems requires toolkits and frameworks that go beyond process modelling.



Several frameworks have a modelling scope that extends beyond process analysis and incorporates user cognitive features in the analysis. For example, Hierarchical Task Analysis (HTA) (see Shepherd, 1998) is a popular modelling technique that deals with the decomposition and analysis of tasks. It is a goal-oriented framework that allows for the modelling of the cognitive and physical processes associated with achieving a system goal by breaking down the processes into sub-goals. HTA is a cognitive engineering tool that is well suited for analysing cognitive tasks in a descriptive way; however, it has not been adapted to describe differences between systems.

Another modelling approach that addresses user cognition is CTA. CTA is a set of methods that includes user perception, cognition, and motor actions required to accomplish a task (Kirschenbaum et al., 2007). Unlike HTA, CTA has been used to compare different system solutions (see Crandall, Klein, and Hoffman 2006). In one example, Mioch, Mistrzyk, & Rister (2010) used CTA to compare two communication methods. In their study, voice control was compared to datalink communications between a cockpit and ground control. This study was accomplished by a semi-formal task analysis and allowed the authors to understand the differences between two cognitive approaches. The authors were able to make recommendations to improve airline safety. Comparisons with CTA were formalized as an approach by Kirschenbaum, Trafton, and Pratt (2007), who developed C2TA. In their study, they argued that modelling and analysing a single domain would *not* have necessarily met the needs of the users. The authors suggested that it is only possible to determine if processes and cycles have common elements compared to other forecasting environments by studying and comparing several similar systems. Thus, modelling, studying, and comparing similar systems resulted in better and more generalizable results. Clear advantages can be found by comparing similar systems.

Vicente (1999) argues that properly understanding work demands and complex systems requires an analysis of cognitive constraints and environmental constraints. Based on this criteria, understanding the cognitive requirements of a task is helpful for systems design but does not frame the work within the context of the task's functional goals, user tradeoffs, cultural and professional values, training requirements, or team composition. Thus, the drawback of CTA and C2TA comparisons is the absence of incorporating environmental constraints into the modelling and analysis. These environmental elements are important contributors to the differences between complex socio-technical systems, as evident when trying to compare healthcare systems and their performances (see Papanicolas and Smith 2013).

An interesting opportunity exists for transferring critical knowledge and lessons through comparisons. However, successful comparisons will need to incorporate the cognitive and environmental aspects of work to support meaningful and useful transfers. Thus, a gap is present in currently available tools: whereas it is possible to compare processes and the cognitive aspects of tasks, and while the

benefits of comparisons have been demonstrated, the impetus is to adapt an ecological framework to support comparisons.

## **6.2 Comparisons with Cognitive Work Analysis**

I have described a need for an ecological framework that could support complex domain comparisons. I propose that CWA is a suitable framework for this purpose due to its multi-phased approach to look at domain constraints that considers both environmental and task constraints.

### **6.2.1 Why CWA for comparisons?**

As an ecological framework, CWA analyses work demands by taking cognitive and environmental constraints into consideration. Thus, CWA provides the benefits of being able to compare cognitive aspects, such as C2TA, while framing the cognitive aspects within an environment and context that render the analysis more meaningful. CWA meets my need to compare complex system domains, provides a set of tools to incorporate a complex environment, and would enable rich comparisons.

As a tool, CWA is an excellent solution for drawing complex comparisons because the framework already incorporates some comparison concepts. For example, during WDA, different types of system constraints are essentially compared through several levels of abstraction, and during ConTA, different types of users (e.g. experts and novices) are contrasted. Thus, the CWA framework already provides mechanisms and tools for comparisons, and the framework itself provides an effective comparison paradigm. At a minimum, comparisons are compatible with the framework's intended use, and comparisons with CWA are a viable adaptation.

### **6.2.2 How to Compare CWA Outputs?**

The phases of CWA output different types of models. Abstraction Hierarchies (AH), Decision Ladder Trees (DLs), Information Flow Maps (IFMs), and the Skill, Rules, Knowledge (SRK) Taxonomies are outputs from the CWA framework that can be compared.

#### **6.2.2.1 Comparison of Abstraction Hierarchies**

Different types of hierarchical structures frequently are used to model complex systems. An AH is specified by a means-end relationship between levels: the purposes for which the system was designed (functional purpose); the intended causal structure of the process in terms of mass, energy, information, or value flows (abstract function); the basic functions the system is designed to achieve (generalized function); the characteristics of the components and the connections between them (physical function); and the appearance and spatial location of those components (physical function). The physical function sometimes is used to model relevant attributes of the components. Moving between levels of the hierarchy

provides a deeper understanding of the system by answering what, why, and how. Abstraction Decomposition Space (ADS) can be used in conjunction with the AH to break down the analysis into different levels of resolution (Vicente, 1999).

Obviously, without appropriate boundaries and context, an AH describing the system constraints of an orchestra would be difficult to compare to an AH describing the system constraints of an air traffic control centre; such a comparison would be meaningless. Thus, assume two system domains have been analysed and share similar scopes and boundaries that make them reasonably similar in order to compare their respective AHs.

As an example of AH comparison, Burns, Bisantz, and Roth (2004) wanted to understand the differences between two sets of AHs, and how the models reflected individual choices in the approach to a domain analysis. The study compared different (but very similar) command and control centres of separate naval combat vessels with the intent of understanding the choices of different CWA modellers. The comparison of AHs was done through a Venn diagram that visually described the common and different elements within each level of the AH. Although the aim of the study was to understand modelling choices in similar systems, the study demonstrated some of the initial mechanics required to compare two complex systems modelled with CWA.

#### 6.2.2.2 Comparison of Decision Ladder Trees

ConTA uses Rasmussen's DL trees (1974). DLs capture the flow of information processing associated with individual control tasks, and enable the use of shortcuts to describe expert and novice decision-making patterns (Vicente, 1999). DLs are represented by a series of data-processing (rectangles) and knowledge states (circles) that represent information-processing patterns. Consecutive states of knowledge are separated by mental activities to transform one state of knowledge to the next. ConTA requires information about the state of the work domain, and outputs the required actions on the work domain (Vicente, 1999).

Ideally all users gain experience and eventually become experts. Instead of trying to determine what would be required for a novice to accomplish a task, a modeller can instead try to determine the preconditions that have to be satisfied for expert performance. A cornerstone of ConTA is to model processing behaviour toward designing computer-based information systems that deliberately induce and support expert action, which thereby lead to gains in cognitive efficiency (Vicente, 1999). DLs are tools designed to achieve the objective of discovering and documenting expert cognitive sequences. Through the use of shortcuts, a practitioner can articulate novice and expert paths, and the DL serves as a tool to illuminate ways to induce expert action.

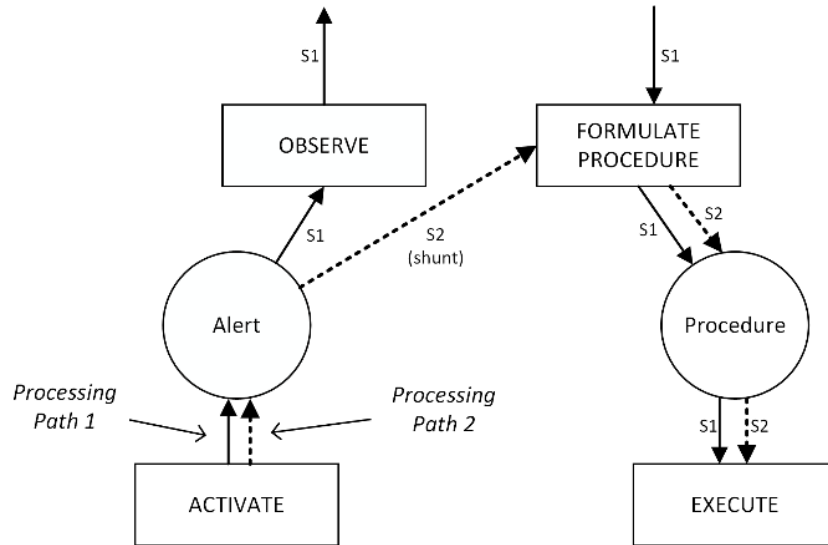
Expertise in a DL can be expressed through the creation of ‘shunting’ effects that bypass higher, abstract levels of thinking. As described by Rasmussen (1974), a trained operator will only occasionally have to move through all the steps of the basic sequence. Shunting results from subconscious or intuitive data processing; “mental activity may take place at a symbolic level, resulting directly in a knowledge state later in the sequence”, or “an association based upon previous experience may result in a leap directly from one state of knowledge to another one in the sequence”. These shunts evolve in the constructive learning process associated with training and are represented by arrows from data processing activities (e.g. boxes) to more abstract states of knowledge (e.g. circles).

In addition to shunts, DLs also allow for the description of ‘leaps’, which represent a direct transition from two separate knowledge states (e.g. going from a circle to a circle). Leaps represent intuitive data processing performed by the subconscious and describe associations that bypass more complex higher-level reasoning. Leaps rely on the large capacity of holistic perception and recognition, as well as the large data processing capacity of the subconscious functions of generalization and modelling (Rasmussen, 1974).

Using CWA to compare specific cognitive processing activities is useful because common tasks, established during WDA, are *not* necessarily the same. By developing and comparing tasks with ConTA, I can create a common template spanning two systems and explore variations in the processing paths in two similar ecologies. This approach is well within the spirit of Rasmussen’s seminal work (1974), which allows for dichotomous representations in DLs. This dichotomy of information, with different processing paths, is already present when comparing novice and expert user performance. The use of shunts and leaps is already common practice to represent differences, and this technique could be used to represent differences between systems, instead of differences between types of users. ConTA allows for a rich comparison of control tasks in an ecological context.

ConTA analyses can be compared by overlaying common elements and showing graphically where systems diverge and reconnect along common paths. To the best of my knowledge, however, there are no examples that I am aware of in which DLs have been compared from two different systems. In the overlay concept, shown in Figure 8, two systems (S1 and S2) share a common DL.

Figure 8. Two paths on a common Decision Ladder (St-Maurice & Burns, 2015).



### 6.2.2.3 Comparison of Information Flow Maps

IFMs are modelling tools from the CWA framework that can be used to describe how a task can be accomplished (Vicente, 1999). As an example of strategy comparison, Vicente (1999) describes a case study in which the strategies of technicians and engineers are contrasted when troubleshooting electronic equipment. The comparison of strategies used by each group is done through a table, in which different criterion or resource requirements used as attributes are shown in rows, and each strategy is shown as a column. This enables an effective high-level comparison of each strategy. IFMs also are used to model teamwork in SOCA. The comparison of team dynamics could be accomplished through a similar table, as well.

### 6.2.2.4 Comparison of SRK Taxonomies

The SRK hierarchy is a taxonomy, not a model. Each level distinguishes categories of human behaviour according to fundamentally different ways of representing the constraints in the environment, and is not a detailed model of psychological processes. Each level of the taxonomy corresponds to a category of human performance. The primary criterion in the development of the taxonomy is usefulness, and not necessarily ‘truth’. As a table, the SRK taxonomy outlines knowledge requirements, rule-based logic, and skill-based abilities. The taxonomy is useful for selecting users and developing effective training (Vicente, 1999).

As a taxonomy is described in a table form, comparing two taxonomies would involve the creation of a table to compare the skills, rules, and knowledge requirements between each system. This

side-by-side comparison of SRK taxonomies would easily highlight common and different user abilities between complex systems.

## 6.3 Comparison of Two Codification Systems

The purpose of the analysis and comparison of the systems was to identify specific ways that the domain of the FHT could produce high-quality data. Higher-quality data would be more accurate, complete, timely, and usable (see chapter 2 for a discussion about primary care data quality). The goal was to compare the FHT to the hospital domain and identify specific attributes that could be transferred from one environment to the other. My goal was not to do a *full* CWA analysis of both systems, but rather to focus on tasks and processes that contributed to the codification and quality of data. Based on the results of chapter 5, I compared WDA, ConTA, StrA, and WCA modelling results. Since chapter 5 did not include a SOCA component, this aspect was not compared here.

### 6.3.1 Comparison Method

#### 6.3.1.1 Common Boundary and Operator Definition

Because I am aiming to compare two systems, I defined similarity as *two systems with common goals*. Some of these shared system goals should be reflected in each system's Functional Purpose (in the WDA phase). I used a set of common goals to ensure the logical integrity of the analysis by determining that all tasks Acted-On the similar work domain (Vicente, 1999) and therefore would be appropriate to compare.

Whereas tension between clinicians and managers previously has been expressed (de Lusignan et al., 2003), I used SMEs from both groups; incorporating clinician and management perspectives that exist within the larger ecosystem was important. However, in the final analysis of codification, the system operator was a manager archetype, and the use of clinical data for secondary purposes was the emphasis of the analysis. Choosing a manager as the operator did not preclude the manager from being a clinician who is delivering care. If the manager was a clinician delivering care, each analysis was described as function-based approached to processes and responsibilities.

The system boundary was restricted to work domain elements that were within the manager's influence and control. Functions and processes that the manager could oversee through policies and procedures were included in the scope, even if indirect.

#### 6.3.1.2 Work Domain Analysis Comparison

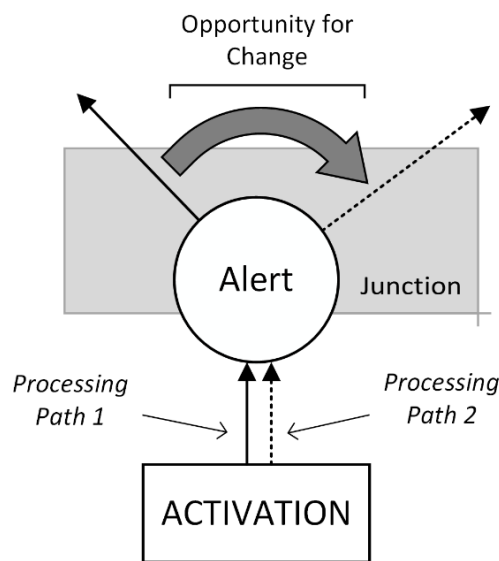
The AHs were compared visually with a Venn diagram (similar to the approach used by Burns, Bisantz, and Roth 2004). This comparison, broken down by abstraction layers, is a summary sheet which

allows the practitioner to highlight similarities and differences between the work domains (St-Maurice & Burns, 2015). This Venn diagram is accompanied by a comparison table that discusses and comments on each noted difference. This enabled a quick but useful comparison of both system domains.

### 6.3.1.3 Control Task Analysis Comparison

Conceptually, the notion of comparing similar tasks is to enable the possibility of overlaying both systems onto a common DL. As my analytical interest is to compare and contrast two systems, a common DL provides a visual opportunity to identify areas of interest. These areas include all points where processing paths diverge in the common DL. To track these areas, I have labelled these sites *junctions*. As a conceptual example, Figure 9 shows that a junction exists at the *Alert* knowledge state, following the *Activation* information-processing activity on the DL. Junctions could appear at any information processing or knowledge state in the DL. In theory, junctions highlight an opportunity to explore how one system could behave more like another.

Figure 9. Opportunity at a junction (St-Maurice & Burns, 2015).



I reviewed my DLs from chapter 5 (see DLs 1.1 to 2.1.2) and worked to apply my common DL approach. Where possible, common DLs were linked, and junctions were identified for further analysis.

### 6.3.1.4 Strategies Analysis Comparison

I abstracted the strategy characteristics into categories to enable comparisons by referring to each set of IFMs from chapter 5 (see IFM 2.1 and IFM 2.2). I noted that different users were performing the tasks in each environment and adopted the social organization comparison paradigm by noting which users were performing tasks and incorporating strategies in each domain.

### 6.3.1.5 Worker Competency Analysis Comparison

After each WCA was completed from chapter 5 (see SRK 1.1 and 2.1), I created a table summarizing my findings for each level of the taxonomy at each information-processing step. I abstracted a summary for each behaviour based on the WCA and identified which specific behaviours were unique to a SRK and which behaviours were common.

## 6.3.2 Results

### 6.3.2.1 Abstraction Hierarchies and Comparisons

A common AH representing patient treatment was developed as a subcomponent of both system domains (see AH 1.A and AH 1.B). AH 2.1 represents data codification at a FHT, and AH 3.1 represents data codification within a hospital. AH 1 is referenced by AH 2.1 and AH 3.1. A Venn diagram comparing the AHs is shown in Figure 10, and a summary of differences is provided in Table 1.

Figure 10. Venn diagram of differences in the WDAs.

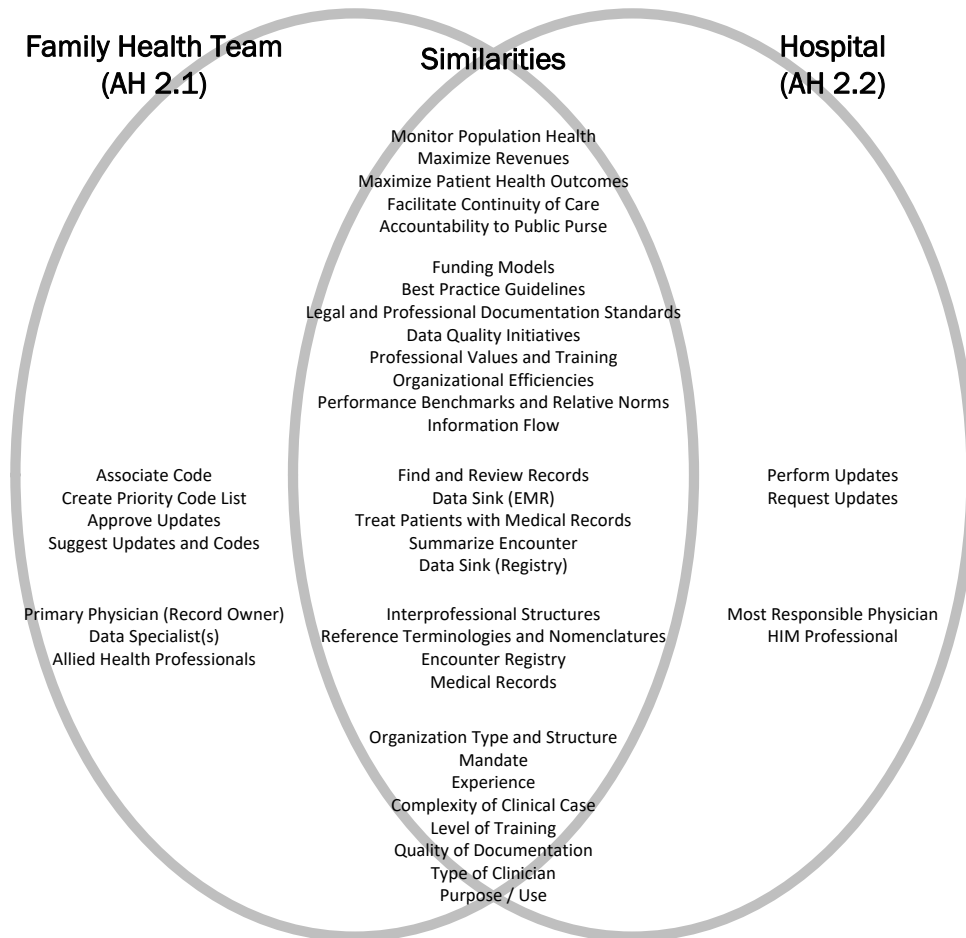




Table 1. Comparison of Work Domains.

Ref	Level	AH 2.1	AH 3.1	Modelled Differences
GF1	Generalized Function	Associate Code	-	In the FHT environment, there is a domain goal to have each clinician associate ICD-9 and ICD-10 codes with elements of their medical records. This is not a generalized function in the hospital environment.
GF2	Generalized Function	Create Priority Code List	-	As part of a data quality initiative at the FHT, there is a 'priority' list that is generated by a regional program to attempt and code the 'most important' diagnoses. This supports the associate code function and limits the scope and complexity for clinicians.
GF3 & GF4	Generalized Functions	Approve Updates & Suggest Updates and Codes	Perform Updates & Request Updates	In the FHT setting, physicians 'own' the medical record. Any changes must be approved after receiving a suggestion. Updates are not required and can be ignored. In acute care, the hospital 'owns' the medical record and forces physicians to complete updates during quality assurance processes. Physicians may lose hospital privileges if they do not perform updates.
PF2	Physical Function	Primary Physician (Record Owner)	Most Responsible Physician	The hospital record is owned by primary care physicians at the FHT, and owned by the organisation in the hospital. This has an impact on several generalized functions.
PF3	Physical Function	Data Specialist(s)	-	At the FHT there are data specialists who support physicians in the codification of data. They provide suggested updates to records. This role is rolled into the HIMP responsibility in the hospital setting.
PF5	Physical Function	Allied Health Professionals	HIMP	At the FHT, coding is performed by clinicians. In the hospital environment, this role is completed by HIMPs.

### 6.3.2.2 Decision Ladder Comparisons

At the FHT, the specific task under study was 'Record Encounter' (see DL 2.1). To complete this task, users had to review their own records and notes (or work from memory) and summarize their encounters into a specialized registry that collects data for government reporting purposes. These clinicians were not physicians, but all kept medical records within a physician's office, were part of a physician's team, and interacted directly with patients on a daily basis. To finish the task, users signed into a website, entered data into required fields, and submitted their data.

In the hospital environment, the specific task under study was ‘Abstract Encounter’ (see DL 2.2). To complete this task, HIMPs had to review patient records and abstract (e.g. summarize) the information. Specifically, they had to open a software application, enter data into required fields, and submit the data by clicking a button.

Two junctions were identified between types of users in the primary care CLs. They are shown graphically Figure 11 and Figure 12.

Figure 11. Junction at System State in ‘Record Encounter’ task.

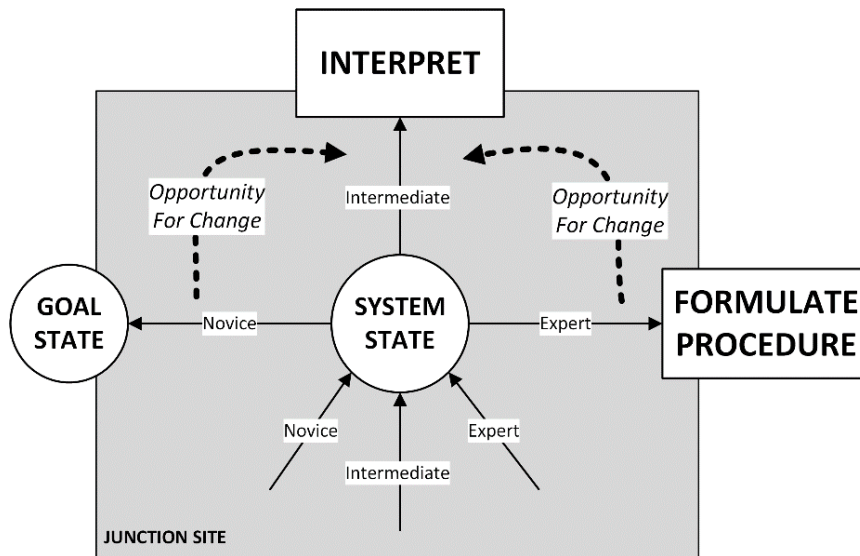
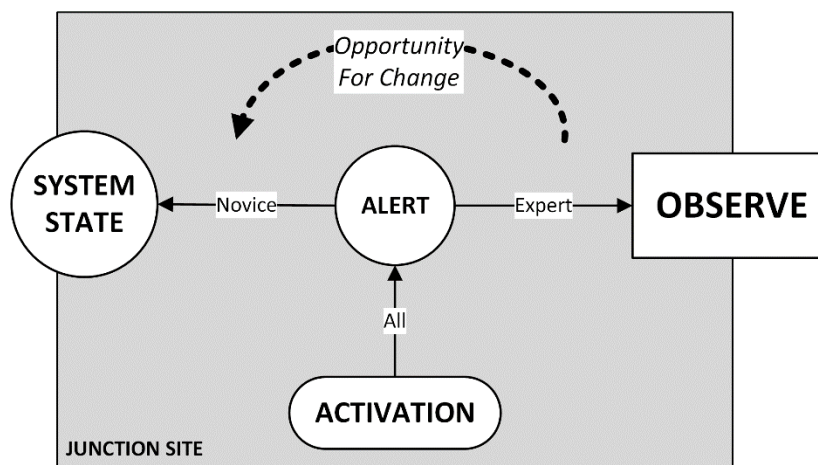


Figure 12. Junction at Alert State in ‘Record Encounter’ task.



### 6.3.2.3 Information Flow Map Comparisons

Table 2 provides an overview of my results by linking strategy characteristics to triggers (e.g. why this strategy was used), strategy adopters (e.g. users using the strategy), and the specific strategies from the StrA that link to the characteristic.

Table 2. Comparison of Strategy Characteristics.

Strategy Characteristics	IFM 1.1			IFM 2.1		
	Strat #	Performed By	Triggers	Strat #	Performed By	Triggers
Same-Day Entry	1, 2, 3	Clinician	Policy	-	-	-
Bulk Entry	4, 5, 6	Clinician	Workload. Experience.	1, 2, 3	HIMP	Policy
Delegation to Others	7	Clinician & Admin Support	Technical abilities. Experience. Workload.	-	-	-
Paper Notes to Support Computer Entry	6, 7	Clinician	Technical abilities. Experience.	2, 3	HIMP	Experience
Complete on Paper and Transcribe	7	Admin Support	Workload of clinicians. Clinic processes.	3	HIMP	Technical abilities. Historic processes.

### 6.3.2.4 SRK Taxonomy Comparisons

After developing an SRK taxonomy for each information processing and knowledge state from ConTA (See SRK 2.1 and SRK 2.2), I created a chart summarizing my findings for each level of the taxonomy at each information-processing step. As shown in Table 3, I abstracted a summary for each behaviour based on the WCA and identified which specific behaviours were unique to a SRK and which behaviours were common.

Table 3. Comparison of Competencies.

Taxonomy		Information Processing Step					
		Observe	Identify	Evaluate	Interpret	Define Task	Formulate Procedure
Skill-Based Behaviour	<i>Record Encounter (SRK 2.1)</i>	-	Create Patient List	Time Mgmt	Balance Priorities, Value, and Time	-	-
	<i>Abstract Encounter (SRK 2.2)</i>	-	Look For Content	Identify Field Values	Look For Specific Content in Records	-	-
	<i>Common Skill Behaviour</i>	Find Records	-	-	-	Launch Software	Place Values Into Fields
Rule-Based Behaviour	<i>Record Encounter (SRK 2.1)</i>	-	Heuristic Validation	Policies and Org Goals	Policies and Org Goals	-	-
	<i>Abstract Encounter (SRK 2.2)</i>	-	Coding Reqs	Project Criteria	Coding Reqs	-	-
	<i>Common Rule Behaviour</i>	Follow Policies & Proc's	-	-	-	Org Reqs	Complete Mandatory Fields
Knowledge-Based Behaviour	<i>Record Encounter (SRK 2.1)</i>	Recall Patient Details	Recall Patient List	Current Schedule and Priorities	Benchmark Experience	ID Registry System	Recall Encounter Details
	<i>Abstract Encounter (SRK 2.2)</i>	Record Layout & Medical Terms	Req'd Docs	Medical Terms and Physiology	Req'd Docs	Special Codes	Medical Terms & ICD-10
	<i>Common Knowledge Behaviour</i>	-	-	-	-	-	-

### 6.3.3 Discussion

#### 6.3.3.1 Work Domain Comparisons

WDA comparisons should allow a practitioner to characterize system differences in a qualitative manner and to explain any differences in system performance or efficiency through differences in the work domain. Performance differences could be explained by variations in the domains at any level of the AHs, and could identify specific constraints that impact either system positively or negatively. Conceptually, WDA comparisons allow for the identification of successful elements within each system and could offer insight regarding how differing elements might impact other domains.

In this example, comparing both sets of AHs showed differences in the Generalized Functions, Abstract Function linkages, and Physical Functions. For example, some of the differences in the Generalized Functions represented different organizational structures. In the case of the hospital, doctors were told to correct or add material to records during quality assurance by the hospital. In the FHT, records were owned by each doctor, and they were in control of all changes. Therefore, doctors were given suggestions about corrections or additions, and were politely asked to consider changes. In both systems, records and codification occurred, but ownership factor impacted how the task was prioritized and valued differently by decision makers. In one environment, coding was required by the organization, whereas in the other, coding was a negotiable action.

Another functional difference between the two systems is that FHT clinicians were asked to summarize encounters in a registry and add codes to their medical records, whereas in the hospital, a team of HIMPs was dedicated to this task. Since the type of employee performing the task was different, different levels of training and available time existed. Based on these role definitions, FHT clinicians were not expected to provide rich data using ICD-9 or ICD-10 codes, whereas this was required in hospitals when the task was handled by HIMPs.

In terms of values in the abstract function of the AH, the act of coding data and inputting codes into medical records is not linked to the 'Professional Values and Training'. This was an intentional model feature and represents observations and feelings about coding data, as expressed by some SMEs. This modelling choice can be interpreted as an opportunity: in the future, it might be possible to engage FHT clinicians into the codification process by helping them understand the value of coding and ensuring they are able to incorporate these values into their professional practice. In fact, several community-driven data initiatives discussed during interviews aim to pursue this goal. In hospitals, since clinicians are not asked to code data, incorporating data abstraction into professional practice would not be a useful goal. However, understanding how HIMPs are trained and how they have coding practices incorporated into their professional practice could be useful for the FHT.

In addition to differences between the models, similarities were noted. Several concepts from the Abstract Functions and Functional Purposes overlapped. Overlaps are a partial artefact from modelling both systems in tandem; as concepts were revealed from interviews in either domain, they could be realistically linked and modelled in both domains. This is similar to the development of common concepts from two systems described by Kirschenbaum, Traflet, and Pratt (2007) when using C2TA; arguably, the overall analysis of each domain is enriched and more applicable through the analysis of two similar systems.

### 6.3.3.2 Control Task Comparisons

The purpose of the comparison of DLs was to identify junctions. Each junction represents an opportunity for deeper analysis, comparison, and potential design interventions. Junctions invite a practitioner to investigate and ask deeper questions about the relationships between the ConTA and WDA phases of analysis. Is the reason for a junction in the DL reflected in the AH? In some situations, one system's path in the common DL may be preferable to the other, as it indicates a more efficient processing path.

During my case study, I selected two similar tasks that I believed would be comparable. At first glance, it appeared that the tasks in both systems involved, generally, very similar processes. However, when attempting to put the DLs from each system into a common DL, it became clear that the tasks, when placed into their respective ecosystems, were contextually different. Specifically, the tasks differed when they reached the information-processing mechanisms at the higher *Interpret* and *Evaluate* information-processing levels. In the case of the FHT task, users processed time management and work priorities at the higher levels of the control task, since these users needed to balance clinical assessments with coding responsibilities; the control task was ultimately about *time management*. In the case of the hospital task, users processed information about special coding cases and special requirements at higher levels. This was based on specific diagnoses, and users had to identify ways to complete the task as completely as possible given a large set of coding requirements; the control task ultimately was about accurately *finding codes and meeting business requirements*. Without producing a common DL, the differences between the tasks became apparent and provided informative insights about each domain.

Instead of placing the two control tasks onto a common DL, I chose to put the ConTA DLs for all types of users into a single common DL for each task. Whereas the hospital DLs did not offer any interesting insight from this approach, in the case of the FHT tasks, two interesting junctions were identified. These junctions are shown in Figure 11 and Figure 12. They demonstrate the concept of comparing DLs.

In the first junction (Figure 11), it is visually obvious that intermediate users were more engaged in their data codification processes and spent more time evaluating the value of their work and deciding how to manage their time and tradeoffs. In contrast, novice users did not consider the value of their task and blindly completed the work, while expert users understood their work value heuristic and no longer engaged in an evaluation, and had predetermined priorities regarding their data. By placing all these paths onto a single DL for comparison, it appears that intermediate users may engage in the deepest thinking about their data quality. From a design perspective, the DLs suggest that novice users should be interrupted when they have reached the *Goal State* knowledge state, and expert users could benefit from support to encourage deeper thinking before they begin to engage in the *Formulate Procedure* information-processing activity.

The second junction (Figure 12) exists because novice users did not enter their data in bulk and therefore did not go through an *Observation* knowledge state to understand the work that must be done. In contrast, expert users tended to bulk enter data and required the additional process of reviewing a sheet or report to understand the work that had to be completed. Based on this knowledge, it would be advantageous to encourage expert users to work like novice users (e.g. enter data on the same day), as this would reduce the cognitive load and complexity of the task.

### 6.3.3.3 Strategies Analysis Comparisons

My analysis showed several strategy characteristics that were unique and common. Primary care strategies were schedule oriented as a result of codification being a secondary task to patient care, and were artefacts of time management requirements and workload decision making. Hospital strategies aimed to improve task efficiency and enhance accuracy. Hospital artefacts were driven by optimization needs and a desire to avoid quality assurance loopbacks and being asked to redo work.

Bulk data entry was a strategy characteristic from both environments. In the hospital, bulk entry of data into the abstracting system is policy driven and a function of process. Data entry is delayed by approximately six weeks, and a consistent lag exists between real-time events and reporting capabilities. At the FHT, bulk entry by clinicians is a strategy to cope with workload but is against policy and is also the least efficient approach because of additional cognitive requirements (e.g. checking which patients have already been inputted and referring back to documentation instead of reporting from memory). The most desirable strategy in each domain was different, although interviews with FHT managers indicated that only half of the encounters entered into the system respected same-day entry policy.

At the FHT, clinicians performed data entry. This is arguably more efficient: clinicians can use their memory and cognition to record information, whereas HIMPs need to review notes and familiarize themselves with each individual episode of care. The benefit, however, of the HIMP is a built-in quality

assurance process that occurs through the secondary interpretation of events and specialized training to ensure standard decision making. By assigning the task to a non-clinician who does not know the patient, there is a verification of information and the removal of bias. The execution of the task by different parties is a key difference in this context and results in data quality differences.

In the hospital abstracting task, experts use paper to take notes before using a computerized interface to input data. Advanced abstracters suggested that anyone with experience would use this approach. This suggests that room is available to provide the computerized interface to support the task and avoid the need to work with paper to take notes during the task.

#### 6.3.3.4 Worker Competency Comparisons

When I first chose the ‘Record Encounter’ and ‘Abstract Encounter’ tasks for comparison, the tasks appeared similar: users were signing into computerized systems, entering data into fields, and submitting the data into a database. These tasks appeared similar for analysis because they are both common skill-based behaviours at the lower cognitive levels in the taxonomy.

However, these tasks were quite different because they feature differences in both the rule-based and knowledge-based behaviours. Rule-based behaviour was similar at lower ends of the cognitive tasks (e.g. following organizational policies and mandates) and following mandatory field rules in the user interface. At higher levels, cognitive tasks required different rule-based behaviours (e.g. interpreting policies vs. reading coding project definitions). Knowledge-based behaviours were completely different between the two SRKs, which clarified that clinicians and HIMPs do not have the same training and work competencies to perform their respective tasks. Despite similar ecologies, the tasks require different types of training and cognitive supports. Whereas both systems generate coded data through a similar interface (e.g. boxes and drop downs), the skill overhead was quite different.

#### 6.3.3.5 Knowledge Transfer

One of the use cases for comparing systems was to understand the causes of different performance characteristics in similar systems, and to facilitate the transfer of knowledge and ideas between systems. The identification of differences can be helpful, but so could the identification of similarities. Understanding how similar complex domains achieves a function can show a generalization of process and context (when the approach is similar) or help generate transferable ideas (when the approach is different). Differences can generate ideas for change or improvement between systems.

In my codification example, record ownership was different in each system. It became clear that the issue of record ownership has a significant effect on the types of processes and responsibilities within the domain. In the hospital, records were owned by the institution. This allowed the hospital to demand



changes to records when they were required for quality assurance purposes. The hospital developed central policies and procedures and expertise, and is able to ensure consistency throughout all records in its possession. At the FHT, individual physicians own the medical record, and they are able to adopt record policies, or not, based on their own discretion. This impacts the generalized functions around chart codification. Assuming that the goal of the complex domains is to codify data, the hospital would be wise to maintain ownership of their records, and the FHT could consider following suit with a new records management framework and mandate; such a change could significantly simplify and improve data codification and policy enforcement.

In comparing the ConTA DLs, in the FHT environment, doctors codify their own data in collaboration with team members. They have access to data specialists and other clinicians. As a process change, could hospitals consider loading some of their codification tasks to clinicians? Interestingly, the domain comparisons offer some insight as to what might happen. Offloading the task completely would result in clinicians worrying about time management, instead of proper and accurate data codification. However, the hospital could consider a two-tier model where clinicians code aspects of the record and are assessed afterward by HIMPs. A two-tiered approach would provide the benefits of the FHT processes (e.g. data that is closer to real time) and the benefits of the hospital (e.g. improved quality assurance and accuracy). Presently, each domain has made different tradeoffs: in hospitals, where HIMPs code data, up to a six-week delay occurs in providing coded data, but the task is guaranteed to be done. At the FHT, the delay could be zero days from associated diagnostic codes to data, but no guarantee is given that it will be done.

### **6.3.4 Future Work and Limitations**

The approach to drawing comparisons through the use of CWA is novel and systematically revealed insights into domain functioning and opportunities that could easily be missed by analysing only one domain. As a proof of concept, my analysis demonstrated that the approach is viable. Future work can begin exploring this approach to transfer information from increasingly disparate systems. For example, could comparisons with CWA help transfer knowledge from aviation to healthcare? This is explored by Kapur et al. (2016) and Clay-Williams and Colligan (2015). An additional case study could analyse a simple task from both systems, such as booking an appointment. My comparison approach to superimpose those tasks within a rich work domain and the outputs of the models could provide valuable insights to improve the process within healthcare. This would demonstrate that increasingly different systems could be successfully compared.

One obvious limitation to my comparison approach is that the CWA models need to be developed with the concept of comparison in mind; the intent to compare has to be ‘cooked into’ the analysis. For

this reason, it is unlikely that off-the-shelf CWA models in the literature could easily be compared. Although a comparison from models in the literature was performed by Burns et al. (2004), the number of published models CWA models would need to grow astronomically before comparing models from the literatures becomes a regular opportunity for investigators.

## 6.4 Suggestions and Recommendations

My comparison of CWA models is a contribution to the literature. To my knowledge, I have presented the first full comparison in this manuscript, and future readers may be interested in some suggestions regarding the future use and adoption of this method.

First, I want to highlight that CWA is an excellent framework for conducting systems analysis. The toolkit is a highly effective set of templates to help communicate findings, and gives modellers a list of questions and inquiries during the analysis of a complex socio-technical system. CWA made me consider cognition in ways I would not have initially considered, and my analysis is richer as a result. Presently, the limitation of CWA is that it is difficult to learn and digest its nuances. I found that CWA is like jumping out of an airplane; you should avoid doing it for the first time without someone accompanying you along the way. Of course, you can do it alone and you'll get to the finish line one way or another, but the quality of the output will vary considerably. With this in mind, I found doing several CWAs simultaneously and seeking comparisons was an excellent approach to appreciating the framework alongside some good coaching from my advisor. Analyzing similar systems was an effective way to quickly deepen my understanding of the approach. Future novice modellers should consider adopting a similar strategy to learn CWA and appreciate its nuances.

More experienced modellers might have more confidence and insight to address the question of similarity between systems. In my case, I may have been a bit aggressive when ensuring that the functional purposes and abstract functions were not merely similar, but identical. I was likely too restrictive in letting my models diverge because I wanted to ensure I could draw a comparison. In another context, perhaps with less on the line, there might be more room to 'play' with the idea of similarity in a less rigid way. It is unclear how this might impact the overall analysis; I would be curious to learn if loosening similarity constraints would pay dividends in insight without sacrificing model consistency and meaning. There is still a lot of room to play with the idea of similarity, and this aspect of my work is not intended to be prescriptive, but flexible. Future modellers should continue to experiment with this idea.

Lastly, the idea of identifying similar tasks was rewarding. In this case, it was very interesting that my CWA described how the abstraction and summary tasks were, in fact, **not** very similar. This was a highlight of my exploration of this idea. Future modellers should not be afraid to attempt comparisons,

and withdraw when it appears there are incompatibilities. The process of assuming a comparison and breaking that assumption is an interesting journey in and of itself.

## 6.5 Chapter Summary

### 6.5.1 Key Findings

**Comparison of CWA models:** I have described an approach to compare complex socio-technical systems through the use of CWA. While several reasons can be seen for conducting a comparison, I have demonstrated that understanding complex differences between two systems can be informative for identifying potential performance differences and transferring knowledge. My approach was used to identify different team structures, training needs, and cognitive tradeoffs that vary between two similar systems. My approach was able to articulate specific differences between the ecologies and output interesting insights.

**Similarities, Differences, and Exchanges:** Understanding how similar complex domains achieve a common function can show a generalization of process and context (when the approach is similar) or help generate transferable ideas (when the approach is different). Differences can generate ideas for change or improvement between systems.

**Don't judge a task by its software:** During my comparison of two similar tasks, I reinforced the old adage from systems design that underneath a user interface and similar 'strokes and clicks', a rich context exists that influences and distinguishes the actions of users: look beyond the user interface! In this case, seemingly similar tasks took place in different contexts, with different priorities and different required competencies.

### 6.5.2 Connections to Case Study and Research Questions

I have three research questions:

1. In primary care, how are individual users influenced by their environment to input high-quality data?
2. What techniques could be used to design systems that persuade users to enter higher-quality data?
3. Is it possible to improve data quality in primary care by persuading users with the user interface?

This chapter continued my review of the codification task environment by drawing comparisons. The process helped me to understand the unique factors that influence primary care users. Primary care users are uniquely influenced by their environment. Some examples include:

- Primary care users have a primary responsibility to see patients. Data coding is a secondary role. Balancing these responsibilities requires time management abilities and, importantly, time.

- The codification of data (through secondary entry in a summary screen or adding codes to records, for example) is not automatically a component of a clinician's professional values or training.
- Record ownership in primary care forces FHTs to make suggestions to physicians and recommend potential codes. Codification is influenced by physician discretion.
- Performing codification tasks is not complex, but completing the task in primary care becomes cognitively complex because users question the value of the work. Expert users develop pre-determined ideas about the value of the task, and if their notions are negative, they will prioritize other activities. Thus, they are influenced by their previous experiences.

### **6.5.3 Connections to Other Chapters**

This chapter built onto the CWA models developed in chapter 5. Some of the comparison insights are used for developing designs in chapter 8.

## **Part C**

### **Measurement, Design, and Improvement**

This part of my dissertation features three chapters that focus on improving data quality within a primary care ecosystem. This was performed during a case study by developing a data quality measurement approach and by analysing historic data to better understand data quality tradeoffs. The results of this work are presented in chapter 7. Based on the results of my models from Part B and information gathered from the case study, I develop several concepts with PD in chapter 8. As a capstone to my dissertation, I show in chapter 9, via a field study, that data quality can be improved by using persuasive techniques.

My manuscript titled ‘An Exploratory Case Study to Understand Primary Care Users and Their Data Quality Tradeoffs’ has been accepted for publication in the *Journal of Data and Information Quality* and will be published in April 2017. A second manuscript that I have produced, titled ‘Using Persuasive Design to Improve Data-entry Behaviour in Primary Care’, was submitted to the *International Journal of Medical Informatics (IJMI)* and is under review at the time of this writing.



## Chapter 7

### Measurement of Data Quality in Primary Care

During my modelling work in the first part of my dissertation, I described the primary care codification ecosystem and compared it to the data codification ecosystem of a community hospital and gained several insights. Specifically, I looked at the ‘Summarize Encounter’ control task and examined the various strategies used by clinicians. Then, I outlined various user competencies required to perform the summarization task as part of clinical practice.

In this chapter I go into greater detail about a specific example of a system. To support additional work and define design priorities, in this chapter I describe my approach to developing several data quality measures for the reporting tool of a FHT (which was previously modelled with CWA). I also do a detailed logistic regression of several metrics to better understand how users manage data quality in primary care and what ecological aspects might influence them. This information, combined with my modelling work from the previous chapters, is used to develop several data-quality-enhancing design concepts for primary care in the chapter 8.

#### 7.1 An Exploratory Case Study to Understand Primary Care Users and Their Data Quality Tradeoffs<sup>8</sup>

**St-Maurice, J, & Burns, C. M. (in press). An Exploratory Case Study to Understand Primary Care Users and Their Data Quality. *Journal of Data and Information Quality*.**

**Abstract:** Primary care data is an important part of the evolving healthcare ecosystem. Generally, users in primary care are expected to provide excellent patient care and record high-quality data. In practice, users must balance sets of priorities regarding care and data. The goal of this study was to understand data quality tradeoffs between timeliness, validity, completeness and use among primary care users. As a case

---

<sup>8</sup> The content of section 2.4, Data Quality, was published as part of the accepted manuscript in the *Journal of Data and Information Quality* (St-Maurice & Burns, in press). It was published as part of the paper’s Introduction. To provide better structure and readability to the dissertation, most of the paper’s introduction was shown in chapter 2, while the remainder of the manuscript is shown here.

study, data quality measures and metrics are developed through a focus group session with managers. After calculating and extracting measurements of data quality from six years of historic data, each measure was modelled with logit binomial regression to show correlations, characterize tradeoffs and investigate data quality interactions. Measures and correlations for completeness, use and timeliness were calculated for 196,967 patient encounters. Based on the analysis, there was a positive relationship between validity and completeness, and a negative relationship between timeliness and use. Use of data and reductions in entry delay were positively associated with completeness and validity. Our results suggest that if users are not provided with sufficient time to record data as part of their regular workflow, they will prioritize spending available time with patients. As a measurement of a primary care systems effectiveness, the negative correlation between use and timeliness points to a self-reinforcing relationship that provides users with little external value. In the future, additional data can be generated from comparable organizations to test several new hypotheses about primary care users.

### **7.1.1 Study Purpose**

In chapter 2, I have described some anecdotal evidence of interactions between data quality dimensions in primary care, and have articulated an impetus to better understand these interactions. For example, if a user invests several hours entering high-quality data that is a perfect representation of a clinical encounter, users would have less time for patient consultations and care (St-Maurice & Burns, 2014). Otherwise, users increase the amount of time per encounter (Brown et al., 2012). More likely, users are making sacrifices, taking shortcuts and engaging in tradeoffs regarding their data to manage their workload and priorities. The objective of this study is to better understand primary care users and their data quality tradeoffs by measuring and correlating different types of data quality. This is done through a case study.

#### **7.1.1.1 Development of Measures and Metrics**

The first part of our study involves defining contextual data quality measures in a primary care environment. We use interviews with experts to ensure our measures are reflective of primary care data quality issues. The purpose of this part of the study is to create the best possible measures of data quality using available metadata.

#### **7.1.1.2 Analysis of Historic Data**

The second part of our study aims to understand primary care users, their tradeoffs, and the impact of tradeoffs on data quality. As a hypothesis, we believe users balance time and priorities to achieve several different goals, such as patient care and data entry. We believe users engage in tradeoffs



and that there is an impact on data quality. For example, as users choose to delay their input of data (e.g. timeliness), we would expect an impact on data accuracy and completeness.

To study user choices, data quality measures and metrics are analysed in regression models. The purpose of this analysis is to numerically characterize the relationship between different data quality dimensions.

### 7.1.1.3 Scope

The scope of our work is to understand clinicians in primary care as users of data and data systems. These clinicians include physicians and allied health professionals (AHPs), such as pharmacists, dietitians, mental health workers and nurses. These users have a dual role of providing patient care, documenting clinical findings and generating data. We also consider managers who use the data generated by clinicians to inform decision making. Our scope does not include clerical staff and does not include patient-generated data, or patients as users of information. As well, we are not including data from recording devices that may generate faulty information and observations.

## 7.1.2 Methods

### 7.1.2.1 Case Study

Case studies are widely used in data quality research (Madnick et al., 2009). Case studies are empirical methods that use a combination of qualitative and quantitative evidence to examine phenomena in a real-world context. The in-depth inquiry of a single instance can lead to a deeper understanding of the subject, generate useful information and generate hypotheses (Flyvbjerg, 2006). Exploratory case studies are intended to be theory generating.

For our case study, we worked with a FHT in Ontario. These organizations employ AHPs who work in family physician's practices. AHPs enable comprehensive health services for patients in each clinical practice by providing services such as counselling, health promotion or lifestyle coaching. The AHPs working for these organizations are required to create normal clinical documentation, but must also record their activities in a structured format to help the FHT generate quarterly reports. As a public organization, these reports are required by the Government to rationalize and maintain funding. As an accountability measure for the organization, these reports measure the number of patients that have been seen by the organization's AHPs and provide tabular data regarding presenting problems, the number of patients seen for follow-up, the referring physician type and basic patient demographics. AHPs collect the information for reporting purposes. The information is not used for billing purposes to an insurer.

Creating these quarterly reports from an Electronic Medical Record (EMR) is challenging for larger FHTs. These larger organizations are comprised of several individual clinics. As a clerical task,

consolidating data from 20 individual EMRs to compile a single report each quarter is time consuming and error-prone. As well, the EMR is not able to capture some of the required information. For example, the EMR is not able to report whether a scheduled patient visit was an ‘initial encounter’ or a ‘follow up visit’, and determining this information may require the manual review of each provider’s schedule. Though retrieving this data may seem like a trivial problem to solve from a technical perspective, the EMR is proprietary and without support from the EMR vendor to gather this data, there is no practical way to generate this information for thousands of patient visits every quarter. For a small FHT with a handful of AHPs, this is a manageable problem with clerical staff and clerical processes. For large FHTs with hundreds of AHPs, this problem can become very complex and time consuming.

As an alternative to manual clerical processes, some larger FHTs have opted to use a web-based tool to capture statistics outside the EMR system. This gives FHTs the flexibility to capture the required data fields and to generate data for fluid reporting requirements. Instead of hiring clerical staff to generate these data points, participating FHTs ask clinicians to record clinical documentation within the EMR and report additional data for reporting purposes into the web-based tool. Functionally, this tool is best described as a web-based survey tool that is specially customized to capture statistical data required for Government reports. A screenshot is available in Appendix G.

Before the adoption of this web-based system, managers reported spending days working with clerical staff tabulating reports with spreadsheets and paper notes. The web-based reporting tool provides a significant process improvement for FHTs opting to use it. Another benefit to capturing information in this separate tool is that it can be used by the FHT’s management team for secondary purposes, such as supporting decision making, assessing performance and catering programming to population needs. The relationship between primary care users, managers and this reporting tool is very similar to situations and challenges described in the literature. Some of the issues parallel how clinicians input structured data into EMRs.

#### 7.1.2.2 Data Source

The source of data for our case study is the web-based reporting tool’s database. Whereas the data is not from a primary care EMR, it is manufactured in tandem to clinical documentation, and collects data from users in the same ecosystem. Unlike an EMR, the database is an excellent source of structured data, can be analysed on a large scale and does not suffer from the nuances of free text. This is similar to other studies that measure data quality through external sources of data (Brouwer et al., 2006).

This data source is interesting because it suffers from a variety of data quality challenges that mimic the challenges observed in medical registries. For example, prior to the study managers expressed concerns about the accuracy of the data in the web-based tool, and also suggested there were several

concerns about the timeliness of information from users. These data quality concerns (e.g. late data entry or inaccurate entries) were thought to have a direct impact on the quality of the reports issued to the Government. As an accountability measure, the quality of these reports is important to the organization's reputation and funding. The data in the system is important, needs to be reliable and therefore needs to meet high data quality standards.

By working with a FHT and the data from its reporting database, we created several models to represent data quality relationships and tradeoffs. Our case study included qualitative methods to define data quality measure through interviews with managers at the FHT, and the use of logistic regression to analyse quantitative data.

### 7.1.2.3 Ethics

A research protocol was submitted to the University of Waterloo's research ethics office, the FHT's privacy officer and the FHT's Board of Directors. After approval by all parties, the study was started by interviewing experts and discussing data quality with a focus group.

### 7.1.2.4 Qualitative Development of Data Quality Measures

We used a qualitative approach to define data quality measures for the FHT's database. This involved a four-step process that included an exploratory meeting, a discussion with the organization's data analysts, a focus group about data quality and an identification of feasible measures.

#### 7.1.2.4.1 Step 1 – Exploratory Meeting with a Subject Matter Expert.

After receiving ethical clearance, we held a preliminary meeting with the Operations Director at the FHT. This contact is responsible for using and interpreting the data in the reporting tool. Having been in this role for over 5 years, we considered this Director a subject matter expert (SME) regarding the web-based tool's data.

Prior to our meeting, the Director was given literature about data quality to review as a primer to the discussion. In addition, we brought a list of potential data quality measures per Wang et al. (1996) to stimulate discussion about data quality within the reporting tool. Over the course of the meeting, several possible data quality dimensions and formulae were explored. This formed an initial 'wish list' of data quality measures that aligned with the SME's contextual data needs.

#### 7.1.2.4.2 Step 2 – Exploratory Measures.

Following our discussion with the SME, we met with the organization's information management team to discuss possible data quality measurement approaches. As an exploratory exercise, analysts created several potential data quality measures and graphs. These measures were created by programming

queries directly into the underlying SQL database. The results of this exploratory exercise were exported as sample data quality measures and supported with sample data. Whereas the SME was not constrained by technical issues and metadata limitations, the analysts were forced to place an emphasis on data quality dimensions that were practically measurable and programmable.

#### 7.1.2.4.3 Step 3 – Management Focus Group.

We facilitated a focus group with managers by using the ‘wish list’ from the SME (step 1) and the sample data quality reports from the analysts (step 2). The focus group was conducted over the course of 90 minutes with 5 middle-managers. The discussion was semi-structured. We prepared questions (shown in Table 4) ahead of time to guide the conversation.

Notes were created over the course of the discussion. These notes were distributed to all participants for comment immediately following the session and all contributors were asked to share any additional feedback. All feedback was documented and grouped thematically. To structure the results, we created a table summarizing results by data quality dimension, organizational concerns, and measurement discussion.

Table 4. Open Ended Questions Used with Focus Group.

Question	Open Ended Question
1	There are several types of data quality dimensions. Examples include accuracy, timeliness, granularity, usefulness, value and consistency. Which dimensions are the most relevant to the data stored by the reporting tool?
2	Based on the top 5 dimensions we’ve discussed [from discussion in #1, show notes to group], what are ways you feel you could calculate or measure this type of data quality? What would the ‘formula’ be?
3	Before this discussion, I calculated several sample data quality measures for [this month last year] with the help of the Information Management team. Here is a copy of those calculated measures [hand out sample measurements and calculations created earlier]. 3a. Which measures do you think are the best representation of data quality in the reporting tool? 3b. Which measures are not a good representation of actual data quality? 3c. Do you have any suggestions for adjusting the calculations? Should other variables be taken into consideration? 3d. How would you improve data quality in the reporting tool?

#### 7.1.2.4.4 Step 4 – Feasibility Analysis.

We mapped the data quality measurements created by the analysts to the results of the focus group discussion. The purpose of this step was to filter measures that were feasible (e.g. calculable) and

measures that were not feasible. To bring structure to these results, we added a feasibility column to our summary table and noted whether each requested data quality measure was feasible or not given the dataset and the available data attributes.

#### 7.1.2.5 Development of Data Quality Variables

After identifying feasible data quality measures with the focus group, each measure needed to be carefully articulated as a quantitative measure. Through collaborations with the data analysts, each feasible measure was defined and implemented through SQL. Effective data quality metrics can be defined at the data item, attribute, record, or database level (Shankaranarayanan & Cai, 2006). Based on previous studies (see Blake and Mangiameli [2011]) the data analysts identified each data quality dimension as a Boolean value for each record.

As part of the development of measures, some concepts had to be further refined after the focus group exercise. For example, measuring usefulness and usability through the generation of reports required refinement and contextualization; when should an entry be flagged as used in a report? In some cases, it was possible to measure a data quality dimension in two ways. For example, the timeliness of data could be measured as same-day Boolean flag (e.g. was the data entered on the day of the appointment?) or could be measured as a delay in days. Both measures were kept for modelling purposes as it was unclear which measure would be most appropriate or effective.

Defining each measure with a meaningful name required careful consideration and research. For example, the focus group provided a measurement approach that could, according to participants, be useful to measure accuracy. However, the measure was better described as validity based on previous studies, such as that of Bray and Parkin (2009).

#### 7.1.2.6 Measures from Historical Data

After finalizing the individual data quality measures and their calculation, the data analysts from the FHT exported data dating back to 2008. Per the study's ethical approvals, the dataset was de-identified and all information pertaining to patients or users was removed or masked to ensure that individuals could not be associated with the information provided for analysis. For each encounter that was recorded in the system, a set of data quality measures was provided with relevant data attributes, such as the entry's username (masked), the month and year of the entry, and the user's program area (e.g. health promotion, nursing, diabetes care, etc.).

#### 7.1.2.7 Data Adjustments

Based on the data quality measurements defined by the organization's management team, some of the raw data collected needed to be adjusted. In some cases the time delay for entering records was either

impossible (smaller than 0) or seen as highly unlikely. The definition of unlikely delay was based on discussions with the focus group, and defined as delays greater than 120 days. Unlikely delay was identified as an input error and not a true representation of the amount of time it took to enter data into the system. In these instances we kept the entry for analysis in our models but recorded a NULL for the time value, and set the validity flag to false.

#### 7.1.2.8 Simple Correlations

To understand some of the user tradeoffs, we performed a simple inter-variable correlation between each data quality dimension. We analysed the dataset using the R statistical software and the *Hmisc* package (Harrell Jr, 2016) and combined the results into a single matrix. We calculated Pearson correlations and performed significance tests using the *rcorr* function.

#### 7.1.2.9 Model Testing

The dataset included data quality measures for individual encounters and multiple observations of data quality for several users and several periods of time. Whereas simple correlations can be informative, they do not take these variations into consideration and the correlations can be significantly skewed. We anticipated variance in users and variances associated with the user's program area. To better understand tradeoffs given the structure of the dataset, a multilevel model was required to compensate for multiple observations.

Multi-level models were built based on the data quality measures. The repeated observations were accounted for by using a multilevel mixed-effects model and a normal correlation structure. We included a random intercept term to control for clustering of variance on individuals (110 users), program area (7 different areas), years (4 years) and months (12 months) over repeated measures.

The models were analysed using the R statistical software and the *lme4* package (Bates et al., 2016). Each model was developed using forward stepwise regression; each model started with no variables and we tested the addition of additional variables by comparing ANOVA F-Scores and Akaike Information Criterion (AIC) values. If additional variables and interactions were not statistically significant, they were not included in the final model. All models were developed and reduced with this approach into their final form.

Not all measures were appropriate to model as dependent variables. For example, validity was measured at over 97% 'valid' and could not be correctly modelled using regression.

### 7.1.3 Results

#### 7.1.3.1 Qualitative Results

Over the course of the focus group session with the FHT, six data quality dimensions were discussed. A summary table with the data quality dimension, the organizational concerns with the data (Question 1 from Table 4) and suggested measurements (Question 2 and 3 from Table 4) are shown in Appendix F. The feasibility of each idea, as discussed with the FHT’s data analysts, is also shown.

#### 7.1.3.2 Data Quality Variables

The definitions for each data quality variable is provided with a series of logical rules in Table 5. These rules represent the underlying logic programmed into SQL queries and use terminology that flows with business processes and interface logic. To clarify some of the business logic and terminology, a screenshot of the entry form is available in the Appendix G. In order to test models, it was important that all data quality variables be calculable from historic records and to ensure the variables did not require system changes or new data points to measure data quality.

Table 5. Data Quality Measures.

<b>Dimension</b>	<b>Definition</b>	<b>Implementation</b>	<b>Variable</b>
Timeliness	Timeliness is a measure of whether the data is out of date and available for use when needed.	<i>Measure 1:</i> Measure the number of days between the encounter and the day the data was entered into the system. Can be rounded up to the closest week. Cannot be smaller than 0.	<i>EntryDelayDays</i> And <i>EntryDelayWks</i>
		<i>Measure 2:</i> Business rules state that users need to enter data on the same day as the encounter. This measure of timelines is a flag regarding whether or not that business objective was achieved.	<i>SameDay_YN</i> – True or False (0/1)
Integrity, Validity, and Accuracy	Validity can be defined as “the quality of being logically or factually sound; soundness or cogency”. As a term in data quality, validity is more primitive than accuracy and relates to the logical state of data.	To be considered valid, an encounter date must meet the following criteria: (1) Encounter date must occur before the entry date (e.g. future dates imply an incorrect date entry). (2) Encounters must be recorded after 2008-01-01 (which is the system’s kickoff date). (3) If the encounter is an initial encounter with a referral date, the referral date must be within six months prior to the encounter.	<i>Valid_YN</i> – True or False (0/1)

<b>Dimension</b>	<b>Definition</b>	<b>Implementation</b>	<b>Variable</b>
Completeness	Completeness refers to the quantity of data elements that have been used to describe the breadth and depth of the clinical encounter.	To be considered complete, an encounter that is an initial encounter must have all required data elements, including a (1) referral source; (2) at least 1 secondary encounter reason; and (3) must not have a primary reason for visits of type 'Other'.	<i>Complete_YN</i> – True or False (0/1)
Usefulness, Usability, and Accessibility	High-quality data must be fit for the use by data consumers or its collection is irrelevant.	The data was flagged as used if the user accessed the report generator tool and entered data on the same day.	<i>Used_YN</i> – True or False (0/1)

### 7.1.3.3 Data Adjustments and Filters

The original dataset spanned seven years (from 2008 to 2015). Over this period of time, 294,001 encounters were recorded by 133 different users. Each user was a member of one or more of the seven different program areas.

Unfortunately, the database only started capturing entry dates as a data attribute in late 2011. Therefore, timeliness could only be calculated for a portion of the dataset. As such, we had to limit our analysis to data spanning from 2012 to 2015 in order to incorporate the timeliness dimension. This subset included 196,967 individual encounters for 110 users working in seven different program areas.

There were 440 records where timeliness was below zero (0.002% of the dataset) and 2,228 records where the timeliness was over 120 days (1.1% of the dataset). A total of 2,668 records were given a time delay of NULL and each of these entries was flagged as inaccurate during further analysis and modelling.

### 7.1.3.4 Descriptive Statistics

There were 196,967 encounters. Based on the criteria from Table 5, individual entries into the system were categorized as being timely (0 for false, 1 for true), valid (0 for false, 1 for true), complete (0 for false, 1 for true), and used (0 for false, 1 for true). The distribution of records in these categories is shown in Table 6 by program area. Many users contributed data to multiple program areas therefore the total number of users shown in Table 6 is greater than 110.

Holding each program area equal from Table 6, we calculated program area averages. As well, we calculated an overall average for all 196,967 records. These results are shown in Table 7.



Table 6. Data Summary by Program Area.

<b>Program Area</b>	<b># Users</b>	<b>Valid (% records)</b>	<b>Complete (% records)</b>	<b>Same Day (% records)</b>	<b>Useful (% records)</b>	<b>Average Delay (in days)</b>
Counselling Services	28	49,537 (96.0%)	8,657 (16.8%)	29,841 (57.8%)	2,043 (4.0%)	7.42
Diabetes Care Services	36	55,159 (99.2%)	8,019 (14.4%)	21,822 (39.3%)	4,586 (8.3%)	11.22
Health Education and Nursing Support	39	41,692 (97.3%)	17,262 (40.3%)	29,899 (69.8%)	6,359 (14.8%)	5.06
Health Promotion Activities	20	10,084 (99.5%)	4 (0.0%)	548 (5.4%)	518 (5.1%)	29.16
Homecare Services	11	9,699 (98.6%)	6,307 (64.1%)	5,640 (57.3%)	458 (4.7%)	3.91
Nutrition Consultations	32	16,868 (98.8%)	1,332 (7.8%)	9,354 (54.8%)	2,387 (14.0%)	4.37
Pharmacy Consultations	11	8,722 (88.5%)	1,159 (11.8%)	2,774 (28.2%)	880 (8.9%)	20.44

Table 7. Measure Averages.

<b>Grouping</b>	<b>Valid (% records)</b>	<b>Complete (% records)</b>	<b>Sameday (% records)</b>	<b>Useful (% records)</b>	<b>Average Delay (in days)</b>
By Program Area	96.8%	22.2%	44.6%	8.5%	11.65
Overall	97.4%	21.7%	50.7%	8.8%	9.26

Additional breakdowns of the data is provided in the Appendix H. Data is broken down by delay grouping (e.g. delay by month, etc.), entry month, entry year and program area.

### 7.1.3.5 Simple Correlations

Pearson's  $r$  correlations were calculated for each variable and are shown in Table 8. Same day and Day Delay were not correlated because both variables represent the same data quality dimension and a correlation would not be logical.

Table 8. Simple Correlations (Pearson's  $r$ ).

	<b>complete_yn</b>	<b>used_yn</b>	<b>valid_yn</b>	<b>sameday_yn</b>	<b>entry_delay_wks</b>
<b>complete_yn</b>	1.00	-0.02***	0.07 ***	0.20 ***	-0.10 ***
<b>used_yn</b>	-0.02 ***	1.00	0.00	-0.11 ***	0.05 ***
<b>valid_yn</b>	0.07 ***	0.00	1.00	0.16 ***	-0.45 ***
<b>sameday_yn</b>	0.20 ***	-0.11 ***	0.16 ***	1.00	N/A
<b>entry_delay_wks</b>	-0.10 ***	0.05 ***	-0.45 ***	N/A	1.00

*Significance Codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.'*

### 7.1.3.6 Data Quality Models

Three models were created using a forward stepwise method for simplification. The delay variable was converted to weeks by dividing the values by 7 and rounding up to the closest week. This was done during model reduction to control for over dispersion. A summary of the final simplified models is shown in Table 9. The measure of timeliness, as a count in weeks, was analysed as a negative binomial and modelled as a count data element. The measures of validity, completeness, same-day entry and use were recorded as binary variables and analysed using a logit binomial model.

Table 9. Data Quality Model Results.

<b>Data Quality Dimension</b>	<b>Dependent Variable</b>	<b>Independent Variables</b>	<b>Random Effects</b>
Completeness	complete_yn	valid_yn + useful_yn	(1 user) + (1 reportingarea) + (1 year) + (1 month)
Timeliness	sameday_yn	useful_yn + complete_yn + valid_yn	(1 user) + (1 reportingarea) + (1 year) + (1 month)
Use	used_yn	complete_yn + sameday_yn	(1 user) + (1 reportingarea) + (1 year) + (1 month)

### 7.1.3.6.1 Completeness as a function of Usefulness and Validity.

Given two levels of the completeness variable ('0' indicating an incomplete entry and '1' indicating a complete entry), multi-level logistic regression was completed and reduced with a forward stepwise method. This exploratory model investigates the relationship between the completeness of an entry, as correlated with other data quality dimensions. Neither the entry delay (in weeks) nor the same day flag for timeliness contributed significantly to the model. The coefficient results, the odds ratios and the random intercept values for users, program areas, months and years for this model are shown in Table 10.

Table 10. Coefficients for Completeness Measure.

<b>Coefficients</b>	<b>Estimate [LL-UL]</b>	<b>Odds Ratio (LL - UL)</b>
(Intercept)	-6.477 *** [-7.65 - -5.299]	0.001 [0.000 - 0.004]
valid_yn (yes)	2.493 *** [2.309 - 2.678]	12.106 [10.065- 14.560]
used_yn(yes)	0.095 ** [0.029 - 0.160]	1.099 [1.030 - 1.175]

*Random effects:* users = 13.645, month = 0.004, year = 0.032, program area = 8.203. *Significance Codes:* 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.'

### 7.1.3.6.2 Timeliness as a function of Use, Completeness and Validity.

Given two levels of the timeliness variable ('0' indicating a late entry and '1' indicating a same day entry), multi-level logistic regression was completed. This exploratory model investigates the relationship between the timeliness of an entry and how timeliness is correlated with data quality dimensions. All data quality dimensions were significant. The coefficient results, the odds ratios and the random intercept values for users, program area, months and years for this model are shown in Table 11.

Table 11. Coefficients for Timeliness Measure.

<b>Coefficients</b>	<b>Estimate [LL-UL]</b>	<b>Odds Ratio (LL - UL)</b>
(Intercept)	-4.927 *** [-5.765 - -4.09 ]	0.007 *** [0.003 - 0.017]
complete_yn (yes)	0.378 *** [0.334 - 0.422]	1.459 *** [1.396 - 1.525]
useful_yn (yes)	-0.667 *** [-0.717 - -0.617]	0.513 *** [0.488 - 0.539]
valid_yn (yes)	4.906 *** [4.653 - 5.16 ]	135.116 *** [104.869 - 174.088]

*Random effects:* users = 4.451, month = 0.011, year = 0.010, program area = 1.286. *Significance Codes:* 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.'

#### 7.1.3.6.4 Use as a function of Completeness, Timeliness, and Validity.

Given two levels of the use variable ('0' indicating no use of the data during input, and '1' indicating use of the data on the same day as it was entered), multi-level logistic regression was completed and reduced with a forward stepwise method. This exploratory model investigates the relationship between the use of data and how use is correlated with other data quality dimensions. In this model, the validity dimension did not significantly contribute to the model and was removed. The coefficient results, the odds ratios and the random intercept values for users, program area, months and years for this model are shown in Table 12.

Table 12. Coefficients for Use Measure.

Coefficients	Estimate [LL-UL]	Odds Ratio (LL - UL)
(Intercept)	-3.231 *** [-4.155 - -2.306]	0.040 *** [0.016 - 0.1 ]
complete_yn(yes)	0.071 * [0.006 - 0.137]	1.074 * [1.006 - 1.146]
sameday_yn (yes)	-0.735 *** [-0.785 - -0.686]	0.479 *** [0.456 - 0.504]

*Random effects:* users = 6.997, month = 0.064, year = 0.119, program area = 1.133.

*Significance Codes:* 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.'

## 7.1.4 Discussion

### 7.1.4.1 Qualitative Results

During the focus group discussion, six themes and data quality dimensions emerged including representation and interpretability, accuracy and validity, completeness, timeliness, believability and usefulness (see Appendix F). The focus group also provided some insight about how each data quality dimension might be measured. The group did not suggest, however, that their approaches to measurement were perfect representations of each data quality dimension. They did feel that their proposed measures were viable proxies for each dimension. Two of the themes discussed with the focus group (i.e. consistency / interpretability and believability) were interesting but according to the data analysts they could not easily be measured with existing data.

The issue of consistency was the first item mentioned during the focus group discussion and appeared to be a significant concern for the group. Believability was raised as a concern later in the discussion, but the group indicated it was less of a concern compared to the other measures. Though both measures fell out of scope for this case study, they are noteworthy in their own right and could be addressed in the future through a survey. A survey could provide a snapshot of current insights regarding

the data's interpretability and believability, but it would not be possible to assess historically. It would also be difficult to correlate these dimensions with the available historic measures and attempt to understand tradeoffs.

The remaining themes were accompanied with suggested approaches for measurement. When we had to characterize these measures as variables, we labelled them as measures of *completeness*, *validity*, *timeliness* and *use*. Timeliness and completeness were good representations of the discussion. Validity, however, was used as a final description of the themes discussed regarding accuracy, integrity and data validity. Initially, the term accuracy was used to describe the variable. However, accuracy is too nuanced and generally refers to whether or not the available data represents reality. Without comparing the database's information to clinical charts, it would not be possible to measure 'true' accuracy. Confirming the logical 'validity' was possible. Using the term validity and discussing valid data incorporated the concepts discussed with the group and aligned with the suggested measurement strategy.

The group discussed the concept of *usefulness* and *usability* of the data and suggested that measuring a user's use of the data through reports was a viable measurement strategy. However, it was acknowledged that measuring 'use' was not a perfect measure of either usefulness or usability. As an argument, the group suggested that a user would not run reports regarding data without a purpose (e.g. if it was not useful to some end), and the user would not be able to report on the data if it was not usable (e.g. if it was not possible to derive some useful meaning or generate a report). Measuring use in this way could also be a measurement of accessibility, as it determines whether or not users know how to look at their data. Overall, the *use* measure represents an interesting attribute of data quality, though it doesn't succinctly fall into a single and clear dimension. At best, measuring data use was seen as a proxy measure for usefulness from the user's perspective, but was not a measure of usefulness for the purposes of Government reporting. At worst, it measures data access. Either way, the proposed measurement maps to the use dimension of data quality.

#### 7.1.4.2 Data Quality Variables

Data quality measures were calculated for each entry in the database using the measures described in Table 5. This provided 196,967 observations of each variable.

##### 7.1.4.2.1 EntryDelay

One of our measures of timeliness was the delay in entering information. If a user saw a patient on a Monday, and they recorded their visit on the on Wednesday, the entry delay would be 2 days. This measure is significant because it shows that over fifty percent of users delay their entry. This measure allows for the discrimination of users who are extremely late versus users who are only a few days late.

The challenge with this measure, however, is a high standard deviation of 9.32 days. This led to difficulties in having models converge. Models were less sensitive and were able to converge when we used a less specific variable by measuring the delay in weeks.

#### 7.1.4.2.2 SameDay\_Y/N

Our second measure of timeline was a simple Boolean value indicating whether or not a user had met the business requirements of a same day entry. Per Table 7, 50.7% of our observations were flagged as meeting the Same Day criteria. Since approximately half of the users submitted their data on the same day, this Boolean expression is a powerful discriminator and useful for modelling.

#### 7.1.4.2.3 Valid\_Y/N

For each entry, validity was measured using the FHT's business rules and the reporting tool's interface logic. These rules were developed in consultation with the data analysts. Generally, this measure would flag, as a Boolean, whether or not the entry was logically possible. For example, it was clearly erroneous to record an encounter that predated the inception of the organization.

Logical errors in the database directly impact the quality of reports provided to the Government. Per Table 7, 97.4% of our observations were flagged as meeting the validity criteria. This validity value is comparable to the results from a medical registry case study which reported 98% accuracy based on a gold standard (Arts, 2002). For modelling purposes, these results were too homogenous to use this variable as a dependent variable in logistic regression.

#### 7.1.4.2.4 Complete\_Y/N

For each entry, completeness could be measured using business rules and logic. These rules were developed in consultation with the data analysts. Generally, this measure would flag, as a Boolean, whether or not an entry had captured all required information with specificity. According to business rules, users were supposed to record a referral source for initial encounters. If they did not, their entry would be flagged as incomplete.

These types of errors were noticed by the focus group when they attempted to provide reports by referring physician. These errors directly impact the quality and specificity of reports provided to the Government because any incomplete data would be characterized as 'Other' or 'Unknown'. Per Table 7, 21.7% of our observations were flagged as meeting the complete criteria. This number is lower than we expected. Our definition for completeness may have been too restrictive as our results differed significantly from other studies about medical registry assessments, which reported 94% completeness (Arts, 2002).

#### 7.1.4.2.5 Used\_Y/N

After reviewing the results from the focus group, this measure was the most difficult to define in practice. Conceptually, we wanted to know if users were accessing and using their data. However, generating a Boolean variable for 196,967 records was not immediately obvious; should we attempt to see if users had generated a report that included the recorded entry? If a user generated a report once per month, how could that be measured on a per entry basis? How could we aggregate the number of report generations against the other types of measures? Conceptually, measuring usefulness is helpful to understand the value of data, but difficult to measure in practice.

The data analysts described a novel strategy to measure the usefulness dimension by looking at instances where a user recorded information and generated a report at the same time. Per Table 7, 8.8% of the data entered was accompanied by a report on the same day. Though this is not a comprehensive measure of usefulness, it was a reasonable proxy to help understand whether or not users utilize their data in a literal sense. A limitation to this measure is its simplicity; for such a rich and nuanced dimension, the measure is a simple binomial and removes some important depth.

#### 7.1.4.3 Simple Correlations

We generated simple Person's  $r$  correlations. There were several  $r$ -values that showed very small and negligible relationships (e.g.  $< 0.10$ ). There was a small, statistically significant positive association ( $r = 0.20$ ) between Completeness (*complete\_yn*) and Timeliness (*sameday\_yn*), suggesting that timely entries were more likely to be complete. There was also a stronger, statistically significant negative association ( $r = -0.45$ ) between *delay\_days* and *validity\_yn*, suggesting that as the delay increased there was a higher likelihood of inputting invalid data.

While these results are interesting, they are not the best representation of the results of the dataset. The parameters in the data vary at several levels (e.g. users, program area, year and month). Simple correlations do not take the multiple observations from multiple users into consideration and treat all observations equally. The results would therefore be skewed by several types of individual differences and users with different quantities of entries. Though the coefficients were consistent with other findings, multilevel regression would offer more appropriate results by accounting for individual differences and were used to understand the relationships between variables.

#### 7.1.4.4 Data Quality Models

We developed and simplified three data quality models using multi-level binomial logistic regression.

##### 7.1.4.4.1 Completeness as a function of Use and Validity

Our model for completeness included *valid\_yn* and *used\_yn*. None of the timeliness measures were statistically significant. There were no interactions between variables.

It is interesting that timeliness and delay did not have a correlation with the completeness variable. This implies that completeness is not impacted by time and that encounters are just as likely to be complete if they are produced on the same day as they are if produced several weeks late. This may be a consequence of the relatively high random effect coefficients for program area and users, which accounts for most of the variability in the data. This would be an indication that completeness is associated with the variability of users and program area workflows. Once the random effects are removed timeliness is not significant. Thus, the variabilities in completeness is better seen as a function of users and program area than timing.

Per Table 10, there is a strong odds ratio of 12.106 between valid entries and completeness, and a weaker odds ratio of 1.099 between the use of data and its completeness. This shows that if data is used and valid, it is more likely to be complete.

##### 7.1.4.4.2 Timeliness as a function of Use, Completeness, and Validity

For this model we used the *sameday\_yn* variable to describe timeliness as the dependent variable. This allowed us to use the same binomial regression techniques as the other models. The final model included all of the other variables, since they were each statistically significant during reduction. There were no interactions between variables.

Per Table 11 there was a very strong odds ratio of 135.116 between the timeliness of data and validity, and a much weaker odds ratio of 1.459 between timeliness and completeness. This result shows that there is a very high likelihood that data was inputted into the system on the same day if the data was valid. Completeness also contributes to the likelihood of data being entered on the same day, but not to the same degree.

Per Table 11 there was a curious negative odds ratio of 0.509 between timeliness and the use of data. This suggests that the odds of an entry being timely is *negatively* correlated with users using their data. In other words, users who enter their data on the same day are less likely to run reports. Thus, users are *either* entering their data on the same day, *or* they are using their data through reports. While this may at first appear to be counter-intuitive, it suggests there may be a causal relationship; users run reports



because they are not entering their data on the same day. This interpretation is consistent with the concepts and concerns discussed during the focus group sessions.

#### 7.1.4.4.3 Use as a function of Completeness, Timeliness, and Validity

Our model for use includes *complete\_yn* and *sameday\_yn*. When choosing the best variable to represent timeliness, the *same\_day* measure was selected because it was statistically significant and provided the best AIC scores compared to the delay variable. The validity *valid\_yn* variable was not statistically significant. There were no interactions between variables.

Per Table 12 there was a small odds ratio of 1.074 between the use of data and its completeness. This implies that users who use have complete data were more likely to have used it. This is consistent with the results of the completeness model. Likewise, there was a negative odds ratio of 0.479 between use and timeliness, which is consistent with the timeliness model. Overall, the results of this model did not present new information, but confirmed the relationships uncovered in the two other models.

#### 7.1.4.4.4 Random Effects

Another interesting result from each model was the random effect measurements. The random effect coefficients represent variability in the model associated with each parameter type. Practically, random effects represent the amount of noise removed from the model, according to each variable. The random effect coefficients are summarized in Table 13.

Table 13. Summary of Random Effect Coefficients.

<b>Data Quality Model</b>	<b>Users</b>	<b>Month</b>	<b>Year</b>	<b>Program Area</b>
Completeness	13.645	0.004	0.032	8.203
Timeliness	4.451	0.011	0.010	1.286
Use	6.997	0.064	0.119	1.133

In each model, we can see that there is a relatively high amount of variability within users and program area. Thus, the measures are strongly influenced by differences in individual clinicians and program area workflows. For completeness, the impact of users is 1.65 times the impact of the program area. For timeliness, the impact of users is 3.46 times the impact of the program area. For use, the impact of users is 6.17 times the impact of the program areas. Relatively speaking, the impact of month and year on each model is very small. The variability caused by program area is consistent with the findings in Table 7.

#### 7.1.4.5 Primary Care Users and Their Tradeoffs

Our ability to measure completeness and validity during this case study was a result of poor interface design. As a solution to some of these data quality problems, it would be very simple to program the user interface to reject impossible entries (to manage validity) or incomplete entries (to manage completeness). From a strictly technical perspective, this may appear to be a valid approach. However, as observed by Hirsch (2012) applying a technical Band-Aid to a complex problem will not necessarily result in improved data quality. From a complex socio-technical perspective, understanding the user and their context is paramount to generating helpful solutions and understanding the value of interventions on the user interface.

Ironically, the results of our case study would have been significantly impacted had users been forced to validate or complete their data. Luckily, users were allowed to make mistakes and demonstrate tradeoffs. This enabled a richer understanding of users and provides some potential generalizations about primary care users. The results may provide insights about user behaviour in systems where it may be less obvious or straight forward to define and enforce valid or complete entries (such as within an EMR).

Our goal in this case study was to understand primary care users who are responsible for recording data into a web-based reporting tool. The purpose of a case study is to engage in an in-depth inquiry of a single instance. Case studies can lead to a deeper understanding of the subject, generate useful information and generate hypotheses (Flyvbjerg, 2006). Though the random effect coefficients showed that there was a lot of variability between users and program areas, the results of our models allow us to generalize certain aspect of primary care users. Our results have helped develop three hypotheses about primary care users.

##### 7.1.4.5.1 Users Prioritize Their Time, Which Impacts Data Validity and Completeness

Our models describe users balancing timeliness, validity and completeness. We calculated that when data is entered on the same day, there is a much higher likelihood that the data will be valid and complete. If users do not achieve the same day benchmark, validity and completeness are impacted negatively.

In an attempt to understand users and their tradeoffs, it is important to consider *why* users are not entering their data on the same day. Are users forgetful, lazy or disgruntled? Probably not (a least, not generally). Therefore, why are users opting out of the same day business requirement? What are users doing instead of entering their data on the same day?

Our data articulates the crux of the primary care problem and the root of this conundrum: clinicians are asked to code data. Generally clinicians see patients as a priority and deliver healthcare services as a vocation. Whereas half of the users are able to incorporate the same day entry into their daily

workflow, the other half is not successfully adopting data entry into their day to day processes. These users are prioritizing other activities over entering data. Hopefully they are spending more time with patients and at least sacrificing data validity and completeness for patient care. In this sense, sacrificing timely data entry would be understandable.

This tradeoff is important to understand when asking clinicians to report their activities and data. If organizations wish to produce high-quality data that is timely, valid and complete, they need to ensure users have sufficient time to tend to data entry tasks. If users are forced to decide between succinctly entering data on the same day instead of spending more time seeing patients, half will opt to see more patients. Underlying our observations is not a user interface concern, but a reality associated with workflow and complex socio-technical decision making. Improving data quality will involve documenting workflows and processes that vary by user and program area, and finding ways to encourage and support same-day data entry.

#### 7.1.4.5.2 Users Can Achieve Validity Without Sacrificing Completeness

During the qualitative interviews, it was suggested that accuracy and validity problems are likely negatively associated with completeness, since forcing users to enter *something* will force them to enter *anything*. We also discussed the example of users picking “Albanian” as a patient’s nationality because it was the first item at the top of the drop down list. Conceptually, there is an idea that if you require users to enter complete data, they will start entering arbitrary values into fields because they either do not have the necessary information, lose interest, or do not care.

Our results point to the contrary of this hypothesis: there is a correlation between users ensuring their data is valid and submitting complete data. Given our results, there is *not* a tradeoff between validity and completeness and users are able to achieve both types of data quality simultaneously. Conscientious users address both requirements simultaneously. As an insight from our case study, we have shown that managers are reasonable to expect both types of data quality simultaneously.

Practically, this has implications for measuring total data quality in systems where there is limited metadata. Generalizing our results, measuring completeness (which, arguably, is not necessarily difficult) could be a good parallel measure of data accuracy and global data quality.

#### 7.1.4.5.3 Use of Data as Compensation

An interesting result from our case study is the increased use of data when there is a delay in entry; there was a negative correlation between entering data on the same day and using data. We also saw a correlation between use and completeness, and use and validity measures. Thus, users who do not

enter their data on the same data run more reports to compensate with use, and this use helps counteract the negative consequences of a delayed entry.

These results offer an interesting insight into users and could explain why users are generating reports: when data input is delayed, users generate reports to understand what has been submitted and what has not. Generally, users who have entered their data on the same day have no reason to review their data and interact with it. As a measure of system effectiveness, this is not a good result: the results suggest that users are not engaging with their data because they find their data interesting or helpful, but because they need support in completing their data entry task.

In a sense, the system is merely self-reinforcing and not providing users with value. An ‘effective’ system would likely have little correlation between data use and data entry, and would demonstrate a richness of use beyond self-reinforcement. Future systems should aspire to break this correlation.

#### 7.1.4.6 Implications and Application of Findings

##### 7.1.4.6.1 Implications for System Design

There are several implications for this work that will interest developers and engineers. Most obviously, we noticed that completion and validity within the database could have been controlled more effectively at the point of entry. The user interface could have been designed in such a way as to reject impossible (invalid) data submissions and force users to provide all required information and ensure completeness. The user interface can also flag questionable entries for the user to review. There are obvious opportunities to improve data quality deficiencies through design interventions that would be simple to implement, and could be implemented pre-emptively in other systems. As discussed, this needs to be approached within a larger context and is not merely about technical interventions.

Another important implication for engineers is to consider building relevant metadata into data structures to enable future data quality measurements. Our case study is a reminder that if there is a need to measure some form of data quality over time, the system will eventually need data to generate measures. In our case study, we created measures based on focus group results and available data points, but were limited in our scope by the system’s data attributes. On one hand, we’ve shown how to gather insight from a database that did not have any direct data quality measurement tools and metadata built within its core design. On the other hand, we’ve shown the limitations of trying to measure data quality in a system that had few tools designed to measure data quality at inception.

In order to successfully encourage and measure data quality, engineers and designers need to consider data quality early in the implementation and design of systems. In the same way that there is a call for ‘privacy by design’ during the implementation of new systems (see Cavoukian et al., 2010), a

similar ‘quality by design’ paradigm would be an appropriate approach for engineers and designers to consider.

#### 7.1.4.6.2 Implications for Healthcare Environments

Our case study presents interesting results specifically for the primary care domain. The dataset that was analysed was manufactured within the same ecosystem, and previous work by St-Maurice and Burns (2014) suggest that there are similar data quality problems within the EMR. Since some of our insights relate to workflow, the implications of this could extend past the FHT’s reporting tool and into the use of EMRs and clinical documentation. Developers and engineers creating EMRs in primary care should carefully consider the processes in which EMRs can be successfully adopted into workflow and understand that a delay in data entry creates more work for clinicians (e.g. increased time to generate and use reports) and decreases data quality (e.g. reduces completeness). Encouraging and supporting prompt data entry is an important goal and understanding users is valuable.

Outside primary care, there are other instances in healthcare where clinicians are asked to input data into a reporting system. Either through the completion of forms or through the double entry into a second system, the workflow described during the case study is common. For example, in Ontario’s long-term care system, nurses are required to complete assessments using RAI-MDS (Hirdes et al., 2000). Similarly, nurses complete RAI-HC for the Department of Veterans Affairs (Hawes, Fries, James, & Guihan, 2007). These medical registries can be improved through processes and policies (Arts, 2002) but may experience some of the same tradeoffs discussed in our findings because the data entry is performed by clinicians.

#### 7.1.4.6.3 Contributions to the ‘New’ World

Arguably the world of healthcare is changing. Healthcare is evolving towards learning and adaptive healthcare systems, interoperability, IBM-Watson and large amounts of patient generated data; we are entering a ‘new’ era of data in healthcare.

From this perspective, this study is very important. Our case study clearly demonstrates that healthcare environments with clinical users still struggle with basics. Unless clinicians are able to enter information accurately and completely on a consistent basis through the healthcare spectrum, the likes of IBM-Watson are not going to have access to data to interpret. Understanding users and how they can better input is a precursor to feeding new data analytics engines. Fixing and understanding the issue of basic data input in healthcare is a precursor to any ambitions to delivering big value in the ‘new world’.

#### 7.1.4.7 Future Work

There are many interesting opportunities for future work with the existing dataset. Our random effect coefficients were grouped into single values but could be further broken down by additional predictive variables. Without acquiring new data, it would be possible to compare months of data entry (e.g. to explore the idea of seasonal data quality) and to compare program areas (e.g. to assess data quality in nurses compared to social workers). Understanding the nuances of different months and program areas could provide valuable insights regarding workflows and the timing of potential interventions. There could also be additional user interviews to provide about our findings. These interviews could help further describe the relationships suggested by our results and find causes of user behaviour.

Another opportunity relates to other FHTs using the same database reporting tool. These comparable organizations represent different geographies and organizational structures. Comparing the results from our case study to data from other ‘sister’ organizations would provide some valuable insight into the data quality phenomena that were observed. Whereas our current work is based on several hundred thousand records, it would be possible to analyse a dataset of two million records broken down by several different primary care organizations. This would support the generalization of our findings and would provide an opportunity to test our hypotheses.

Now that we have a better understanding of primary care users and their tradeoffs, our results could be used to identify opportunities to develop new designs and enhance processes that support data quality. How could we encourage users to enter their data on the same day? Would it have the anticipated benefits? What design paradigms would be effective for encouraging and rewarding users? We plan to conduct A/B tests on design concepts to better understand how data quality could be improved by encouraging same-day entries.

Finally, we have described a hypothetical link between our dataset and data from primary care EMRs. Our hope is that our findings would apply to other systems used by clinicians in the primary care ecosystem. Understanding the user’s interaction with the reporting database could be insightful in trying to understand their relationship with health records. However, this relationship needs to be defined and explored more carefully in a separate study. This would involve comparing the data quality measures and results from an EMR and looking for similarities and differences. If such a study was successful in demonstrating a link between several primary care tools, we could further generalize our findings.

#### 7.1.5 Conclusions

The crux of data quality problems in primary care is that clinicians are asked to manufacture coded data in addition to their health care responsibilities. These users balance several priorities and are part of a complex socio-technical system. To understand users and their tradeoffs in this environment, we

used qualitative and quantitative methods to understand how these users tradeoff and balance their priorities, and how this impacts data quality in primary care. We defined several variables based on available data in a web-based reporting tool.

We have shown that delayed entries have a negative impact on data quality, and that users in primary care should be accorded sufficient time to complete data reporting workflows. Users should not be placed in a position where they must decide between patient care and data input, as half of the users will opt to provide additional patient care. As well, our results debunk a myth that suggests users will either provide accurate data, or complete data; there was no evidence of tension between these two measures. We also observed a negative relationship between use and timely data entry. As a measurement of a systems effectiveness, this points to a self-reinforcing data repository that provides users with little value. A better system delivering more value to users should not show a correlation between data use and data entry.

These results have useful implications for designers and engineers that are engaged in designing data quality measures, creating primary care EMRs and analysing medical registry data. Future work will include testing designs that encourage specific behaviours that are linked with higher data quality.

## 7.2 Chapter Summary

### 7.2.1 Key Findings

**A Complex Context, A Simple Interface:** The interface through which users are asked to input data in this example is straightforward. There is no reason that an experienced user would not be able to enter data within 30 seconds. The limiting factor with this system is not technology, but ecological and cognitive. This work further supports the idea that codification itself is ‘easy’ on the surface (e.g. clicking boxes and selecting items in menus), but that the underlying environment and context that requires clinicians to code their own data is problematic and challenging. If it is difficult for users to enter timely and accurate data in a structured, simple system, there are reasons to question the ability of users to enter equally simple data into an EMR.

**User Tradeoffs:** It is clear that the users are not ‘good robots’ who follow all organizational rules. The results of this chapter begin to clarify and articulate reasons for poor data by looking at different use behaviours and tradeoffs.

**Benchmarks:** The results from this chapter provide definitions and benchmarks for data quality measures that will be helpful for identifying priorities and measuring future improvements.

**CWA Models:** The behaviours and tradeoffs described through this work are consistent with the findings of the modelling work from Part 1 of the dissertation. For example, the fact that users prioritize

their tasks and delay data entry aligns with the varying levels of expertise discussed in the ConTA models (Appendix C) and the StrA IFMs (Appendix D).

### **7.2.2 Connections to Case Study and Research Questions**

I have three research questions:

1. In primary care, how are individual users influenced by their environment to input high-quality data?
2. What techniques could be used to design systems that persuade users to enter higher-quality data?
3. Is it possible to improve data quality in primary care by persuading users with the user interface?

This chapter was a case study and provided a real example of data, data quality, user decision making, and ecological factors. This chapter goes into further detail compared to the previous chapter, provides greater details about how a codification task takes place in an environment, and shows how users are influenced by aspects of their environment and context (e.g. available time). The results also provide insight to what type of intervention may be required to improve data quality in a real system. Techniques to improve data quality will require a baseline, which is provided here.

Users have been characterized as making dynamic choices when using the system, even though the technology is essentially just a basic web survey. The task itself is easy, but in the context of a clinical practice, the task is challenging to complete on time, varies highly in quality, and varies from program area to program area. The findings only are well explained with a complex socio-technical lens.

### **7.2.3 Connections to Other Chapters**

In chapter 8, design concepts are developed based on the findings of the previous CWA models. The results of this chapter are used to identify design goals toward specific behaviour changes. The next chapter continues to examine this example by developing specific PD to improve data quality based on these results.



## **Chapter 8**

### **Design Concepts for Data Quality Improvement**

In the previous chapter, I analysed a case of data quality in a primary care system. I use the results of my analysis in this chapter by developing several mockups and design concepts that are intended to help improve data quality by encouraging specific behaviours. The mockups were also generated by interpreting the results of my CWA models.

#### **8.1 Persuasive Design**

##### **8.1.1 Using Social Persuasion with Clinicians**

In the introduction to my dissertation, I mentioned several examples of data quality improvement. ABE consisted of local meetings where comparative data was presented to individual physician practices (de Lusignan et al., 2006) and was described as the most successful change agent toward clinician attitudes toward data quality (de Lusignan, 2005). More recently, a data quality feedback tool generated comparative data quality feedback and was successful at improve recording (van der Bij et al., 2016).

All of the previous examples involved a comparison of users to generate successful behaviour change. This is part of two basic social processes—persuasion and social comparisons—that are effective at creating lasting change. Thus, the effectiveness of these studies can be explained by social psychology processes, which provide an understanding of the processes that result in changes to personal belief, opinion, self-efficacy, and behaviour (Suls & Bruchmann, 2013). These studies strongly suggest that persuasion and social comparisons are effective approaches to changing clinician behaviour. The studies also provide evidence that persuasion would be an effective approach to improving data quality.

##### **8.1.2 Persuasive Technology**

Persuasive systems may be defined as “computerized software or information systems designed to reinforce, change or shape attitudes or behaviors or both without using coercion or deception” (Oinas-Kukkonen & Harjumaa, 2008). Persuasive Design Techniques (PDTs) are a set of techniques that can be adopted by system engineers to create Behaviour Change Support Systems (BCSS) to help users adopt a

new behaviour, or to adapt their current behaviour (Oinas-Kukkonen, 2010). PDTs are built on several psychological theories, such as Herzberg's theory of motivation and Maslow's hierarchy of needs.

Four notable PD models have been developed (Torning, 2013), including the Design with Intent method (Lockton, Harrison, & Stanton, 2010), the Persuasive Systems Design (PSD) Process Model (Oinas-Kukkonen & Harjumaa, 2009), Fogg's (2009b) Eight-Step Design Process, and Fogg & Hreha's (2010) Behaviour Wizard model. In a comparison, Torning (2013) found that the Design with Intent Method and the PSD model met more classic design criteria compared to Fogg and were better grounded in design theory. (This is not surprising, as Fogg was trained as a behavioural psychologist and not a computer programmer.) In the literature, specifically in healthcare applications of PD, the PSD model is more regularly referenced.

Using technology to persuade users has many advantages. Systems have varied benefits as persuaders, such as interactivity and an ability to adapt persuasion strategies according to users' actions. This kind of interactivity has not been possible in traditional media (Harjumaa, 2014). As well, per Fogg, (2003), systems can be more effective than human persuaders because they

- are persistent
- offer anonymity
- exist in locations and contexts that are not possible for humans
- quickly adapt to large amounts of data
- simultaneously attempt several modalities to influence people
- are able to scale to a large user base

In healthcare, the use of PDT has focused primarily on aiming to change patient behaviour and positively change health outcomes. As an example, in 2012, 101 studies had looked at 83 different interventions on web-based products that aimed to persuade users toward better health (Kelders, Kok, Ossebaard, & Van Gemert-Pijnen, 2012). In contrast, few examples that use PD to influence clinicians in healthcare are found.

As persuasive approaches to improving data quality have been successful, it makes sense to offload the persuasion to technology. Extending previous work on data quality with PDT, formal behavioural change models and design framework models could be a highly effective approach to improving data quality in primary care. At present, I do not know of any existing uses of PD to improve data quality with clinical users.

### **8.1.3 Fogg's Behavioural Change Model**

PD can be used to create persuasive technology through the analysis of behaviour models. Fogg (2009a) suggests that behaviour change is a function of motivation, ability, and triggers (i.e. Behaviour =

fx [motivation, ability, triggers]) and proposes the Fogg Behaviour Model (FBM). The idea is that each element of behaviour change must be triggered simultaneously, or the behaviour will not take place (B. Fogg, 2009a). Fogg further explains that if behaviour change does not occur, at least one of the elements from his BMAT model is missing.

According to Fogg's (2009a) BMAT model, it is important that the triggers occur in a context where users have enough motivation and ability to perform tasks related to behaviour change. For example, if a user is highly motivated to do a task, but not capable of performing it because of a skill gap, a trigger (e.g. a prompt) will not result in behaviour change. Ability is sometimes also referred to as capacity. In the same way, if a trigger is applied to a user who lacks the required resources to perform the task (such as time or money), a trigger (e.g. a prompt) will not result in behaviour change. Fogg suggests that behaviour change is most challenging with tasks that are hard to do, and motivation is low. Generally, motivation and ability are limiting reagents to behaviour change. Fogg's model can be used to identify obstacles that inhibit people from performing a specific behaviour.

The FDM is helpful because it clarifies *when* behaviour change will occur.

#### **8.1.4 Persuasive System Design Model**

Based on several behavioural change models (including the FBM), Oinas-Kukkonen and Harjumaa (2009) developed a framework for designing persuasive systems, known as the PSD model. Their model presents a set of PD features, and helps align the intent, events, and strategies that can be used to persuade users. The PSD model is helpful because it categorizes design features into the following four categories:

- Primary task support: features that support users in performing their primary task
- Dialogue support: features that support users while interacting with the system
- Credibility support: features that make systems more credible and, therefore, more persuasive
- Social support: features that motivate users by increasing social influence through the system

A combination of features from these four categories can be used to achieve persuasive goals. The PSD model helps identify *how* a user may be persuaded, but requires a design to answer *who* the users are, and *why* the change is required in order to build an appropriate persuasion context. Oinas-Kukkonen and Harjumaa also discuss key foundations to PD, such as understanding that: information technology is always doing 'something'; a commitment and consistency is needed for PD; direct and indirect routes to behaviour change are possible; and solutions should be incremental, open, unobtrusive, and easy-to-use.

### 8.1.5 Literature Gaps

Oinas-kukkonen & Harjumaa (2009, p. 489) state that “without carefully analysing the persuasion context, it will be hard or even impossible to recognize inconsistencies in a user’s thinking, discern opportune and/or inopportune moments for delivering messages, and effectively persuade”. The authors continue by describing the importance of recognizing the intent of persuasion, understanding the persuasion event, and defining the strategies to use. The value of understanding the use context, user context, and technology context of persuasion is highlighted.

The gap with the Oinas-Kukkonen & Harjumaa (2009) discussion of the persuasion context is that it discusses the importance of understanding context, but it fails to deliver specifics about how to analyse a domain or system and extract the necessary elements to *build* a persuasive context. No toolkits for systems analysis are suggested to accompany the PSD.

Recently, Alahäivälä & Oinas-Kukkonen (2016) noted that studies often fail to provide a systematic analysis of contextual factors, and that systematic analyses of the persuasive contexts have been lacking. Other studies also have noted a gap in models that can inform the design of PDT (Mohr, Schueller, Montague, Burns, & Rashidi, 2014). As well, literature reviews have focused on extracting, analysing, and categorizing persuasive system features but have not characterized system analysis approaches to support PD (Lehto & Oinas-Kukkonen, 2011). It is not clear if the details around systems analysis are underreported, or done arbitrarily. Generally, a formal systems analysis framework does not appear to be associated with PD. This gap was highlighted by Fogg in 2009 when he suggested that there are not many well defined processes for designing persuasive technology, and that practitioners regularly adapt methods from other fields (B. Fogg, 2009b). While some work has addressed the design aspects of PD, it does not appear that there are many well-defined processes that take designers from analysis to implementation.

## 8.2 Defining the Persuasion Context with Cognitive Work Analysis

Throughout my dissertation, I have defined data quality as a complex socio-technical system. Currently, no obvious approaches are available for linking my results to PD. As such, I needed to develop an approach for linking these frameworks. Luckily, the idea of defining a ‘persuasion context’ per Oinas-Kukkonen & Harjumaa (2009) is very compatible with the *raison d’être* of CWA. Since the CWA framework provides a systematic approach to understanding context, environment, and cognition, it easily addresses many of the information requirements described by Oinas-Kukkonen & Harjumaa. As well, the idea of tying CWA to FBM and PSD has previously been explored by Rezai & Burns (2014), though with only a few phases of the CWA framework.

To link CWA, FBM, and PSD, I took the work of Rezai & Burns (2014) a step further by adopting a Who, What, Why, When, How (WWWWH) paradigm. For each of the questions of the WWWWH, I linked appropriate sources of information from either CWA or the FBM. The answers to the WWWWH approach provided a persuasive context, which could then be used by the PSD. My ecological approach takes advantage of the strengths of each framework: CWA provides insight about context, environment, and cognition; the FBM provides information about when change will occur; and the PSD provides tools and design ideas that can create a change in behaviour. The combination of these frameworks filled the analysis-to-design spectrum with a series of useful tools and sources of information. My approach is described in greater detail in Table 14.

Table 14. Linking Frameworks for an Ecological Approach to Persuasive Design.

	PSD			FBM	Analytical Need	CWA Context	CWA Analysis
	Intent	Event	Strat.				
Who	Persuader	User			Identify the persuader and the user, or class of users, who are the target of the persuasive intervention.	Ecological	ConTA (DLs) SOCA (IFMs)
What	Change Type	Technology			Identify what behaviours need to change and what the new target action or behaviour looks like.	Cognitive	ConTA (DLs) DL comparisons (junctions)
Why		Use			Contextualize the reasons for the task in the complex system. Why did old behaviours develop and what are the constraints on new behaviour?	Ecological	WDA (AH)
When				Motivators	Identify motivating factors within the environment.	Ecological	WDA (AH)
				Abilities	Identify user abilities and capabilities. Identify constrained resources (e.g. time, money, etc.).	Cognitive	ConTA (DL) WCA (SRK)
				Triggers	Identify reasons users adopt specific behaviours or strategies.	Ecological	ConTA (DLs) StrA (IFMs)
How			Message Route		How will we create a change? What design principles and strategies would be appropriate?		

My approach is similar to the framework described by Mohr et al. (2014). The Mohr model aimed to “provide a framework for the translation of treatment and intervention aims into an implementable treatment model”. Similar to my approach, the Mohr model is based on a WWWWH paradigm with some minor differences. For example, Mohr et al. capture behaviour change strategies under ‘How (Conceptual)’ and define the specific intervention heuristics under ‘How (Technical)’. In my approach, I describe behavioural change strategies under ‘When’ (e.g. when will the change occur) and use the PSD to identify the technical heuristics and persuasive strategies. Mohr et al. use ‘When’ to describe the user and task completions, where I use ‘What’ to articulate the specific behaviour changes (e.g. What needs to change).

The Mohr approach does not provide a set of tools to define data sources, and analytical tools to systematically gather information; like the work by Oinas-Kukkonen & Harjuma (2009), it provides a theoretic basis for needing the information for each question. As well, the Mohr’s use of the ‘How’ for both theory and application blurs the elegance of the model. It also attempts to further break down the analysis into theoretical and instantiation groupings. In my case, a breakdown of theoretical and instantiated components is not required, because my approach assumes that most of the necessary information is available from a rich CWA analysis and that the theoretical aspects of the analysis already have been incorporated.

I describe each step of my approach in greater detail next and provide examples from my CWA of data codification.

### **8.2.1 Who is the target?**

This part of my approach involves identifying the user, or class of users, that is the target of the persuasive intervention. The types of users generally should be captured within the ecological analysis performed during the WDA of a CWA, and should be attached to the processes under study.

In my overarching case study, the users are the clinicians who are employing a FHT’s reporting tool. These users clearly are linked to the task in AH 2.1 (Appendix B). More generally, however, identify which user or group of users might be more challenging within the confines of a complex socio-technical system. In situations involving team dynamics, identifying users might involve reviewing multiple AHs or using SOCA to understand which user is the target of PD.

### **8.2.2 What is the change?**

This part of my approach involves identifying what behaviours need to change and what the new target action or behaviour needs to be. This information can be taken from the ConTA phase of CWA. This is a good fit because ConTA is really modelling *what* must be executed and/or processed by the

system and controller and describes these goals regardless of who is supposed to conduct the activity, or how (McIlroy & Stanton, 2015).

In addition to using ConTA to identify interesting tasks, I can use my comparison approach. Recall, I compared different DLs from different domains or levels of expertise to identify junctions. These junctions identify specific design opportunities that can inform current and desired user behaviours. The opportunities are in effective persuasive opportunities and represent where a system might benefit from a change in user behaviour.

As an example, one of the junctions identified a difference between expert and novice users at the Alert state. Based on this junction, the opportunity exists to have expert users favor a novice approach to entry and process data on the same day. Thus, the specific change in behaviour would be to have more-experienced users enter their data on the same day once they have been alerted (by virtue of their workflow) that data must be entered.

### **8.2.3 Why is the task completed?**

This part of my approach involves contextualizing the reasons for old behaviour and the constraints on new behaviour. In addition, the goal is to understand why the behaviour needs to change and what are the potential limitations. This information should be captured in ecological analysis performed during the WDA of a CWA, and should manifest itself as a series of abstract functions and functional goals.

As an example, in my AH, the record encounter task is linked to the abstract functions of ‘Benchmarks and Norms’ and ‘Professional Values’. Both of these links in the AH help to contextualize the task and offer insight into why it is important. In this case, entering data is a professional responsibility associated with patient care, and enables benchmarks. Establishing and measuring norms and benchmarks are associated with accountability; therefore, the task is important. In terms of understanding constraints and reasons for old behaviour, the link to ‘Professional Values’ and ‘Training’ is insightful. Building and moderating behaviour through a sense of “duty” or by developing the sense of a professional norm could be valuable approaches to PD.

### **8.2.4 When will the change occur?**

To understand the conditions under which behaviour will change, I refer to the FBM. This model suggests that behaviour change is a function of motivation, ability, and triggers. The FBM is built on the idea that behaviour change only occurs when there is a combination of proper motivation, ability, and trigger.

#### 8.2.4.1 Motivators

This part of my approach involves identifying what motivators may already exist within the user's environment. The concept is to catalogue existing motivators that are currently present and develop a better appreciation for potential motivating factors. Enhancing current motivators or introducing new motivators might be effective strategies. Fogg (2009a) describes motivators as either sensation (pleasure/pain), anticipation (hope/fear), or belonging (acceptance/rejection). He suggests that these core motivators are central to the human experience.

Information about motivators may be captured in ecological analysis performed during the WDA of a CWA. Motivators may manifest themselves as a series of abstract functions and functional goals and might take the form of tradeoffs or balances.

If my goal was to have users enter their data more regularly, I would look at the record encounter task in the Generalized Function and see that the task is tied to benchmarks and norms. The relationship between the task and the abstract function is that users are responsible for meeting organizational benchmarks; failing to report data could result in disciplinary action. According to the FBM, this type of motivation would be fear of disciplinary action. It is possible that other forms of motivation, other than fear, could be effective at improving data quality.

#### 8.2.4.2 Abilities

This part of my approach involves identifying user abilities and capabilities and identifying constrained resources (e.g. time, money, etc.). Generally, the issue of user abilities is nicely explored by CWA's WCA with the SRK taxonomy.

If my goal was to have users enter their data more regularly, I would identify time as a constraint, and users would need time management abilities. Referring to SRK 1.1 in Appendix E: steps 7 and 8 describe time management capabilities for users.

#### 8.2.4.3 Triggers

This part of my approach involves finding reasons that users adopt or change behaviours. My approach to triggers is similar to my approach to motivators. The concept is to catalogue existing triggers that are currently present in the system. Enhancing current triggers or introducing new triggers might be effective strategies. Removing negative triggers that encourage undesirable behaviour in the system also would be an effective design approach, once they are identified.

In my example, I analysed different triggers for strategies during my comparison of CWAs. Based on my comparison of StrA in Table 2, I can see that users select strategies based on organizational



policies, workload, experience, technical abilities, and practice workflows. Changes and proxy-persuasions for any of these factors, supported with PD, could be viable trigger mechanisms.

### **8.2.5 Answering ‘How’ by Identifying the Persuasion Strategy**

To transform my persuasion context into designs, my approach uses the PSD. By using PSD to identify and categorize persuasion strategies, I am effectively completing the WWWWH paradigm by answering ‘How’. Identifying the persuasion strategy involves selecting a route, which is either indirect or direct, and selecting a message (Oinas-Kukkonen & Harjumaa, 2009). The route and the message must be compatible with the persuasion context.

## **8.3 Persuasive Designs to Improve Data Quality**

The results of chapter 7 are an important part of my research because they allowed me to identify specific design objectives. This enabled me to continue my case study by measuring the impacts of new designs in chapter 9. In this section, I develop my persuasion context with the results of my CWA and identify three design goals based on my findings from chapter 7. I use PD principles from the PSD model to create mockups of interventions that can improve data quality. My list of mockups and design ideas reveal possibilities, and are not intended to represent the entire breadth of potential PD intervention that could possibly be applied to the reporting system. My criteria was to identify ‘low hanging fruit’ that could be introduced easily to the reporting tool during a field study (chapter 9).

### **8.3.1 Goal 1: Reduce Entry Delay**

Based on the results of chapter 7, this was clearly a significant issue with timeliness in the FHT’s reporting tool. Per Table 7, between 2012 and 2015 only 50.7% of users were entering their data on the same day as their encounter, and there was an average 9.26 delay in days for entering data. Based on the interview results in Appendix F and the correlations from chapter 7, reducing the delay should improve data quality within the system and is a valid design goal.

I used my ecological approach from the previous section to define the persuasion context. I employed information from my CWA about data codification. I focused on using information from the junction described in Figure 12. The persuasion context is shown in Table 15 and builds on the framework mapping established in Table 14.

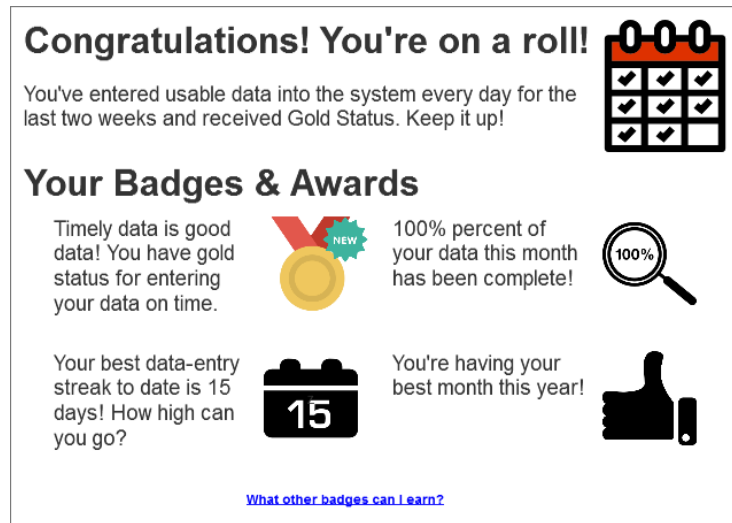
Table 15. Persuasion Context to Reduce Entry Delay.

Question		Answer	References
Who		Our target users are AHPs entering data into the FHT reporting tool. There are no complex team dynamics or multiple classes of users.	AH 2.1 in Appendix B.
What		Based on the comparison of Novice and Expert users, there is an opportunity to have expert users favor a novice approach to entry and process data on the same day. This occurs at the Alert level of the ConTA, when users have finished a patient encounter.	ConTA 1.1.1 - 1.1.3 in Appendix C. Junctions in Figure 11 in chapter 6.
Why		Summarizing the data is related to benchmarks and norms. The task will help the organization be accountable. Timely data will allow the organization to respond to needs more quickly. Professional Values and Training provide potential insightful constraints on the change. Building and moderating behaviour through a sense of ‘duty’ or by developing the sense of a professional norm could be a valuable approach to PD.	AH 2.1 in Appendix B.
When	Motivation	Users have professional values which will lead them to input data. Users are responsible for meeting organizational benchmarks; failing to report data could result in disciplinary action.	AH 2.1 in Appendix B.
	Abilities	Users need to prioritize their time and engage in time management to change this behaviour. They need time and time management abilities.	Steps 7 and 8 of SRK 1.1 in Appendix E.
	Triggers	Users make strategic choices (and are triggered) by organizational policies, workload, experience, technical abilities, and practice workflows.	Comparison of StrA in Table 2
How	Message	My message primarily asks users to enter data on the same day. I want to appeal to their sense of professional duty and want to establish the behaviour as a professional norm. I want users to think about entering their data right away and avoid the bulk entry strategy. I want to encourage them to use the same-day workflow.	Persuasion Context Analysis
	Route	Direct or Indirect.	Persuasion Context Analysis
Strategy		Reduce entry delay, a dialogue-based persuasion strategy could be appropriate. Effective approaches might include Praise, Rewards (virtual), or Suggestions. Reduce entry delay, a persuasion strategy based on social support, could also be appropriate. Effective design principles might include Social Comparison, Normative influence, and Social Facilitation.	PSD

My first design uses the praise principle, which states that “by offering praise, a system can make users more open to persuasion” (Oinas-Kukkonen & Harjumaa, 2009, p. 493). It attempts to encourage

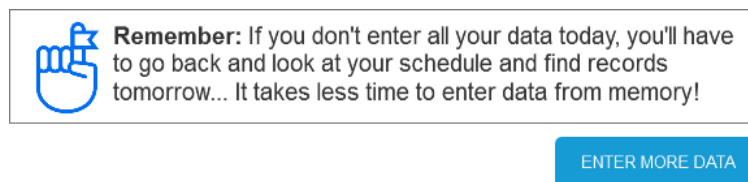
and normalize entering data on the same day. The PD should encourage users to think about keeping their statistics and award numbers high and change their workflows and entry strategies. The route is direct. This concept is shown in Figure 13.

Figure 13. A use of the praise dialogue principle to reduce delay.



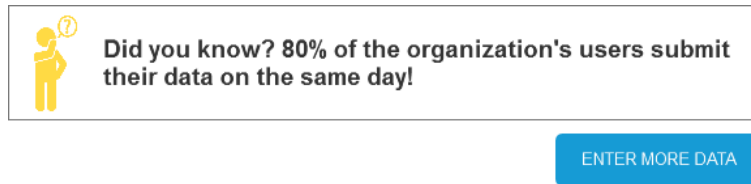
My second design uses the suggestion principle, which states that “Systems offering fitting suggestions will have greater persuasive powers” (Oinas-Kukkonen & Harjumaa, 2009, p. 493). It attempts to encourage and normalize entering data on the same day and appeals to a user’s time management skills. The PD is designed to encourage users to change their workflow and adopt the same-day strategy from StrA. The route is direct. This concept is shown in Figure 14.

Figure 14. A suggestion dialogue to reduce delay.



My third design uses the social facilitation principle, which states that “system users are more likely to perform target behaviour if they discern via the system that others are performing the behaviour along with them”(Oinas-Kukkonen & Harjumaa, 2009, p. 495). It attempts to normalize the same-day entry behaviour by showing how common it occurs in the organization. It also tries to show that the behaviour is a normal professional standard and part of a user’s professional “duty”. This is an example of an indirect route. This concept is shown in Figure 15.

Figure 15. A simple use of the social facilitation principle to reduce entry delay.



### 8.3.2 Goal 2: Increase Completeness and Accuracy, and Enhance Believability

Based on the interview results in Appendix F and the correlations from chapter 7, value would be seen in improving accuracy and completeness within the system. Per Table 7, between 2012 and 2015, only 21% met the completeness requirement. At the same time, interviews (Appendix F) noted that sometimes a trust issue is experienced with the data, and the data needs to be double-checked with information in EMR system; they also noted issues with interpretability of the data. Based on these comments, it is not clear if users understand how data is measured as timely, complete, or valid. To support a better understanding of data quality and to build trust, users should be presented with opportunities to improve and understand their current score or status and understand what inputted data ‘means’ in the bigger picture.

I focused on using information from the junction described in Figure 11. The persuasion context is shown in Table 16; it builds on the framework mapping established in Table 14.

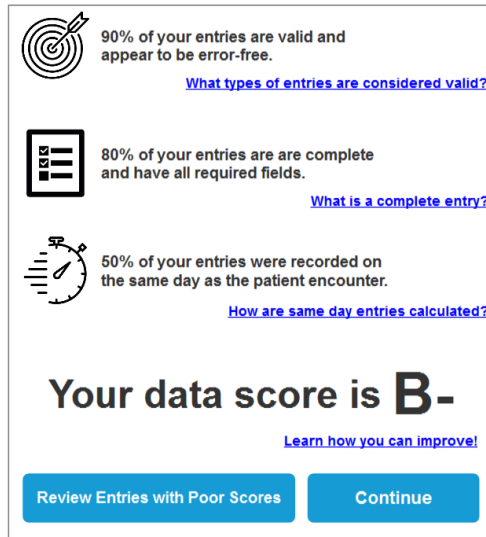
Table 16. Persuasion Context to Increase Completeness, Accuracy and Enhance Believability.

Question	Answer	References
Who	Our target users are AHPs entering data into the FHT reporting tool. There are no complex team dynamics or classes of users.	AHPs shown in AH 2.1 in Appendix B.
What	Based on the comparison of Novice, Intermediate, and Expert users, there was an opportunity to have users adopt intermediate user behaviour by interpreting tasks more deeply after reaching the system level of the DL. Users need to more carefully consider the data they are entering.	ConTA 1.1.1 - 1.1.3 in Appendix C. Junctions in Figure 10 in chapter 6.
Why	Summarizing the data is related to benchmarks and norms. The task will help the organization be accountable and allocate resources to maximize health outcomes and population health. Accurate data will help to monitor population health effectively. Professional Values and Training provide potential insightful constraints on the change. Building and moderating behaviour through a sense of ‘duty’ or by developing the sense of a professional norm could be valuable goal for a PD.	AH 2.1 in Appendix B.

Question		Answer	References
When	Motivation	Users have professional values which will lead them to input data. Users are responsible for meeting organizational benchmarks; failing to report data could result in disciplinary action.	AH 2.1 in Appendix B.
	Abilities	Users need to prioritize their time and engage in time management to change this behaviour. They need time and time management abilities.	Steps 7 and 8 of SRK 1.1 in Appendix E.
	Triggers	Users make strategic choices (and are triggered) by organizational policies, workload, experience, technical abilities, and practice workflows.	Comparison of StrA in Table 2.
How	Message	My message needs to make users prioritize completeness and accuracy of data. I want to appeal to their sense of professional duty and want to establish the behaviour as a professional norm. I want users to understand the quality of their data and how it might impact their benchmarks. I want users to think more deeply about their data. I want users to trust the system and value the data they input into the system.	Persuasion Context Analysis
	Route	Direct.	Persuasion Context Analysis
Strategy		Improve completeness and accuracy, a persuasion strategy based on primary task support, could be appropriate. Effective approaches might include task simplification, tailoring alerts, and self-monitoring. Flagging input errors and aiming for task simplification are relatively straightforward paradigms. Users described problems with believability (e.g. trusting the system's recording of data) and expressed difficulty understanding how data was being tabulated by management. To clarify how data is captured and how it will be used, Oinas-kukkonen & Harjumaa (2009) suggest applying system credibility principles, such as verifiability and trustworthiness.	PSD

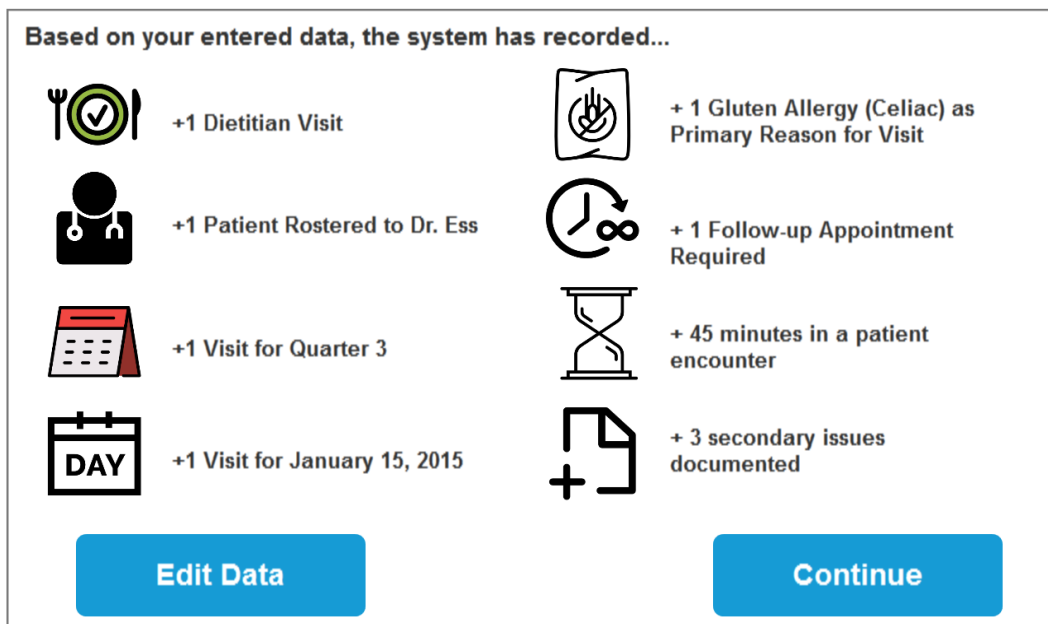
My first design uses the self-monitoring principle, which states that “a system that keeps track of one’s own performance or status supports the user in achieving goals” (Oinas-Kukkonen & Harjumaa, 2009, p. 492). The design attempts to educate users about the quality of their data by providing them with links to learn more about data quality measures and expectation. The design tries to establish a sense of professionalism around data quality by providing a score. The design also invites users to fix poorly scored entries and improve their data quality. The concept is shown in Figure 16.

Figure 16. A scoring mechanism to encourage self-monitoring.



In my second design, I adapt the concept of verifiability and trustworthiness principles to this context by showing users what data they have inputted into the system and confirming how it will be interpreted. I want users to better appreciate the data they have inputted. This design also supports data accuracy and completeness, as it invites users to edit their submission if anything is missing. The use of icons can make it easier to identify specific data characteristics by category. The concept is shown in Figure 17.

Figure 17. A confirmation screen to develop credibility.



### 8.3.3 Goal 3: Increase Data Engagement and Use

Previous studies have found a positive relationship between use and data quality. For example, ABE proved to be an effective tool for improving data quality in primary care (de Lusignan et al., 2006) by providing users with a baseline during meetings, educating users about how data is used and recorded, and establishing goals. Thus, increased attention and focus on data, engagement of stakeholders, and comparisons had positive impacts on data quality. Facilitating these processes would be a good use of PD.

The idea of data use also was supported by the interview results from Appendix F. Data included in the correlations in chapter 7 did indicate a positive relationship between use and completeness, suggesting that improving use would have positive impacts on data within the reporting tool.

#### 8.3.3.1 Persuasion Context

I focused on using information from the junction described in Figure 11. The persuasion context is shown in Table 17; it builds on the framework mapping established in Table 14.

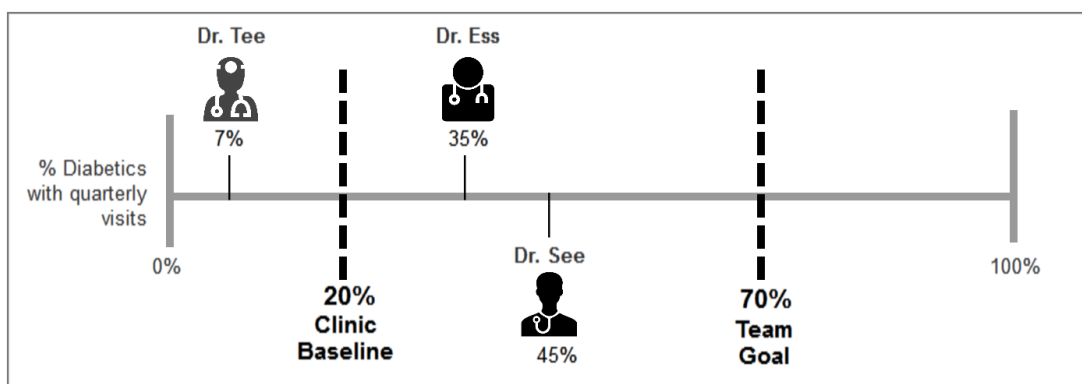
Table 17. Persuasion Context to Enhance Use and Engagement.

Question		Answer	References
Who		Our target users are AHPs entering data into the FHT reporting tool. There are no complex team dynamics or classes of users.	AHPs shown in AH 2.1 in Appendix B.
What		Based on the comparison of Novice, Intermediate, and Expert users, there was an opportunity to have users adopt intermediate user behaviour by interpreting tasks more deeply after reaching the system level of the DL. Engaging with data would influence the Evaluate phase of the control task, where users balance their time and priorities and determine whether they should invest more time to record data accurately. A PD should help users to <i>want</i> to record data accurately.	ConTA 1.1.1 - 1.1.3 in Appendix C. Junctions in Figure 10 in chapter 6.
Why		If users access their data, this will help them understand their present benchmark and relative norms. Entering the data also supports the objective of monitoring population health.	AH 2.1 in Appendix B.
When	Motivation	Users have professional values which will lead them to input data. Users are responsible for meeting organizational benchmarks; failing to report data could result in disciplinary action. Users should have some motivation to want to see their data.	AH 2.1 in Appendix B.
	Abilities	Users need to prioritize their time and engage in time management to use their data (e.g. generating reports takes time). They need time and time management abilities.	Steps 7 and 8 of SRK 1.1 in Appendix E.
	Triggers	Users make strategic choices (and are triggered) by organizational policies, workload, experience, technical abilities, and practice workflows.	Comparison of StrA in Table 2.

Question		Answer	References
How	Message	My message needs to lead users to think more deeply about their data. I want users to understand how the data could be used to improve population health. I want users to engage with their data.	Persuasion Context Analysis
	Route	Direct.	Persuasion Context Analysis
Strategy		Based on the success of previous studies, social support design principles could be highly effective strategies for getting users more engaged with their data. Social comparison, normative influence, social facilitation, competition, and recognition could each be effective design principles.	PSD

My first design uses the social comparison principle, which states that “system users will have a greater motivation to perform the target behaviour if they can compare their performance with the performance of others” (Oinas-Kukkonen & Harjumaa, 2009, p. 495). The goal is to help users understand the value of data from a population health perspective. The design establishes different measures for users and the clinic, and shows a team goal. The design should encourage users to enter data and integrate the task into their workflow to have their information properly reflected. Users should want to ensure that their status vis-à-vis their peers, their clinic, and their goal is as good as possible. The concept is shown in Figure 18.

Figure 18. A comparison of an arbitrary indicator to improve engagement.

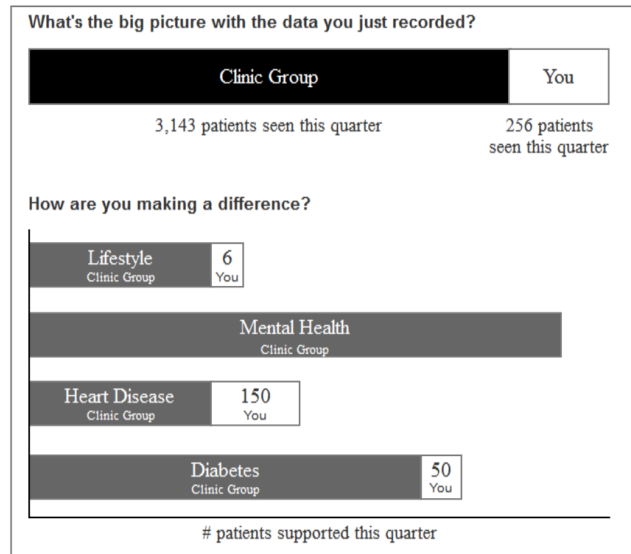


My second design uses the social facilitation principle, which says that “system users are more likely to perform target behaviour if they discern via the system that others are performing the behaviour along with them” (Oinas-Kukkonen & Harjumaa, 2009, p. 495). This display shows that other, similar users are seeing patients and inputting data. It also puts the user’s data into perspective and shows how they have contributed to the team effort. In this sense, the principle cooperation principle, defined as “A



system can motivate users to adopt a target attitude or behaviour by leveraging human beings’ natural drive to co-operate” (Oinas-Kukkonen & Harjumaa, 2009, p. 495), is also used. The design is shown in Figure 19.

Figure 19. Social facilitation of visit volume to improve engagement.



## 8.4 Discussion

### 8.4.1 An Ecological Approach to Persuasive Design

Table 14 links CWA to PD. I call this mapping between the frameworks an ‘ecological approach’ to PD because it uses CWA to identify useful information from an ecological (and cognitive analysis) framework. I am purposefully *not* referring to my approach as a ‘conceptual framework’ because I have merely performed a mapping exercise between my CWAs and the information requirements of the PSD and FBM. My purpose for using a WWWWH approach was to gather useful information and contextualize the data to help a designer (e.g. me) consolidate data and pick an appropriate persuasion principle given the context. After further use and additional testing, Table 14 could form the basis for a broader conceptual framework to support PD.

### 8.4.2 Limitations

My ‘ecological’ approach provides a convenient way to combine CWA, FBM, and PSD. However, the approach is built on the assumption that CWA has systematically analysed the correct breadth and depth of the problem space, and that it has been done in a way that sufficiently supports PSD’s persuasion context. Not all possible tasks were modelled with ConTA and StrA during my CWA. If there was interest in using PD in another area of the codification domain, additional analysis would be

required. In practice, there would likely be a lot of back and forth between modelling work with CWA and incorporating the information into PD. Not all CWA analyses will be compatible with my proposed approach, and CWA models would need to provide sufficient depth to adopt the WWWWH paradigm.

### 8.4.3 What about Ecological Interface Design?

In its original conceptualization and description, EID only used the WDA and WCA phases of a CWA analysis. The phases of ConTA, StrA, and SOCA do not fit the original intent of the framework. In a sense, some phases of CWA are orphaned. Thus, developing a relationship between ConTA, StrA, SOCA, and PDT through my framework is interesting because it extends the relationship between all phases of CWA and design principles. Therefore, my approach is an interesting contribution to the CWA community because it establishes a link to design. It is also interesting to note that the constraints and philosophies behind EID are compatible with PDT; there is no obvious contradiction.

## 8.5 Chapter Summary

### 8.5.1 Key Findings

**Ecological Approach to Persuasive Design:** Since there are very few system analysis frameworks that are intended to provide a persuasion context, I mapped content requirements from existing PD frameworks to various phases of CWA and showed how CWA can inform PD. This is shown in Table 14.

**Design Concepts:** I developed several design concepts that use different design principles. Mockups were presented and were based on my approach to defining the persuasive framework.

### 8.5.2 Connections to Case Study and Research Questions

I have three research questions:

1. In primary care, how are individual users influenced by their environment to input high-quality data?
2. What techniques could be used to design systems that persuade users to enter higher-quality data?
3. Is it possible to improve data quality in primary care by persuading users with the user interface?

In this chapter, I have proposed a technique to design an effective persuasive system by combining CWA, FBM, and PSD. I have shown how this approach could be used to design persuasive elements and which PD principles could change behaviours that would increase data quality.

### **8.5.3 Connections to Other Chapters**

In chapter 9, I develop a mockup for a FHT based on my design concepts from this chapter. The FHT made some adjustments, developed a prototype, and deployed the solution as part of a field study. This provided an opportunity to test the design concepts from this chapter.



## **Chapter 9**

### **Improving Data Quality in Primary Care**

In the previous chapter, I developed several design concepts. In a sense, this chapter is a capstone to my dissertation that ties all elements together: by using modelling work (chapters 3 to 6) and developing data quality baselines and measures (chapter 7), I developed conceptual PDs that are intended to improve data quality (chapter 8). Here, I demonstrate the effectiveness of this process. In this chapter, I present my methods and findings regarding a field study at a FHT and show that data quality can be improved through PD. I was very lucky to have the opportunity to deploy some of my design concepts into a live system and to measure the impacts my intervention had on data quality measures over a period of time.

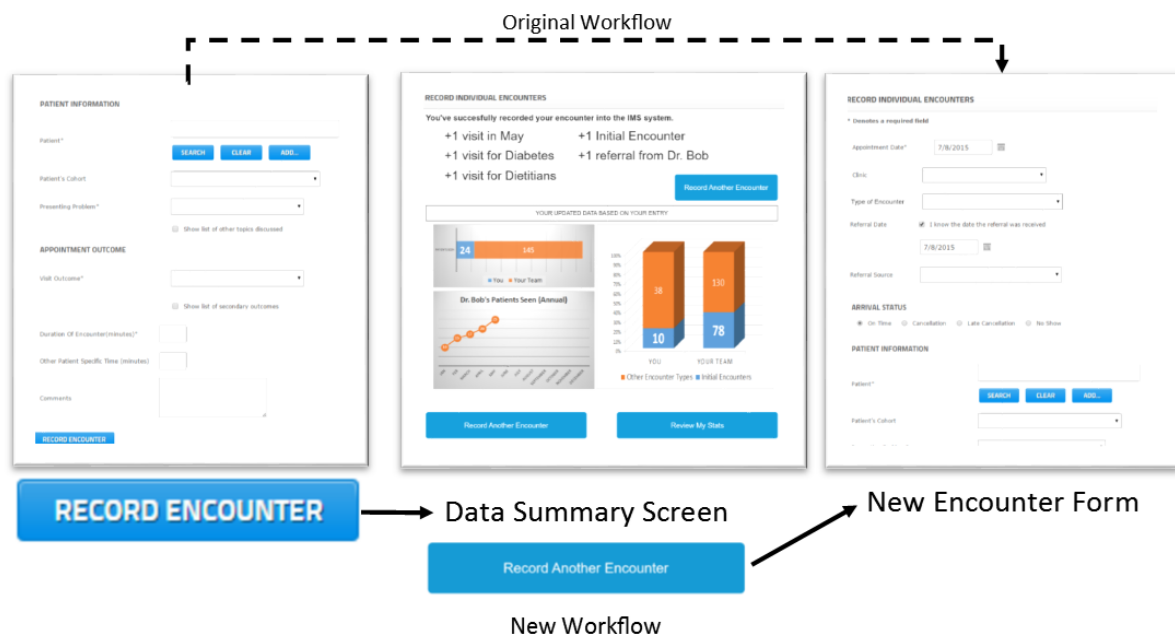
#### **9.1 Field Study**

In chapter 7, I described a case study that used data from a FHT. After the case study, the organization was interested in exploring ways to improve data quality. As a collaborative project, a field study was conducted to measure the effectiveness of the introduction of a new screen on the UI, with included PD elements, on data quality. The new feature was deployed to the same web-based reporting tool described in chapter 7.

Conceptually, the FHT wanted to deploy a minimalistic intervention that would have minimal impacts on workflow and total entry time. As a challenge, the FHT planned to unobtrusively improve data quality without amending fields or overall system behaviour. It was also important that any changes not require re-training. The goal of introducing a minimal change was compatible for research purposes because it would reduce the amount of noise introduced into the system. For example, if the organization had changed the behaviour of input fields or forced users to be ‘complete’ by making all fields mandatory, comparative measures would not be helpful. In a worst-case scenario, forcing users to change their input habits by altering the way data is inputted could introduce input errors and worsen the system’s data quality, as well as cancel any positive impacts of a new UI element. This obviously would be counter-productive to my dissertation and research goals.

Before developing any mockups or prototypes with the FHT, we decided we should try to improve data quality by introducing information into a new summary screen. This intervention would be purely informative, with no interactive components. The new feature would require little user interaction, could be easily skipped, and was to be minimally disruptive. In practice, this meant introducing the summary screen into the data entry workflow. Users would see the summary screen after each entry. Whereas normally users clicking ‘Record Encounter’ at the end of the input form would be brought to a new encounter form, the change showed the summary screen and asked users to click ‘Record Another Encounter’. The new UX is shown in Figure 20.

Figure 20. New User Experience Workflow.



## 9.2 Experimental Design

Conducting an analysis of data from a real system has obvious benefits: any phenomena are observed in context, and the results are more meaningful. Instead of trying to control for an experiment that reproduces results similar to the real world, results can be interpreted based on real data. It is easier to argue that the impacts of an intervention would apply to a real system when impacts have, in fact, been measured on a real system. However, this approach is not without challenge. To measure the effective of the summary screen, I needed to consider experimental designs that were appropriate in the circumstances and respected several constraints.

### **9.2.1 Pre-Post Test**

Most studies of PD designs involve patients who are randomly assigned to control and treatment groups. These studies are designed as randomized control trials (RCTs) and compare exposure to a PD treatment vs. no treatment, or ‘classic’ treatments (such as counselling) to a PD treatment (Lehto & Oinas-Kukkonen, 2011).

From a PD perspective in healthcare, my situation was somewhat unique. Whereas most studies evaluate the effectiveness of PD on patient outcomes, I was looking to measure the impacts of PD on clinicians using a system. Per the results of chapter 7, I knew there was a significant variation in the types of users and program areas and that measuring a statistical difference between two groups would require more users than were available. As well, randomizing groups in an organization could cause confusion and cross-contamination between users working in the same practices. Working with the FHT to communicate with users who would experience the change would be logistically challenging, and it was not clear how questions or concerns could be easily addressed or triaged. A non-randomized approach (e.g. by program area) might resolve some of the logistical challenges, but would make results difficult to interpret because of significant differences between groups.

After consulting with a statistician and reviewing texts specifically related to evaluating user interface changes (see Sauro & Lewis, 2016), I decided to measure the effectiveness of the intervention as a pre-post repeated measures experimental design. With approximately 50 users, a pre-post measure using a paired t-test would provide statistically significant results and provide some evidence of changes in data quality over the two periods. I averaged the data for users eight weeks before and after the intervention and used this as the basis for my analysis.

### **9.2.2 Statistical Process Control**

While a paired t-test might demonstrate the impacts of PD in a system, noise within the data might skew results positively or negatively. This is because the data is coming from a complex socio-technical system where a number of variables could change over the course of a 16-week study and impact data quality measures. For example, management meetings, programming changes, organizational behaviour, strategic direction, and management priorities could easily introduce changes over the course of months. As well, it should be expected that patient volumes and care needs fluctuate seasonally (e.g. higher volumes for the flu in the winter, and lower volumes for assessments around the holidays as staff take vacation). Thus, a pre-post analysis of an intervention may be appropriate and insightful, but it might not be sufficiently compelling; results could be attributed to normal change, noise, or a ‘lucky break’ in selecting a study window that happened to produce good experimental results.

Based on the results from chapter 7, it is obvious that I could have access to a large amount of historical data to help measure the effectiveness of my PD. Classically, an approach to understanding some of these complex situations would be to analyse data experimental results against historical data. A multi-factor mixed-effect ANOVA might be able to explain the impacts of time (e.g. year, month, week) on dependent variables and be separated from an intervention value; changes could be attributed to the intervention and a combination of several other factors. In practice, with real data, the drawback to this approach is that it assumes all noise inducing variables can be identified, and that all of these variables are presently and historically measurable with existing data points. Lacking a catalogue of weekly measures to represent complex socio-technical dimensions and the user's environment, a mixed-effect approach was not especially practical, statistically valid, nor insightful. As a novel approach to this problem, I decided to borrow from healthcare quality improvement literature and use Statistical Process Control (SPC) as a measurement instrument (Provost & Murray, 2011; Thor et al., 2007).

The notion of SPC is to measure process variance in two categories. The first type of variance in SPC is chance variation (also known as common cause variation). This category of variation is caused by phenomena that are always present within a system. This variation is a normal process change that should always exist within a system, and the impact should be measured within predictable limits. Generally, chance variation is anticipated noise associated with normal system operations. The second type of variance in SPC is assignable cause variation (also known as special cause variation). This category of variation is caused by phenomena that are not typically, or historically, present in a system. This variation will change a process and can be described as a signal or non-random variation. Generally, assignable cause variation is associated with changes to the system's operation (Montgomery, 2001).

A common analytical tool for SPC, the XmR chart, consists of two graphs. The first graph in the chart is a measure of a variable over time (e.g. X). This graph shows the mean calculated value for the analysis period, an upper control limit (UCL), and a lower control limit (LCL). A line graph is shown over a period of time. If values are above or below the control limits, they represent assignable cause variation. Values between the control limits represent chance variation. The second graph in the chart shows the moving range (e.g. mR) between each value in the X graph. A mean value for the period and a UCL are also shown. These graphs represent the absolute value of the change from period to period and can be used to identify significant variation. Variation above the UCL is abnormal (Montgomery, 2001).

Trends and changes within the XmR chart that are within the UCL and LCL can be information. If an upward trend occurs starting at a specific time interval, it can be useful to attribute these changes to an assigned cause. For example, if an annual chart shows an increase or decrease in performance at fiscal year end, the assigned cause may relate to a financial incentive or organizational pressure. Identifying



different assigned causes can provide insights to a process and be quite interesting (Montgomery, 2001). Many examples of the use of SPC in healthcare (Thor et al., 2007) are available.

SPC is intended to be used when measuring data from a real system. Philosophically, it can capture the challenges and nuances associated with measuring complex socio-technical systems over time. SPC often is used to measure quality improvement by a team. It becomes part of a quality improvement philosophy, and it is used by teams during weekly or monthly meetings to track progress, identify potential signal changes and causes, and improve processes. In my case, SPC lends itself well to contextualizing historical data and provides a suitable approach to contextualizing the results of my intervention. This use of SPC is intended to further support the pre-post analysis and provide evidence that the observed changes are not associated with random noise or normal variations by incorporating historic data into the analysis. I want to show that I am causing a non-random variation and a phenomena of my own making.

### **9.3 Methods**

Several components to this field study existed. First, qualitative research and interviews were used to design and develop the new summary screen. Second, the effectiveness of the new summary screen was determined with a repeated measure study design, and multiple measures were used to quantify its success. SPC was used to obtain the measure outcomes relative to the previous year. Lastly, user feedback was gathered through a survey, which included an adapted survey from the System Usability Score (SUS) measurement tool and free text comments.

#### **9.3.1 Ethics**

The field study was submitted to the University of Waterloo Ethics Board and approved prior to meeting with users, designing changes, or gathering data. The collaboration was set up as a secondary analysis of data and was primarily initiated, developed, and deployed by the FHT. Users were not required to opt into the study, as it was conducted as part of the organization's normal software revision and update cycle. I served a consultancy role to facilitate the design of a summary screen and to assess its impact as a third party. All shared data was anonymized, and the identity of users was withheld by the FHT. I had no direct contact with participants. The study introduced very low risk to study participants.

#### **9.3.2 User Interface Design**

As a primary goal, the FHT was interested in improving the timeliness of their data. The organizations wanted to use PD to reduce the delay between seeing patients and recording data into the system and wanted to improve the ratio of same-day entries. Improving accuracy, completeness, and use were considered secondary objectives of the intervention.

The summary screen was designed in collaboration with the FHT's operations manager and the FHT's information management team. After agreeing on the concept of a summary screen and a new UX (see section 9.1), a mockup including PD elements was developed based on ideas from chapter 8. To facilitate a brainstorming session, the design concept was turned into a worksheet with open-ended questions, graphical elements, and UX ideas. (This worksheet is shown in Appendix I). The worksheet was used during a discussion with the Operations Manager, who took the worksheet to a manager's meeting to gather additional feedback.

A second meeting took place with the Operations Manager and the information management team after there was time to gather feedback; this took approximately six weeks. Based on the discussion of that meeting, a second worksheet was developed and provided to the development team. (A copy of this second worksheet is shown in Appendix J.) This worksheet was used to create a prototype; the process took approximately six weeks. The prototype was shown to the Operations Manager and his team, and several revisions were made until the design met all stakeholders' requirements.

### **9.3.3 Data Quality Measures**

The data quality measures for this field study are shown in Table 18. Although these measures were inspired from the results presented in chapter 7, a 12-month gap existed between the studies. During that time, the FHT refined some of the measures based on feedback, further analysis, and adjustments to business rules. For example, the delay measure in chapter 7 was a 'raw' delay in data entry. In the field study, it was possible to adjust the delay to take weekends and holidays into consideration. For example, if an appointment took place on a Friday and was entered on a Monday, the delay could be calculated as having an adjusted delay of 1 day and a raw delay of 3 days. Both measures are presented in the results. The completeness measure was also simplified by removing the secondary reason for encounter requirement because business rules had changed. The usage measure was useful for logistic regression, but it presented problems as a weekly measure. It was normalized as a ratio between the number of reports generated by a user and the total number of entries. The measure is expressed as reports generated per entry.

Table 18. Field Study Measures.

Measure Name	Definition
Same-Day Entry	Percentage of records that were entered on the same day.
Average Delay Per Encounter	The average amount of delay between a patient visit and entering the data into the system, measured in days. Raw measures and adjusted values are shown to compensate for weekends and holidays.
% Complete	Number of records that were measured as complete. An entry was measured as complete if the reason for visit was not specified as 'Other'. If the visit was an initial encounter, the referral source must be specified.
% Valid	Percentage of records that were measured as valid. A record was considered valid if the appointment date occurred after the entry date, if the appointment date occurred after January 1st 2008, and if the amount of time between the appointment date and entry date was less than 4 months. If the encounter was an initial visit, the time between the referral date and appointment date must be less than 6 months to be considered valid.
Reports Generated per Entry	This is a standardized measure of report generation as a function of entries into the system. Grouped by week, the measure is reports per entry. It is a measure of data usage. The idea is to adjust for high users of the system (e.g. people who record more data would probably run more reports).

The data measurement criteria were programmed into queries in SQL. All data points, including historical data, used the same measure definitions.

### 9.3.4 Deployment

Once the design and measures were finalized, the FHT went through internal processes to test the new summary screen and receive approval from the management team before a final launch. Once all assurances and approvals were received, an e-mail notification with a video introduction to the summary screen was sent to all users of the system. The next business day, the intervention was deployed simultaneously to all users as a hard launch.

### 9.3.5 Data Collection

After a preliminary analysis of the data in early January, a full dataset was provided in mid-February. Since users were known to delay their entries, it was important to determine whether or not the dataset was complete. Obviously, users who were late in entering their data would not be captured in data collection, and this would positively skew the results. As a preliminary analysis, the number of recorded encounters and number of individual users entering data by week were analysed. Based on this data, an average was calculated. Weeks at the end of data export having insufficient data were removed from the analysis. Data was aggregated by user, and averaged by week.

### 9.3.6 Pre-Post Analysis

To measure the effectiveness of the intervention, the field study was set up using a repeated measures experimental design. Measures were calculated for paired-users for eight weeks prior and nine weeks after the implementation. One week was added to the post-implementation period because of an annual holiday break for the organization's employees. The total number of entries (e.g. patient visits) was compared from each period to ensure that no the difference was not significant. The measures for each user were averaged for the pre and post periods. Data for pre-measurement and post-measurement measures were analysed with a paired t-test. P-values, Cohen's D, and power was calculated by using the Real-Statistics Resource Excel plugin available at [real-statistics.com](http://real-statistics.com).

### 9.3.7 XmR Charts

XmR charts were generated with the R statistical software and the *qicharts* package. This was done after reproducing the results of an example (see Wheeler, 2012) with the software to ensure its accuracy. The chart generation was scripted and automated to take data directly from a secondary SQL database that performed the data grouping.

XmR charts were created for variables that were significant in the paired t-test. To give context to the results, the XmR charts were generated with data from the year prior to the intervention, and the period following the intervention. All data points were used to calculate the average, UCL, and LCL values. The implementation of the user interface (UI) change was graphically marked on the XmR chart with a black line and the note "UI CHANGE". Charts were created by breaking down the data by week and by month.

When reviewing the XmR charts, some weeks were known to artificially skew the analysis and were removed. For example, Week 1 in January 2016 only included January 1, which was a holiday. Data generated for this week would have been a significant outlier. As well, Week 53 in 2015 only included 2 working days (after accounting for statutory holidays), and were likely staff vacations during this period. Data from these weeks were outside the control limits but were not associated with a control process and therefore were removed.

When reviewing the results of the XmR charts, several variations could be attributed to assigned causes. Two managers that were familiar with the organization, its culture, and its data initiatives were asked to comment on likely explanations for changes.

### 9.3.8 User Survey

A survey was created and distributed to users in mid-January. Users were invited to anonymously answer structured questions and provide free-text feedback about the web-based reporting tool and the

new summary screen. The purpose was to debrief users and obtain feedback about how the changes influenced the users.

I took a simple approach to analysing the free text responses. I categorized comments by identifying them as either positive, neutral, or negative. I also identified comments that were related to the system's performance outside the context of the new summary screen, and comments that were directly relevant to the intervention. I also flagged answers that discussed the desired behaviour changes, such as timeliness, or elicited an emotional response. After categorizing responses, I referred to them to further understand my qualitative results as appropriate.

The structured area of the survey included 10 questions about the reporting tool that were adapted from the standard System Usability Scale (SUS), and 10 questions that were specifically about the new summary screen. I changed the nouns in the SUS survey to refer specifically to either the existing system, or the new summary screen. The questions adapted from the SUS tool were analysed separately to generate a SUS score for the reporting tool and the summary screen. Confidence intervals of 95% were calculated. The modified SUS questions are shown in Appendix K.

Additional Likert-scaled questions were asked to evaluate the impact of the system on user motivation and perceived data quality. The mode, median, and mean for each question was calculated. The questions are shown in Appendix K.

## **9.4 Results**

### **9.4.1 User Interface**

The summary screen, as designed, implemented, and deployed, is shown in Appendix J.

### **9.4.2 Deployment**

The intervention was deployed on November 28, 2016. The introductory video was seen 40 times by mid-February. The video can be viewed at <http://bit.ly/2lOtXCu>. No complaints or concerns were immediately reported, and the deployment was considered successful and without incident.

### **9.4.3 Data Collection**

Over the course of 2016, each week had an average of 52.7 users inputting data into the reporting tool and an average visit volume of 1198. Based on these weekly averages, user entries and recorded patient volume was assessed for each week after the intervention. The data appeared to be complete and consistent up to the end of January. Data provided in February did not appear to be complete at the time the data was provided and was dropped for further analysis.

To aggregate the data for XmR, I calculated data quality measures for each user for each week, and then averaged the result. This ‘by user’ approach is consistent with the approaches used by Sauro & Lewis (2016) when evaluation UI changes, and it balances the data by users equally. For example, if one user saw 40 patients and recorded them all late, and another user saw 10 patients and recorded them all on time, we would balance the measure: one user would be on time 100% of the time, and the other would be 0% on time; the average would be 50%. The data is slightly more complex to generate from historic records but provides a balanced measure that represents the behaviour of a set of users.

#### 9.4.4 Pre-Post Analysis

The results of the paired t-test for the pre-post analysis of the implementation of the summary screen are presented in Table 19. The data quality measures are based on 53 paired users. The report generation data is based on 58 users.

Table 19. Pre vs. Post Results with Paired T-Test.

	<b>Pre</b>	<b>Post</b>	<b>Change</b>	<b>p-Value</b>	<b>Power</b>	<b>Cohen D</b>
<b>Average Delay Per Encounter, Raw (n = 53)</b>	5.8	2.4	-3.4	0.006	0.815	0.393
<b>Average Delay Per Encounter, Adjusted (n = 53)</b>	2.8	1.4	-1.4	0.014	0.779	0.375
<b>Percent of Records Entered Same Day (n = 53)</b>	62.9%	73.2%	+10.3%	< 0.001	0.996	0.632
<b>Percent of Records Measured Complete (n = 53)</b>	86.4%	81.6%	-4.8%	< 0.001	0.978	0.545
<b>Percent of Records Measured Valid (n = 53)</b>	98.9%	99.6%	+0.7%	0.045	0.537	0.282
<b>Reports Generated Per 100 Visits (n = 58)</b>	5.2	7.2	+2.0	0.151	0.289	0.193
<b>Average Number of Reports Generated (n = 58)</b>	5.8	7.1	+1.3	0.411	0.124	0.109

The average number of entries per user for the pre-period was 336.62 and for the post-period was 314.31. The difference of 22.31 entries was not significant ( $p = 0.231$ ).

### 9.4.5 XmR Control Charts

XmR Control Charts are shown in Figure 21 to Figure 26 for each data quality. The data was also grouped by month where possible, as shown in Appendix L.

Figure 21. XmR for Timeliness Measure (Recording Delay, Raw).

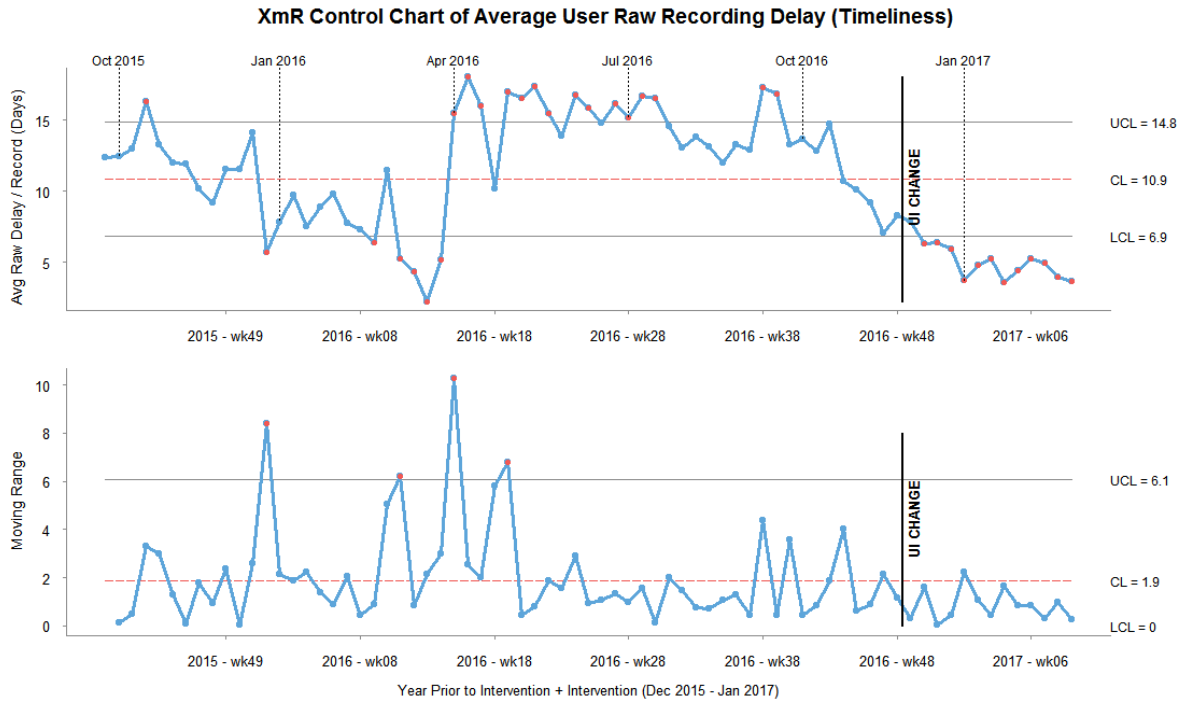


Figure 22. XmR for Timeliness Measure (Recording Delay, Adjusted).

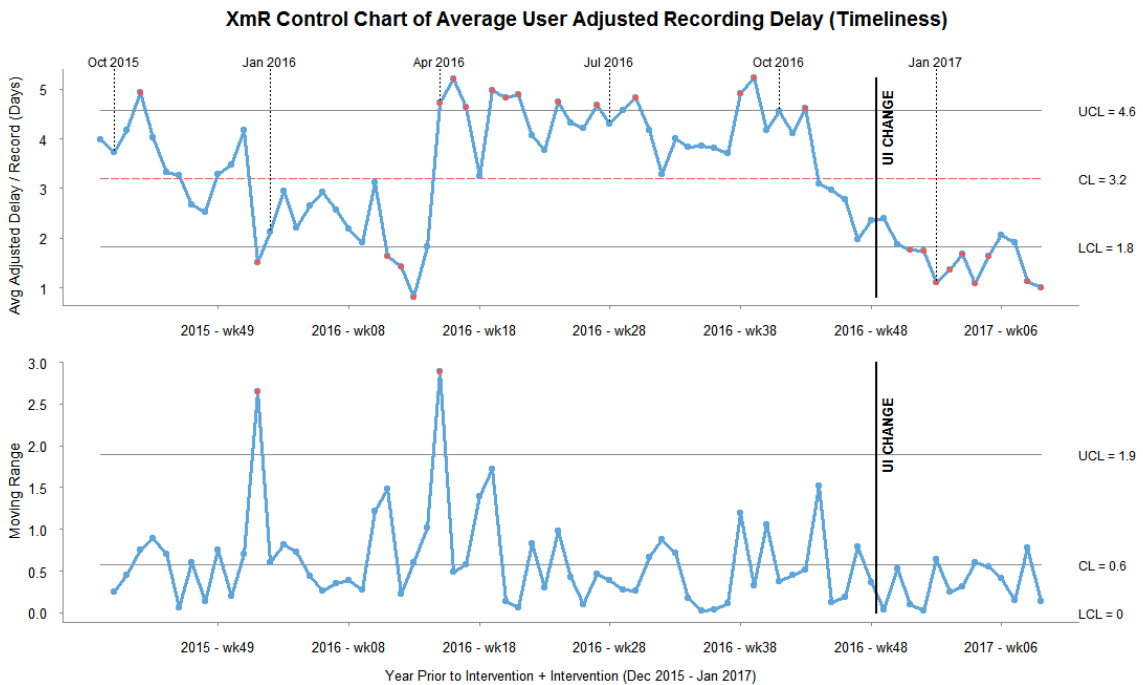


Figure 23. XmR for Timeliness Measure (Percent Same-Day Entry).

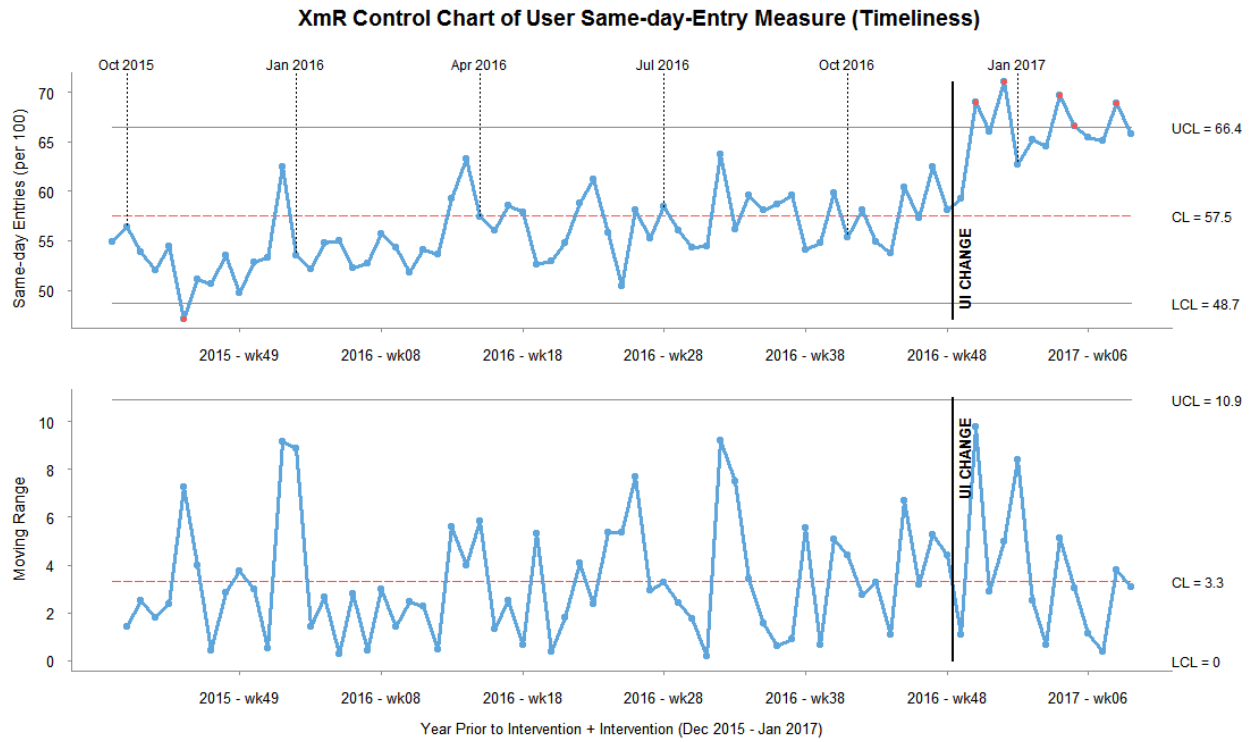


Figure 24. XmR for Completeness Measure.

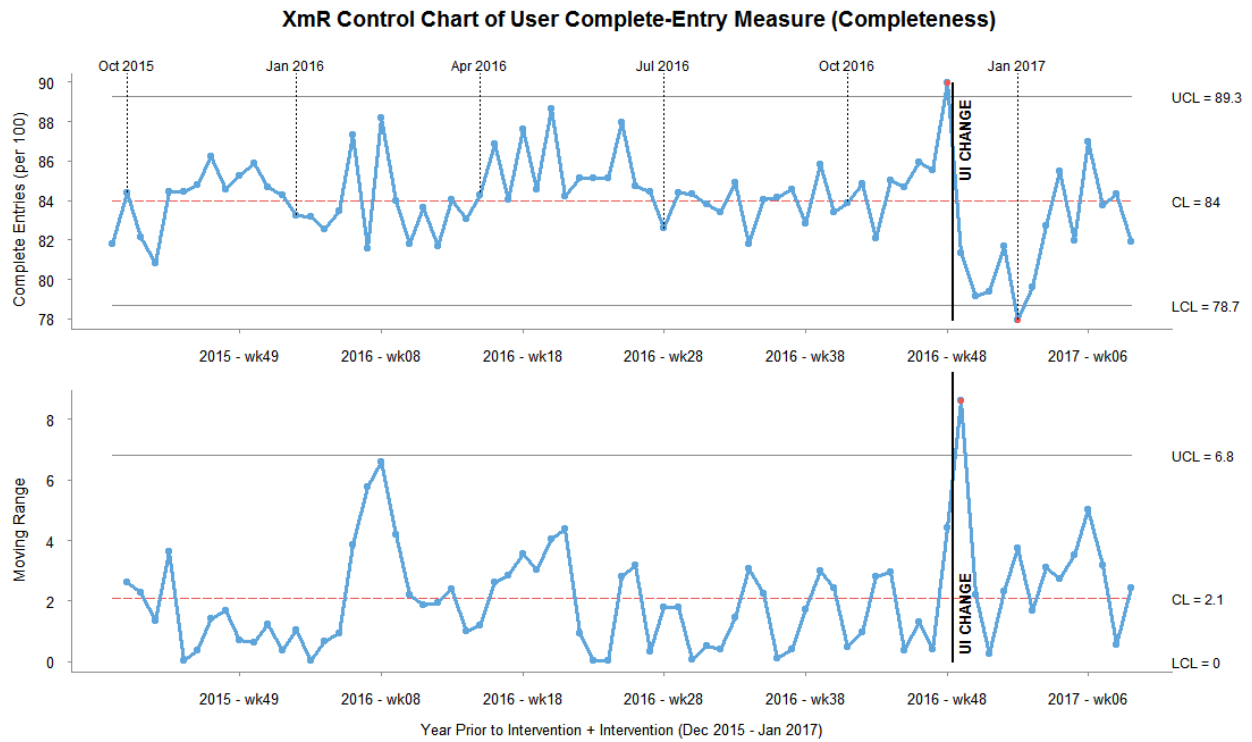




Figure 25. XmR for Validity Measure.

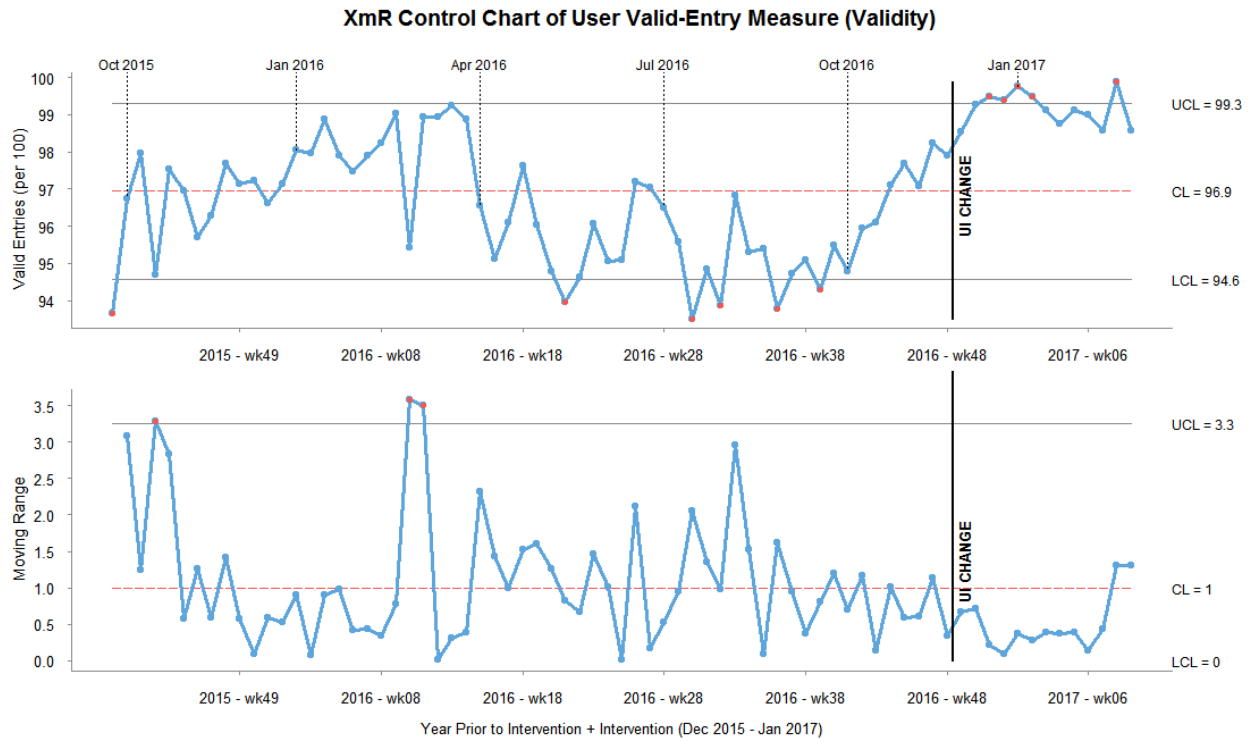
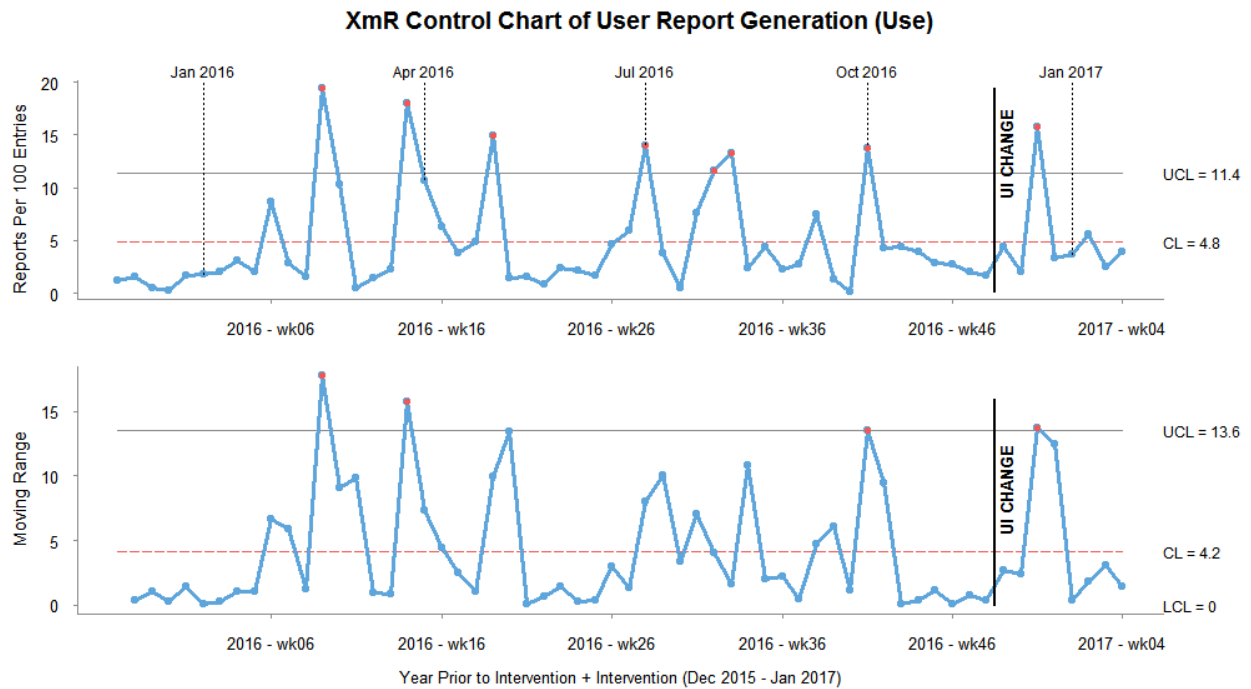


Figure 26. XmR of Use Measure.



### **9.4.6 User Feedback**

Seventeen users responded to the survey. A summary of the answers and a copy of all text comments is provided in Appendix M.

The 95% confidence intervals of the SUS score for the reporting tool was 58.10 – 70.90. This was based on 15 users, since two users did not answer all questions in the SUS survey questions relating to the reporting tool. The 95% confidence intervals of the SUS score for the new summary screen was 57.84 – 68.63. This was based on all 17 user responses. The interpretation of these SUS scores is discussed later.

All 17 respondents answered the other Likert-scale questions. For each question, average, 95% confidence intervals, median, and mode were calculated and included in Appendix M.

## **9.5 Discussion**

### **9.5.1 User Interface**

#### **9.5.1.1 Description of final content**

In the final design of the summary screen, the user has to scroll through the content of the summary screen before seeing a second ‘Record encounter’ button and being able to submit more data. The summary screen was divided into three sections.

The first section was called ‘Your Updated Data Based On Your Entry’. This section provided a confirmation of some of the recorded data (e.g. ‘+1 Visit for the month’, and ‘+1 Initial Encounter’). Below this confirmation area was a summary of the total visits recorded for the month. In addition, an ‘Edit Entry’ was placed below the text.

The second section was called ‘How did this change your current reporting statistics?’ This section featured three graphs that management felt presented the most important information they wanted to ensure users were tracking. This included a pie chart that broke down the user’s no-show rate from the last three months, the user follow-up ratio from the last three months, and a graph of scheduled visits (no shows and actual encounters) over the last two weeks. Beneath these charts, users could click “Review my stats” and generate more complex reports. The second section was designed to engage users with their data and encourage use. The intent was to show users their data. The time span of two weeks was created to match the badge calculation. If this time period was too long, users would never be able to effectively improve their statistics. For example, showing that data was late over the course of two months would take two months to clear. If there was a negative period of data, a user would be reminded of their negative measures and might be discouraged from entering their data. If the time span had been too short, it might be too easy for users to improve their scores.

The third section of the summary screen was called ‘Badges and Awards’. The current iteration only shows a ‘Same day’ badge which is programmed to display and reward the percentage of same-day records. Different badges are presented at the 70, 80, and 90 marks. The text provides a current same-day percentage measure. As long as a user remains between 90% and 100% same day, they will keep the ‘top’ badge available to them. This is intentional to allow users to keep their badges even if they miss one or two appointments; the badges are attempting to avoid discouraging users. This section also displayed the percentage of users that enter their data on the same day.

#### 9.5.1.2 Design Evolution

From the original mockup presented to the FHT (Appendix I) to the final implementation (Appendix J), several notable changes were seen through the consultation process.

First, the consultation resulted in the removal of social support design principles. In the original worksheet, users were compared to their team. This was the adoption of the normative influence principle and was designed to make users have the same goal. Through discussion and debate, the FHT determined that the culture of the organization was not ready for comparisons of any kind. They considered that individual users would be upset by their data if it was low and did not embrace the idea of the social facilitation principle.

Another change occurred in the first section. Originally, a confirmation of all inputted data elements was intended to appear. Instead, the FHT wanted to highlight only three specific data attributes that were important for users to track and watch. This change simplified the display for users and drew their attention to the most important metrics, from a management perspective.

Finally, the original mockup included badges to encourage timeliness, same-day entries, and use. These badges were chosen as examples, and the FHT was encouraged to brainstorm other potential badges that might be appropriate. The FHT decided to start with a single badge as a proof of concept that drew attention to their primary objective of reducing delay.

Overall, the original mockup and final implementation are comparable, and the insights from the organization and decisions of what to keep and what to reject are insightful and interesting in and of themselves.

#### 9.5.1.3 Links to Previous Design Concepts

Despite evolution of the design, several of the original PD concepts were still visible.

4. The credibility support principle, as shown in Figure 17, is visible in a simplified form in the first section.

5. The graphs in the second section of the summary screen are similar to the concepts for social facilitation in Figure 19, but without the social comparison. In its present form, the concept is better described as the primary task support self-monitoring principle.
6. The third section features badges, which is an implementation of the praise dialog principle from Figure 13.
7. The “Did you know” comment at the bottom of the third section is an implementation of the social facilitation principle, as described in Figure 15.

Many of the design concepts developed in chapter 8 appeared on the summary screen.

#### 9.5.1.4 Unexpected Impacts

Anecdotally, there was one unexpected impact from the use of badges, which represents a paradox. The challenge is that the system cannot know how many patient visits a user was supposed to have over a period of time. Therefore, if all the current data from a user was entered on the same day over the previous week, their badge would score them at having a 100% same day. However, say a user did not enter their data for two days: upon signing into the system, they would still see that 100% of their data had been entered on the same day based on their previous performances. In this use case, entering data (the desired task and behaviour) would *drop* their score. Therefore, the PD is incentivizing *against* entering data, because the PD suddenly provides a negative feeling, instead of a rewarding feeling. The negative feeling could be nicely aligned with FBM as a curious negative motivator. However, the user has no opportunity to avoid the negative impact, as they have missed their window. It is unclear if they would be better in the future because of this situation, or if they would be less inclined to pay attention to the PD.

#### 9.5.2 Time Interval for XmR

For generating the XmR graphs, I calculated the mean and the control limits based on data from the year prior to the implementation of the summary screen. This made sense, as I wanted to capture the normal variability associated with a one-year cycle, such as the impact of quarterly reporting, patient volumes, staff vacation, etc. At the same time, I did not want to capture more than a previous year because the system and organization has had an evolving mandate. I argue that in this situation a previous year is a reasonable timespan, and that this choice balances several constraints and measures ‘normalcy’ in the measure and provides an opportunity to diagnose meaningful changes in the process. My results are consistent regardless of the time intervals selected, and regardless of the breakdown by week or month (See Appendix L).

### 9.5.3 Data Analysis

In the pre-post analysis, there were statistically significant changes to the same-day entry percentage, accuracy percentage, completeness percentage, and average delay. The impact on data use was not significant.

#### 9.5.3.1 Timeliness Measure (Delay in Days)

According to the pre-post analysis, the intervention reduced the raw delay in days from 5.8 to 2.4 days. This is a 60% reduction in the delay. The test was statistically significant ( $p = 0.006$ ) with a power of 0.815. The Cohen's D of 0.393 would be considered a medium-to-high effect size. The adjusted delay was reduced from 2.8 to 1.4, which is a 50% reduction. This test was also statistically significant ( $p = 0.014$ ) with a power of 0.779 and an effect size of 0.375.

The XmR Control Charts for this measure is probably the most interesting because there are clearly assignable causes of change. The change that is immediately obvious occurs in April 2016. Two points were below the LCL, and within two weeks, three points were above the UCL. As well, between April and July 2016, 10 of the 14 points were above the UCL. Between July and the UI change in late November, only one point is outside the range. After the UI change, eight of the nine data points were below the LCL. There appears to be a downward trend from October 2016 to January 2017, which intersects the UI change. While it appears that the UI change contributed to the reduction of the average delay, the downward trend suggests another process change may have been taking place. I reached out to a Director at the FHT to understand potential causes for the changes seen in the XmR chart, as well as any other potential processes that may have been at play in the months prior to the UI change.

During conversations with the FHT, it became clear that many changes in delay were aligned with fiscal reporting periods. In April, the FHT needed to submit a report to the MOHLTC and therefore required timely data from all users by the end of March. Managers throughout the organization were asked to remind users to input their data on time (FHT Director, personal communication, February 15, 2017). This explains the values below the LCL in March. When asked about the jump and values above the UCLs in the third quarter of 2016, the Director explained that “the organization participated in a Ministry initiative by which we didn't have to submit a Q1 report [in 2016] (April-June) and [instead] submitted a combined report for both [reporting quarters at the] end of September [...] This is why you see that delay in May-July period”. Furthermore, the Director said that “I put a focus on data [in the reporting tool] in August” and “started putting data quality results on [...] on all team's meeting agendas in both September and October highlighting [average same day entry percentage measures]”. The Director thought that the summary screen had made an impact on reducing delay (FHT Director, personal communication, February 15, 2017).

The comments from the Director clarify several patterns in the data. The improvements in March and the degradation in April were due to reporting due dates with the Government. The high delay between April and July can be attributed to the organization not needing to submit a report and therefore paying less attention to data entry. The gradual reduction starting in September can be attributed to bringing greater attention to delay during meetings. Despite this information, there is still evidence in the control chart to suggest that UI change had a significant impact on the delay; the results prior to the UI change were within normal ranges and have historical precedents. The data after the UI change represents an unusual improvement, with a majority of points falling below the LCL. The XmR charts support the idea that there was a reduction in delay, but the result from the paired t-test may be skewed positively as it included a combination of processes. Attributing the entire reduction in delay to the UI change is probably unreasonable.

The data supports the idea that users are operating within a complex socio-technical system and that data quality measures can be impacted by environmental factors, such as reporting periods and management priorities. The data suggests that users change their behaviour accordingly. Interestingly, these observations support the ideas extracted from StrA that suggested different processes and entry patterns were a function of organizational behaviour (see Appendix D).

The historic data prior to the intervention is almost as interesting as the results of the intervention. It demonstrates that pressure and persuasion has been applied previously in a cyclical way by the organization's management team, and it shows that users have reacted to persuasion. PD is expected to be a better persuader by virtue of its constant pressure. The results suggest that PD can reproduce and augment normal organizational persuasion.

#### 9.5.3.2 Timeliness Measure (Same-Day Entry)

According to the pre-post analysis, the intervention increased the percentage of same-day entries by 10.3%. The test was statistically significant ( $p < 0.001$ ) with a power of 0.996. The Cohen's D of 0.632 would be considered a large effect.

The XmR chart tells a compelling story. Prior to the change, the system signal was relatively stable. Small spikes occurred prior to the end of each quarter, which is consistent with the assignable causes in those time frames from the delay data. Prior to the implementation, there were no obvious trends and there were no out-of-control signals. After the change, a majority of the points were above the UCL, and all points were above the previous average. Based on the results of the XmR charts and the paired t-test, the evidence is compelling that the intervention increased the number of same-day entries within the system. These results could be attributed to PD components on the summary screen which were specifically designed to encourage this behaviour.

### 9.5.3.3 Completeness Measure

According to the pre-post analysis, the intervention decreased the percentage of complete records by 4.8%. The test was statistically significant ( $p < 0.001$ ) with a power of 0.978. The Cohen's D of 0.545 would be considered a large effect.

The XmR control chart shows that the completeness measure was a relatively stable measure over the previous year. No remarkable spikes or changes occurred, and no obvious trends or out-of-control signals were seen. The week prior to the implementation, the completeness hit a high-point. It is not clear why this occurred, but a single measure above the UCL is not necessarily significant (Montgomery, 2001). At the UI change mark in the XmR chart, a significant impact on the completeness measure is seen. The moving range value is above the UCL, which indicates that the impact and change from period to period was significant. One value after the implementation was below the LCL. Interestingly, it appears that after the initial 'shock' of the change, the completeness variable appeared to be returning to normal, fluctuating above and below the average, similar to the previous year. The results of the t-test show a drop after the implementation and an average reduction in the 8 week window. The XmR presents more data after the intervention and indicates stabilization. In this sense, the XmR provides more context to the results and provides additional insight about the results.

The measure has two components: whether the 'Other' option was selected, and whether the referral date was entered. The changes in this XmR do not align with organizational reporting quarters and likely are attributable directly to user behaviour. I contacted the FHT to enquire about which measure may be more significant in this variance. The data analyst responded that "it appears it's because of the referral issue. Out of a total of 3,670 initial visits after the implementation, only 250 had a reason for visit that had 'other' [as the reason for the encounter]. Of those same 3,670 initial visits, [approximately] 1,870 of them had a referral date of 1900-01-01 or earlier" (FHT Data analyst, personal communication, February 16, 2017). I enquired how it was possible for users to enter a date of 1900-01-01, as this seemed quite odd. "There is an option where you can [click] 'I don't know when the referral source was' [and the system therefore records] the date of 1900-01-01. [We implemented this feature because] staff pushed back saying that they don't always have the referral date handy so they need [the 'I don't know'] option. [...] If [users] leave the default option of 'I know when the referral date is' [the system] forces [users] to [enter] a date" (FHT Data analyst, personal communication, February 16, 2017). Thus, it appears that immediately after the summary screen introduction, the complete measure was reduced because there was a statistically significant change in the amount of entries recorded with the 'I don't know the date' instead of entering the referral date for initial encounters.

It is very interesting that a passive change to the user interface (e.g. a non-interactive summary screen) changed user behaviour in this way. The summary screen appears *after* users enter the information

and select referral details. This may represent a reaction to the ‘Same Day’ badge on the summary screen: in order to hit this metric as quickly and easily as possible, users abandon the referral date to optimize their time. To counter-balance this behaviour, a ‘Completeness’ badge may be appropriate.

The behaviour changes in this case seemed less sensitive to environmental factors, and related more strongly to individual user behaviour. The changes could be attributable to task flow, user prioritization, and cognition. Based on discussions with the FHT, it does not appear that this reaction was orchestrated; it appears that a statistically significant number of users responded independently to this stimulus in the same manner. This gives support to some of the findings of the ConTA analysis from the first part of the dissertation (see Appendix C and Appendix E) which described time management behaviours and skills, and predicted the idea of users balancing task value against entry task. In the future, it would be interesting to correlate these behaviours with measures of expertise (e.g. length of employment).

The drop in the completeness measure, and its recovery, is a curious and interesting result. As a hypothesis to explain the data, it is possible that users were pressured to ‘catch up’ in their data entry immediately after the intervention; users would need to clear a ‘backlog’ of data entry to reduce their entry delay for future visits. This means that there would be a higher volume of work and data to enter immediately after the intervention, which would require a short-term change to entry-behaviour to adapt to the new goal. Given a higher volume to meet the entry delay goals, and without additional time to enter data, users likely cut corners to reduce their overall entry time and clicked the “I don’t know” radio button to reduce their overall work time. This hypothesis would be consistent with the time management phenomenon that was measured and modelled earlier in my dissertation. The drop in the completeness measure appears short lived and appears to resolve itself.

#### 9.5.3.4 Validity Measure

According to the pre-post analysis, the intervention increased the accuracy measure by 0.7%. The test was statistically significant ( $p < 0.045$ ) with a power of 0.537. Rejecting the null hypothesis should be done cautiously. The Cohen’s D of 0.282 would be considered a small-to-medium effect.

The XmR control chart shows that the validity measure was stable during the study period (e.g. eight weeks prior to and after the UI change); the data looks symmetrical over the UI change marker, with a small change and some mild improvements. This is consistent with the results of the paired t-test. Looking over the previous year in the XmR chart, the accuracy measure seems erratic. It does not appear to follow quarterly reporting periods, and there are out-of-bound signals from week to week. When broken down by month (Appendix L), the XmR appears much more stable and less erratic, but still lacks



any clear patterns or trends. These results support the concerns about this measure from chapter 7. The current data does not show a meaningful improvement to this data quality dimension.

#### 9.5.3.5 Use Measure

According to the pre-post analysis, no impact on report use was seen. The XmR control chart shows that the use measure was stable during the year prior to the UI change, and that the UI change had no noticeable impacts on trends or patterns. The results of the t-test and the XmR control are consistent. Users did not generate more reports as a function of the summary screen and were not encouraged to look at their data more or less often.

### 9.5.4 User Feedback

The survey provided an opportunity to understand user perspective about the new summary screen. A majority of respondents had used the system for at least two years and reported signing into the system at least on a weekly basis.

#### 9.5.4.1 SUS Scores

The reporting tool's SUS average score was 64.5, with confidence intervals ranging from 58.10 to 70.90. The new summary screen average was 63.24, with confidence intervals between 57.84 and 68.63. The scores clearly overlap and are very similar. No significant difference is seen between the current system and the summary screen.

Trying to assign subjective measures to these scores can be challenging; a SUS score of 64.5 should not be interpreted as a 64.5 out of 100. SUS scores are better measured through percentiles compared to other studies and SUS scores. Bangor et al. (2009) associated subjective rating scales with SUS scores to help practitioners interpret individual scores. For a web application, a score of 64.5 falls below the average and would receive subjective rating of 'OK' or 'marginal'. On a letter grade scale, the results would translate to a D (Bangor et al., 2009).

These results have two take-aways. First, the summary screen did not represent a degradation of the overall system's usability, but did not meaningfully improve the UX either. Second, there are opportunities for improving both the web-based reporting tool and the new summary screen. Users provided helpful suggestions in the free text comments that identify clear usability issues that could easily be addressed.

#### 9.5.4.2 Miscellaneous Likert-Scale Questions

Assuming that a neutral rating average is 3 on the Likert-Scale, all of the non-SUS questions had 95% confidence intervals that overlapped a neutral agreement rating. The only notable result about the

reporting tool is the question ‘I feel like the data I am entering into the reporting tool is collected and used for a meaningful purpose,’ which was skewed toward agreement. The only question that was skewed toward agreement is regarding the new summary screen, ‘it took more time to enter data into the system because of this screen’. Thus, there is a performance problem with the summary screen. This is not totally surprising, since an additional screen has been inserted into the workflow for data entry. A mild agreement with this statement is reasonable.

Generally, users appear neutral about the new summary screen: they appear to be neither enthusiastic nor angry about the intervention. Since the summary screen was intended to be a passive introduction into the system, these results are positive and encouraging.

#### 9.5.4.3 Free Text Comments

Many interesting responses were in the free text area of the survey. Five comments mentioned concerns about the performance impacts of the summary screen, including ‘adds time and doesn’t change practice’, ‘the summary screen added lag time and [made the process] more cumbersome’, ‘[the extra time required] really adds up!’ and ‘it feels slower to load pages and enter data’. Based on these comments and the Likert-scale answers, there does appear to be a legitimate concern about the performance of the system. The FHT received this feedback and is presently working on adjusting their queries with table-valued functions to reduce the load time by 80% (FHT Data analyst, personal communication, February 16, 2017).

In other comments, users provided suggestions for UX and user interface adjustments to the summary screen. For example, two users suggested having the summary screen appear only once a day, instead of after every encounter, or enabling a daily, weekly, and monthly view. These suggestions are not unreasonable and could easily be implemented by the FHT. Taking the summary screen out of the workflow would address all of these concerns, but it is not clear if this continues to provide the same effect on user behaviour.

Several users articulated positive feedback and gave the intended behaviour change heuristics the summary screen was intended to encourage. For example, ‘I find I am now entering stats every 7-9 days instead of every 9-14 days’, ‘I like seeing the graphs - I’m a visual person and this helps to summarize what I view as important info about my practice’, ‘It was nice to see incentives on the screen of reaching goals and receiving badges’, and ‘I feel it could act as a motivator to those who have not [been timely] in the past’. Some users suggested there was only an initial impact with comments such as ‘at first it helped somewhat; now I again rely on my own motivation’. These comments align with the PSD proposed by Oinas-Kukkonen & Harjumaa (2009): there are different kinds of behaviour change (e.g. one time, short term, long term), and different kinds of interventions are appropriate for each. While it appears clear that

the summary screen introduced a change in behaviour, further work will be required to properly categorize the change as either short or long term.

Other comments in the survey were concerning. One user reported that it felt like there was new pressure on data entry. This is not an incorrect impression, but associating pressure to enter data with the summary screen was unexpected. Based on comments from the Director at the FHT, there had been reminders to staff pre-dating the implementation of the summary screen (FHT Director, personal communication, February 15, 2017). By this comment, it appears some users saw the summary screen as an accentuation of management reminders to enter data on time, and had a negative reaction. This is further described by another respondent who said the summary screen made them feel “anxious and unhappy”, and complained that “the summary screen just makes me anxious”.

Anxiety and unhappiness from users are very strong words. However, the true cause of anxiety is not the summary screen or the data, but the user’s performance and statistics. Specifically, the user complained that the summary screen caused anxiety because the system reminded them that they had no-show visits on their record. This would be akin to a student expressing anxiety over seeing their grades posted on a learning management platform. Regardless, if users feel that the summary screen is tracking their progress closely as a proxy manager, it is understandable that performance tracking could cause anxiety. This correlates with some of the historical findings of the XmR for delay, which showed that increased pressure reduced delay; however, this comment suggests the adaption was not necessarily a positive response to management’s oversight.

In terms of improving the summary screen, a few design heuristics may help alleviate some of this anxiety. Currently, the data provided is only a measure of a single user’s data. This was intentional, as management did not feel that the organization was ready to compare users; management had anticipated negative outcomes to users seeing their performance compared to other users. However, in this case a comparison might help alleviate some of the anxiety of this user by normalizing their results. If a user is worried about their performance, would it not be helpful for them to see the performance of other, similar users? A comparison paradigm could help build a user’s confidence, compliance, and engagement, and reduce potential anxieties about their own data. This concept is incorporated into part of the PSD model, which describes normative influence and social facilitation as design principles. However, the impact of these principles would have a different response. Regardless, this idea would not be difficult to incorporate into the summary screen.

Contrasting responses were given regarding the summary screen. Whereas some users expressed seeing a carrot, others saw a stick. Based on the data and outcomes, this would be an example where performance and preference are not correlated; it appears performance is occurring where preference is not.

### 9.5.5 Primary Care User Hypotheses

Case studies can lead to a deeper understanding of the subject, generate useful information, and generate hypotheses (Flyvbjerg, 2006). With the results from this field study, my hypotheses about tradeoffs can be improved. Through logistic regression in chapter 7, I developed three hypothesis about primary care users. These hypotheses included (1) Users entering timely data will trade off other types of data quality (e.g. validity), (2) Users have the ability to achieve validity and completeness at the same time, and (3) Users are presently using their data as a necessity to input the data when it is late. This chapter has allowed me to further build on these hypotheses.

The first hypothesis, that users prioritize their time which impacts data quality, was further supported by my results. I did observe tradeoffs and time conservation when the implementation was deployed (per the delay XmR, Figure 21 and Figure 22). However, the users sacrificed completeness (at the intervention, per the XmR, Figure 24). Validity was not significantly affected. In chapter 7, I hypothesized that users entering timely data (e.g. users prioritizing their time) impacted validity and completeness. In this case, I have evidence that this prioritization affects completeness. Thus, more generally, receiving and encouraging timely data in the FHT's reporting system has a cost; I have evidence that completeness and validity can be impacted by timeliness.

My second hypothesis, that users have the ability to achieve validity and completeness at the same time, may require further consideration. During the intervention, validity held but completeness dropped immediately, before recovering after a transient period. The XmR for completeness (Figure 24) is an interesting artifact and suggests that the intervention itself introduced an immediate change in how users behave. Users may have decided to quickly get their data up to date, and took shortcuts to achieve this goal. Thus, accuracy and completeness can be achieved simultaneously if there are no changes in the ecosystem.

Finally, I proposed that users are presently using their data as a necessity to input the data when it is late. Conceptually, the idea was that as delay increased, users would need to run more reports. If this was true, as the delay decreased after the intervention, I would have expected the number of reports generated to drop. However, no statistically significant change in the number of reports created was seen (Figure 26). This hypothesis would require further testing, but presently it is not compatible with my results from this chapter.

### 9.5.6 Limitations

The measures used for data quality were imperfect. This is similar to the issues discussed in chapter 7. On one hand, I've shown how to gather insight from a database source that was never designed for objective data quality measurements. Alternatively, I've shown *the limitations* of trying to measure

data quality in a system that was never designed for objective data quality measurements. The validity measure, for example, was mostly measured as valid and was not a good discriminator. Suggesting that the summary screen had a minimal impact on data validity may not be a correct conclusion because the measure may not fully capture this data quality concept. The same measure was removed from regression analysis in chapter 7.

Another limitation is the number of measurements after the implementation. Whereas eight weeks of measures after the introduction of the summary screen allowed me to show that a behaviour change occurred that was related to data input and data quality, further measures will be needed to qualify this change as either short or long term. This would require ongoing work, as different types of persuasive elements would need to be tested over several years to measure the length of effects.

Several practical benefits to engaging in applied research with a single FHT were experienced. I was lucky to have such an engaged and interested partner to work with on both a case study and a field study. Doing both studies with the same organization, however, may have introduced other signals in the process control that may have amplified some results. For example, I suspected a secondary process at play in the delay measure. It is unlikely that the 60% drop in delay is fully attributable to the summary screen. This secondary process may have been partially induced by the results of chapter 7. In a perfect world, two separate organizations would have been involved to avoid this issue, to better regulate this potential problem, and improve the separation of the results of the summary screen intervention from other processes.

### **9.5.7 Future Work**

As a result of this field study, design improvements and additional evaluations would be both interesting and appropriate.

#### **9.5.7.1 Improvements to the Summary Screen**

Several design improvements would improve the usability of the summary screen and further improve the data quality measures in the system:

1. The current version only has a ‘Same Day Percentage’ badge. This seems to be effective at improving this measure of data quality. As a next step, other badges for other data quality dimensions could also be introduced. The next obvious badge would be a ‘Complete Percentage’ badge that could provide users with additional feedback to improve data quality.
2. Several users suggested adjusting the summary screen so that it does not appear after every encounter. The summary screen could be programmed to appear once per day, or the content

could be moved to the right side of the input screen. This would communicate the information without introducing an additional click in the input process.

3. Introducing careful comparisons to user data could help them better understand the meaning of their data and give better context. For example, if a user has a 15% no-show rate, it would be helpful to understand the clinic's normal no-show rate and their profession's normal no-show rate. This change would likely entail changing the existing pie charts into bar graphs.
4. Several users complained of system performance issues. These should be resolved by programmers to minimize the impact of the new screen on overall time requirements and the recording task.
5. Several users made usability improvement suggestions. They were mostly special use cases that arise from users moving between physical locations. These changes should improve the reporting tool's base usability.

#### 9.5.7.2 Additional Analyses and Evaluation

My results have been quite encouraging and interesting. Several important avenues can be taken to continue investigating the findings:

1. The measures have shown immediate, short-term improvement. Oinas-kukkonen & Harjumaa (2009) suggest that long-term behaviour change is very difficult. Therefore, it will be interesting to see if there are lasting effects, and if those effects last for a long period of time. A return to 'normal' behaviour within several months, or not, would be quite interesting and would provide evidence of long- vs. short-term behaviour impacts. This would provide an opportunity to further debrief users, and adjust the summary screen to maximize the length of the impact.
2. A few other FHTs in the province of Ontario are using the same report tool for statistical purposes. It would be interesting to do a reproduction study and work with an organization that has not had a previous history of data quality improvement with me. Once the summary screen has been through a few more iterations, this will be worth negotiating. This would help understand how much of the 60% reduction in delay is attributable to the UI change.
3. Several UI elements have been deployed simultaneously. It is unclear which specific element on the summary screen is causing a change, or if a combination of persuasive elements is at play simultaneously.

## 9.6 Chapter Summary

### 9.6.1 Key Findings

Several interesting insights result from this study.

**Intervention Evaluation Method:** A simple approach to evaluate a design would be to do a pre-post comparison of a set of users. A paired t-test can provide statistical measures, but in a noisy system it is possible to either have positive or negative findings based on system noise. To compensate for this, I used SPC to put the data into context with a complimentary toolkit. What was interesting was that the results I calculated with a simple t-test matched the results I found in analysing my dataset with SPC. In this case, there was a correlation of results. In addition, SPC provided opportunities for some interesting insights about the system by prompting me to ask questions about events over a period of time. This would not have been possible with a ‘classic’ experimental design. In future studies, SPC could be a valuable tool for assessing the impact of human factors interventions in a real system because it is relatively simple and easy to communicate.

**Simple Intervention with Measurable Results:** The intervention that was deployed into the reporting tool’s workflow made an impact on data quality. Users were only provided with information, and the system’s input areas and business logic were not changed to achieve a result. The summary screen was primarily intended to address the FHT’s most pressing data quality concern, timeliness, and was successful. A simply summary screen presented to users can impact data quality and change behaviour.

**Context & Environment:** The analysis of assignable causes in my entry delay SPC allowed me to measure periodic organizational impacts on data quality. Quarterly trend changes were associated with reporting cycles and organisational culture. It was clear when management was interested in tracking data, and when it was not. This supports the idea that data quality is part of a bigger socio-technical system and that the coding task is impacted by the user’s environment.

**User Tradeoffs:** The drop in the completeness measure was fascinating because, upon investigation, it was caused by users adapting to their constraints in a similar way. Multiple users adapted to the new same-day heuristic by picking the ‘I do not know the referral course’ option. This supports the idea that data quality is part of a bigger socio-technical system and that the coding task is impacted by the user cognition and decision making.

**Neutral Feedback:** According to the Likert-scale questions on the survey, users were relatively neutral about the summary screen. The SUS score for both the reporting system and the system screen could be translated as merely ‘OK’ and could be improved. Other than some performance issues for loading data on the summary screen, the change was neither positive nor negative. The free text

comments suggests that some users saw carrots (e.g. motivators and rewards), while others saw sticks (e.g. pressure to enter data).

**Performance Anxiety:** Some concerning claims said that the intervention made users feel unhappy and anxious about their data. Originally, comparators were not introduced into the system because management felt it would introduce competition between users and cause stress. Comparators could also have the alternate effect of normalizing data for users and reducing their anxiety about data. This idea would need to be further investigated.

### **9.6.2 Connections to Case Study and Research Questions**

I have three research questions:

1. In primary care, how are individual users influenced by their environment to input high-quality data?
2. What techniques could be used to design systems that persuade users to enter higher-quality data?
3. Is it possible to improve data quality in primary care by persuading users with the user interface?

Based on the results of this chapter, I discovered other interesting ways in which individual users are influenced by their environment and was able to measure the impacts of organizational pressure and user decision making. By introducing a new screen and not changing any of the input fields on the reporting tool, I was able to improve data quality by eluding to ecological variables. Thus, data quality is not a technical problem, but a complex socio-technical problem and the output of a series of processes, policies, technology, and workflows.

This chapter has shown that it is possible to improve and change data quality through changes to the user interface. In this case, PDT has been effective at making a measurable change and impacting several data quality measures. PDT was deployed without changing input screens or forcing a task change within the interface.

This chapter has shown that PDT is an effective technique for persuading users. However, this study was not able to capture which element or design concept, specifically, contributed to an improvement. A combination of CWA, FBM, and PSD appear to have supported the design of effective solutions during the case study. There is merit to using PD to improve data quality in primary care.

### **9.6.3 Connections to Other Chapters**

In chapter 8, I linked my work with CWA to my designs, and my design concepts are linked to the results of this chapter. CWA is linked forward to my results. However, my results have also validated some of the findings I incorporated into my CWAs. Specifically, during StrA, I identified that users are triggered to adopt different strategies according to organizational pressure. As well, the choice of



strategies has an impact on input delay. When I saw the results of the XmR for delay (Figure 21 and Figure 22) and received some context for the patterns from the FHT, I was able to validate the previous modelling choices and the association between the results of my chapters.

Another example was my ConTA's characterization of user time management, and how this behaviour manifested itself when the summary screen was deployed. The fact that a significant number of users reacted in the same way and chose to trade off their time against data quality in similar ways was interesting and unexpected.



## **Part D**

### **Conclusion**

Part D of my dissertation provides a conclusion by summarizing my findings and contributions, and discussing the generalizability of my results.



## **Chapter 10**

### **Conclusions**

To conclude my dissertation, I will highlight some key themes and ideas, review my contribution to the literature, discuss some limitations, and describe future research.

#### **10.1 Case Study and Research Questions**

During my dissertation, I used an instrumental case study approach to answer three research questions.

##### **10.1.1 How are individual users influenced by their environment to input high-quality data?**

My case study investigated the contextual conditions relevant to understanding data quality, and how users are influenced by their environment. CWA was an excellent framework to document my findings. I was able to identify and document several ways in which users are influenced by their environment. This was accomplished through a WDA of patient treatment and two CWAs of data codification. Appendices A through E offer a rich description of influencing factors. Interesting insights and comparisons also were discussed in chapter 6 and further describe the ways that the primary care environment is unique compared to the codification environment at a community hospital.

##### **10.1.2 What techniques could be used to design systems that persuade users to enter higher-quality data?**

The techniques that can be used to persuade users span the entirety of the dissertation. CWA, as an analytical technique, provides very useful information and context. Pairing this technique with the PSD through my ecological approach to PD also showed promise. The full suite of techniques include a systems analysis with CWA, a measurement of data quality as a baseline with qualitative and quantitative approaches, and a use of PDT to develop interventions.

##### **10.1.3 Is it possible to improve data quality in primary care by persuading users with the user interface?**

During my field study in chapter 9, I was able to demonstrate that it is possible to improve data quality, specifically the timeliness dimension, by using a PD approach.

##### **10.1.4 Other Findings**

Through my case study, I also gathered other insights about data quality in primary care including:

1. The crux of the data quality challenge in primary care is that clinicians are asked to manage time and balance priorities. Coding data involves more than entering a code into a UI; coding data incorporates several ecological aspects (such as role assignment and training, record ownership, professional values and standards, funding models, and quality improvement initiatives) and user decision making (such as choosing time management strategies).
2. Hospitals code data effectively with dedicated teams, but struggle with providing timely data; HIMPs produce reliable and high-quality data, but real-time data is not a reality.
3. I have observed consistent tradeoffs in data quality in primary care. In this sense, it appears that asking clinicians to enter high-quality structured data can feel like a game of whack-a-mole. Achieving all facets of data quality is very difficult; data consumers will need to prioritize their most pressing data quality concerns.
4. Some very basic data entry issues must be resolved before the era of IBM-Watson and Big Data in healthcare can truly begin.

Through the case study, I have benefited from working with a FHT. It is fortunate that the FHT's reporting system was simple enough to use for longitudinal data quality research but sophisticated enough to offer insights about data quality and coding within a primary care environment. It would have been challenging to study user behaviour and their tradeoffs, and design solutions with PD, as well as deploy intervention in manageable research cycles.

## **10.2 Summary of Contributions**

I have made numerous contributions to the field of human factors engineering and data quality, and have produced a number of papers that are currently under review.

### **10.2.1 Contributions to Human Factors Engineering**

I have made the following contribution to the field of Human Factors Engineering:

1. I demonstrated that CWA is a flexible and versatile tool. My models of patient treatment and data quality are the first of their kind and help broaden the applications and use-cases for CWA.
2. I built on the work of Burns et al. (2004) and developed an approach for comparing CWA analyses. I demonstrated how this approach can be used to extract interesting and unique insights from similar domains and inform design.

3. I built on work by Rezai & Burns (2014) and developed an approach to link CWA, the FBM, and PSD. This ecological approach to PD allowed me to design persuasion elements that were effective at changing user behaviour in a complex context.
4. I demonstrated the use of SPC as a measurement approach in a field study and validated this approach against a more ‘classic’ experimental design.

### **10.2.2 Contributions to Health Informatics and Data Quality**

I have made the following contributions to the field of Health Informatics:

1. I developed a patient treatment AH that adopts a biopsychosocial perspective and can be used to design better EMR systems in primary care.
2. I used CWA to describe the environmental factors that influence data codification and data entry.
3. I developed an approach for developing measures for a primary care reporting system, and reported baseline measures for use and comparison by other engineers.
4. I developed several PD concepts that could be introduced into primary care systems to improve data quality.
5. I built on existing approaches to improving data quality through social persuasion (e.g. de Lusignan et al., 2006; van der Bij et al., 2016) by incorporating them into an effective PDT in a primary care system. This is the first use of the PSD to identify design principles for changing clinicians’ behaviour regarding data entry and is advantageous over previous approaches because the use of technology makes the persuasion consistent and systematic to all users.

### **10.3 Generalization of Findings**

There will be questions about the generalizability of my results. In many ways, my work has addressed the ‘low hanging fruit’ in a FHT reporting system, and forms a proof of concept. I have shown that it is possible to analyse a socio-technical system, and emulate successful social processes with technology to change data entry behaviours. My models were specific to this context, and it is unlikely that a third party could adopt my models and persuasion context verbatim to replicate results elsewhere. In this sense, my work is a recipe and a template for future work. Applying this work to EMR solutions will require assessing the applicability of the CWA models and determining appropriate design goals for a persuasive intervention. Some differences between areas will be trivial; though privacy laws are not identical from jurisdiction to jurisdiction, these differences are unlikely to obfuscate the overall analysis

and findings. Similarly, terminology issues, such as using the acronym EMR or EHR, are not fundamental barriers to understanding and appreciating the work.

Generally, my CWA models are a starting point for understanding other primary care environments, and could be used to understand hospital coding environments. Using my models and findings in Canada and Europe, where there are similar single payer healthcare systems, should not require a foundational rework. However, using my models in the United States may prove challenging. For example, primary care in the United States transforms and codes data based on the requirements of several private insurance companies. This introduces the notion of ‘upcoding’ with respect to intentional biasing of data for economic purposes (Hersh et al., 2013). As my models were developed to represent a single-payer scenario, where physicians are essentially independent entrepreneurs in an uncompetitive monopoly, several nuances and relationships between data, health, economics, and system sustainability could be fundamentally different. Readers from the United States will need to carefully consider the premises of my models, and may have success with incorporating ideas from behavioural economics.

Regardless of the specific jurisdiction and context of future readers, appreciating system differences and understanding how other systems operate could prove very valuable. This is arguably a significant strength of my work and could be interesting on the wider international stage. The idea of understanding international differences is fascinating, and could be an effective use of CWA comparisons in the future to assess generalizability.

## **10.4 Future Research**

Based on the results of my research, some interesting opportunities are provided for further research.

### **10.4.1 Other Areas of Healthcare**

In many areas of healthcare, clinicians are asked to enter and codify data. Some of these situations require the use of a built-in form, or require data entry into a secondary system. For example, in the long-term care sector, nurses are being asked to fill out specific patient assessments to capture data in a format called RAI-MDS (Guthrie et al., 2014; Hirdes et al., 2000). As well, in primary care, it is common to use custom forms to codify data. These systems and use cases mimic quite closely the system I studied in my dissertation. My work has described an approach to establishing data quality benchmarks, comparing workflows and ecologies, and looking for PDs to improve data quality. It could be applied to other areas, which could bring value to the healthcare system.

### **10.4.2 Better Persuasion for Better Data**

My approach to using PDs is at its infancy and will need to develop alongside innovations to EMRs and developments in the PD community. Presently, there are few examples of persuading clinical



users, and the literature is placing an emphasis on consumer side PD. As my results were very encouraging, I would like to continue perfecting the use of PD principles to change the behaviours of clinicians and improve data quality. There are obviously many more concepts and directions to explore. Identifying which concepts are more effective than others would be very valuable work. The issue of short- vs. long-term impacts on behaviours would also be interesting to catalogue and understand.

Based on previous studies, social comparisons might prove to be highly effective; this would need to be explored with partners interested in changing organizational norms and exploring the polarizing issue of user and performance comparisons. Some of this work might involve continuing to work with current colleagues to further develop the current Summary Screen. Some of this work might include finding ways replicate my results within EMRs. Government agencies also might be interested in incorporating PD principles as inexpensive approaches to improving overall data quality in their registries.

Generally, years of work remain to grow to our full understanding of how to incentivize data quality through the use of PDT strategies, as part of a larger and more complex ecosystem.

#### **10.4.3 Comparison of Additional Systems**

I would like to continue developing my comparison approach and work on additional examples. I genuinely believe that some very interesting insights could be transferred between complex systems with this approach. The approach will need to be further refined and perfected, but I believe it could be a helpful addition to CWA when used to compare more disparate systems.



## References

- Ahlstrom, U. (2005). Work domain analysis for air traffic controller weather displays. *Journal of Safety Research*, 36(2), 159–169.
- Alahäivälä, T., & Oinas-Kukkonen, H. (2016). Understanding persuasion contexts in health gamification: a systematic analysis of gamified health behavior change support systems literature. *International Journal of Medical Informatics*, 96, 62–70.
- Alanen, M., & Porres, I. (2003). Difference and Union of Models. *UML 2003 - The Unified Modeling Language. Modeling Languages and Applications*, 2863(527), 2–17. [http://doi.org/10.1007/978-3-540-45221-8\\_2](http://doi.org/10.1007/978-3-540-45221-8_2)
- Arts, D. G. T. (2002). Defining and Improving Data Quality in Medical Registries: A Literature Review, Case Study, and Generic Framework. *Journal of the American Medical Informatics Association*, 9(6), 600–611. <http://doi.org/10.1197/jamia.M1087>
- Ashoori, M., Burns, C. M., D'Entremont, B., & Momtahan, K. (2014). Using team cognitive work analysis to reveal healthcare team interactions in a birthing unit. *Ergonomics*, 57(7), 973–986. <http://doi.org/10.1080/00140139.2014.909949>
- Bangor, A., Kortum, P., & Miller, J. (2009). Determining what individual SUS scores mean: Adding an adjective rating scale. *Journal of Usability Studies*, 4(3), 114–123.
- Bates, D., Maechler, M., Bolker, B., & Walker, S. (2014). lme4: Linear mixed-effects models using Eigen and S4. R package version, 1(7).
- Batini, C., & Scannapieco, M. (2006). *Data Quality Concepts, Methodologies and Techniques*. Springer Science & Business Media.
- Baxter, P., & Jack, S. (2008). Qualitative case study methodology: Study design and implementation for novice researchers. *The Qualitative Report*, 13(4), 544–559.
- Beevis, D., Vicente, K. J., & Dinadis, N. (1998). An exploratory application of ecological interface design to aircraft systems. In *Proceedings of the NATO RTO Symposium on Collaborative Crew Performance in Complex Operational Systems*. DTIC Document.

- Bettencourt-Silva, J. H., Clark, J., Cooper, C. S., Mills, R., Rayward-Smith, V. J., & de la Iglesia, B. (2015). Building Data-Driven Pathways From Routinely Collected Hospital Data: A Case Study on Prostate Cancer. *JMIR Medical Informatics*, 3(3), e26. <http://doi.org/10.2196/medinform.4221>
- Biro, S., Williamson, T., Leggett, J. A., Barber, D., Morkem, R., Moore, K., Janssen, I. (2016). Utility of linking primary care electronic medical records with Canadian census data to study the determinants of chronic disease: an example based on socioeconomic status and obesity. *BMC Medical Informatics and Decision Making*, 16(1), 1. <http://doi.org/10.1186/s12911-016-0272-9>
- Birrell, S. A., Young, M. S., Jenkins, D. P., & Stanton, N. A. (2012). Cognitive Work Analysis for safe and efficient driving. *Theoretical Issues in Ergonomics Science*, 13(4), 430–449.
- Bisantz, A. M., & Mazaeva, N. (2008). Work domain analysis using the abstraction hierarchy: Two contrasting cases. *Applications of Cognitive Work Analysis*, 49–68.
- Bisantz, A. M., Roth, E., Brickman, B., Gosbee, L. L., Hettinger, L., & McKinney, J. (2003). Integrating cognitive analyses in a large-scale system design process. *International Journal of Human-Computer Studies*, 58(2), 177–206.
- Blake, R., & Mangiameli, P. (2011). The Effects and Interactions of Data Quality and Problem Complexity on Classification. *Journal of Data and Information Quality*, 2(2), 1–28. <http://doi.org/10.1145/1891879.1891881>
- Bray, F., & Parkin, D. M. (2009). Evaluation of data quality in the cancer registry: Principles and methods. Part I: Comparability, validity and timeliness. *European Journal of Cancer*, 45(5), 747–755. <http://doi.org/10.1016/j.ejca.2008.11.032>
- Brouwer, H. J., Bindels, P. J. E., & Weert, H. C. (2006). Data quality improvement in general practice. *Family Practice*, 23(5), 529–36. <http://doi.org/10.1093/fampra/cml040>
- Brown, S., Rosenbloom, T. S., Hardenbrook, S. P., Clark, T., Fielstein, E., Elkin, P., & Speroff, T. (2012). Documentation quality and time costs: A Randomized Controlled Trial of Structured Entry versus Dictation. *Journal of Data and Information Quality*, 3(1), 1–17. <http://doi.org/10.1145/2166788.2166790>

- Burns, C. M., Bisantz, A. M., & Roth, E. M. (2004). Lessons From a Comparison of Work Domain Models: Representational Choices and Their Implications. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 46(4), 711–727. <http://doi.org/10.1518/hfes.46.4.711.56810>
- Burns, C. M., Enomoto, Y., & Momtahan, K. (2009). A Cognitive Work Analysis of Cardiac Care Nurses in Teletriage. In *Applications of Cognitive Work Analysis* (pp. 149–174). Boca Raton: CRC Press.
- Burns, C. M., & Hajdukiewicz, J. (2004). *Ecological interface design*. CRC Press.
- Butz, J., Brick, D., Rinehart-Thompson, L. A., Brodник, M., Agnew, A. M., & Patterson, E. S. (2016). Differences in coder and physician perspectives on the transition to ICD-10-CM/PCS: A survey study. *Health Policy and Technology*, 5(3), 251-259.
- Canadian Institute for Health Information. (n.d.). Data Quality. Retrieved from <http://www.cihi.ca/cihi-ext-portal/internet/en/tabbedcontent/standards+and+data+submission/data+quality/cihi021513>
- Carayon, P., Wetterneck, T. B., Rivera-Rodriguez, A. J., Hundt, A. S., Hoonakker, P., Holden, R., & Gurses, A. P. (2014). Human factors systems approach to healthcare quality and patient safety. *Applied Ergonomics*, 45(1), 14–25. <http://doi.org/10.1016/j.apergo.2013.04.023>
- Cavoukian, A., Taylor, S., & Abrams, M. E. (2010). Privacy by Design: essential for organizational accountability and strong business practices. *Identity in the Information Society*, 3(2), 405–413.
- CBC News. (2009). EHealth scandal a \$1B waste: auditor. Retrieved from <http://www.cbc.ca/news/canada/toronto/ehealth-scandal-a-1b-waste-auditor-1.808640>
- Charles, C., Gafni, A., & Whelan, T. (1999). Decision-making in the physician-patient encounter: Revisiting the shared treatment decision-making model. *Social Science and Medicine*, 49(5), 651–661. [http://doi.org/10.1016/S0277-9536\(99\)00145-8](http://doi.org/10.1016/S0277-9536(99)00145-8)
- Chow, R., & Vicente, K. (2002). A field study of emergency ambulance dispatching: Implications for decision support. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 46(3), 313–317. <http://doi.org/10.1177/154193120204600321>
- Chu, W. W., Hsu, C., Cardenas, A. F., & Taira, R. K. (1998). Knowledge-based Image Retrieval with Spatial and Temporal Constucts, 10(6). <http://doi.org/10.1109/69.738355>

- Clay-Williams, R., & Colligan, L. (2015). Back to basics: checklists in aviation and healthcare. *BMJ Quality & Safety*, 24(7), 428–431.
- College of Physicians and Surgeons of Ontario. (2012a). Mandatory and Permissive Reporting. Retrieved October 1, 2016, from <http://www.cpso.on.ca/policies-publications/policy/mandatory-and-permissive-reporting>
- College of Physicians and Surgeons of Ontario. (2012b). Medical Records. Retrieved October 1, 2016, from <http://www.cpso.on.ca/policies-publications/policy/medical-records>
- Collins, C., & Janssens, K. (2012). Creating a General (Family) Practice Epidemiological Database in Ireland - Data Quality Issue Management. *Journal of Data and Information Quality*, 4(1), 1–9. <http://doi.org/10.1145/2378016.2378018>
- Crandall, B., Klein, G. A., & Hoffman, R. R. (2006). *Working Minds: A Practitioner's Guide to Cognitive Task Analysis*. MIT Press.
- de Lusignan, S. (2003). What Is Primary Care Informatics? *Journal of the American Medical Informatics Association*, 10(4), 304–309. <http://doi.org/10.1197/jamia.M1187>
- de Lusignan, S. (2005). The barriers to clinical coding in general practice: a literature review. *Medical Informatics and the Internet in Medicine*, 30(2), 89–97. <http://doi.org/10.1080/14639230500298651>
- de Lusignan, S., Belsey, J., Hague, N., Dhoul, N., & van Vlymen, J. (2006). Audit-based education to reduce suboptimal management of cholesterol in primary care: a before and after study. *Journal of Public Health (Oxford, England)*, 28(4), 361–9. <http://doi.org/10.1093/pubmed/fdl052>
- de Lusignan, S., Khunti, K., Belsey, J., Hattersley, A., van Vlymen, J., Gallagher, H., Majeed, A. (2010). A method of identifying and correcting miscoding, misclassification and misdiagnosis in diabetes: a pilot and validation study of routinely collected data. *Diabetic Medicine : A Journal of the British Diabetic Association*, 27(2), 203–9. <http://doi.org/10.1111/j.1464-5491.2009.02917.x>
- de Lusignan, S., Stephens, P., & Majeed, A. (2004). Does Feedback Improve the Quality of Computerized Medical Records in Primary Care ? *Journal of the American Medical Informatics Association*, 9(4), 395–401. <http://doi.org/10.1197/jamia.M1023>. Although
- de Lusignan, S., & van Weel, C. (2006). The use of routinely collected computer data for research in primary care: opportunities and challenges. *Family Practice*, 23(2), 253–63.

- de Lusignan, S., Wells, S. E., Hague, N. J., & Thiru, K. (2003). Managers See the Problems Associated with Coding Clinical Data as a Technical Issue whilst Clinicians also See Cultural Barriers. *Methods Information in Medicine*, (42), 416–422. <http://doi.org/10.1267/METH03040416>
- Edwards, A., & Elwyn, G. (2006). Inside the black box of shared decision making: distinguishing between the process of involvement and who makes the decision. *Health Expectations*, 9(4), 307–320. <http://doi.org/10.1111/j.1369-7625.2006.00401.x>
- Effken, J. A., Loeb, R. G., Kang, Y., & Lin, Z.-C. (2008). Clinical information displays to improve ICU outcomes. *International Journal of Medical Informatics*, 77(11), 765–777. <http://doi.org/10.1016/j.ijmedinf.2008.05.004>
- Effken, J., Loeb, R., Johnson, K., Johnson, S., & Reyna, V. (2001). Using cognitive work analysis to design clinical displays. *Studies in Health Technology and Informatics*, 84(February), 127–131. <http://doi.org/10.3233/978-1-60750-928-8-127>
- Elwyn, G., Frosch, D., Thomson, R., Joseph-Williams, N., Lloyd, A., Kinnersley, P., Barry, M. (2012). Shared Decision Making: A Model for Clinical Practice. *Journal of General Internal Medicine*, 27(10), 1361–1367. <http://doi.org/10.1007/s11606-012-2077-6>
- Fidel, R., & Pejtersen, A. M. (2004). From information behaviour research to the design of information systems: the Cognitive Work Analysis framework. *Information Research*, 10(1). Retrieved from <http://www.informationr.net/ir/10-1/paper210.html>
- Fisher, C. W., Lauria, E. J. M., & Matheus, C. C. (2009). An Accuracy Metric: Percentages, Randomness, and Probabilities. *Journal of Data and Information Quality*, 1(3), 16:1-16:21. <http://doi.org/10.1145/1659225.1659229>
- Flyvbjerg, B. (2006). Five Misunderstandings About Case-Study Research. *Qualitative Inquiry*, 12(2), 219–245. <http://doi.org/10.1177/1077800405284363>
- Fogg, B. (2003). *Persuasive technology: using computers to change what we think and do*. Morgan Kaufmann Publishers. San Francisco: ACM.
- Fogg, B. (2009a). A behavior model for persuasive design. *Proceedings of the 4th International Conference on Persuasive Technology - Persuasive '09*, 1. <http://doi.org/10.1145/1541948.1541999>

- Fogg, B. (2009b). Creating persuasive technologies: an eight-step design process. *Proceedings of the 4th International Conference on Persuasive Technology - Persuasive '09*. Retrieved from [http://202.154.59.182/ejournal/files/an eight-step design process.pdf](http://202.154.59.182/ejournal/files/an%20eight-step%20design%20process.pdf)
- Fogg, B. J., & Hreha, J. (2010). Behavior wizard: A method for matching target behaviors with solutions. In *International Conference on Persuasive Technology* (pp. 117–131). Springer.
- Gagnon, M.-P., Desmartis, M., Labrecque, M., Légaré, F., Lamothe, L., Fortin, J.-P., Duplantie, J. (2010). Implementation of an electronic medical record in family practice: a case study. *Informatics in Primary Care, 18*(1), 31–40.
- Gellman, B., & Miller, G. (2013). U.S. spy network's successes, failures and objectives detailed in "black budget" summary. Retrieved from [http://articles.washingtonpost.com/2013-08-29/world/41709796\\_1\\_intelligence-community-intelligence-spending-national-intelligence-program](http://articles.washingtonpost.com/2013-08-29/world/41709796_1_intelligence-community-intelligence-spending-national-intelligence-program)
- Gould, B. E., & Dyer, R. (2014). *Gould's Pathophysiology for the Health Professions* (5e ed.). Elsevier Health Sciences.
- Grant, K. (2015). Ontario's curious shift away from family health teams. Retrieved from <http://www.theglobeandmail.com/life/health-and-fitness/health/ontarios-curious-shift-away-from-family-health-teams/article22989363/>
- Guthrie, D. M., Pitman, R., Fletcher, P. C., Hirdes, J. P., Stolee, P., Poss, J. W., Ezekiel, H. J. (2014). Data sharing between home care professionals: a feasibility study using the RAI Home Care instrument. *BMC Geriatrics, 14*(1), 81.
- Hajdukiewicz, J. R. (1998). Development of a structured approach for patient monitoring in the operating room. National Library of Canada= Bibliothèque nationale du Canada.
- Hajdukiewicz, J. R., Doyle, D. J., Milgram, P., Vicente, K. J., & Burns, C. M. (1998). A Work Domain Analysis of Patient Monitoring in the Operating Room. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting, 42*(14), 1038–1042.  
<http://doi.org/10.1177/154193129804201407>
- Hajdukiewicz, J. R., & Vicente\*, K. J. (2004). A theoretical note on the relationship between work domain analysis and task analysis. *Theoretical Issues in Ergonomics Science, 5*(6), 527–538.



- Hajdukiewicz, J. R., Vicente, K. J., Doyle, D. J., Milgram, P., & Burns, C. M. (2001). Modeling a medical environment: An ontology for integrated medical informatics design. *International Journal of Medical Informatics*, 62(1), 79–99. [http://doi.org/10.1016/S1386-5056\(01\)00128-9](http://doi.org/10.1016/S1386-5056(01)00128-9)
- Hall, T. J., Rudolph, J. W., & Cao, C. G. L. (2006). Fixation and Attention Allocation in Anesthesiology Crisis Management: An Abstraction Hierarchy Perspective. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 50(10), 1064–1067. <http://doi.org/10.1177/154193120605001013>
- Harjumaa, M. (2014). *On the development of persuasive systems - A framework for designing and evaluating behavior change support systems and its applicability for e-Health*. VTT Science 68. <http://doi.org/978-951-38-8169-6>
- Harrell Jr, F. E. (2016). Package “Hmisc.”
- Haug, A., & Stentoft Arlbjørn, J. (2011). Barriers to master data quality. *Journal of Enterprise Information Management*, 24(3), 288–303. <http://doi.org/10.1108/17410391111122862>
- Hawes, C., Fries, B. E., James, M. L., & Guihan, M. (2007). Prospects and pitfalls: use of the RAI-HC assessment by the Department of Veterans Affairs for home care clients. *The Gerontologist*, 47(3), 378–387.
- Hersh, W. R., Weiner, M. G., Embi, P. J., Logan, J. R., Payne, P. R., Bernstam, E. V., Lehmann, H.P., Hripcsak, G., Hartzog, T.H., Cimino, J.J. & Saltz, J. H. (2013). Caveats for the use of operational electronic health record data in comparative effectiveness research. *Medical care*, 51(8 0 3), S30.
- Higgins, P. G. (1998). Extending cognitive work analysis to manufacturing scheduling. In *Computer Human Interaction Conference, 1998. Proceedings. 1998 Australasian* (pp. 236–243). IEEE.
- Hignett, S., Carayon, P., Buckle, P., & Catchpole, K. (2013). State of science: human factors and ergonomics in healthcare. *Ergonomics*. Taylor & Francis. <http://doi.org/10.1080/00140139.2013.822932>
- Hirdes, J. P., Fries, B. E., Morris, J. N., Steel, K., Mor, V., Frijters, D., Teare, G. (2000). Integrated health information systems based on the RAI/MDS series of instruments. In *Healthcare management forum* (Vol. 12, pp. 30–40). Elsevier.

- Hirsch, M. D. (2012). Inaccurate EHR info can spoil research, public health efforts. Retrieved from <http://www.fierceemr.com/story/inaccurate-ehr-info-can-spoil-research-public-health-efforts/2012-10-17>
- Hogan, W. R., & Wagner, M. M. (1997). Accuracy of data in computer-based patient records. *Journal of the American Medical Informatics Association : JAMIA*, 4(5), 342–55.
- Jamieson, G. A., & Vicente, K. J. (2001). Ecological interface design for petrochemical applications: supporting operator adaptation, continuous learning, and distributed, collaborative work. *Computers & Chemical Engineering*, 25(7), 1055–1074.
- Jansson, A., Olsson, E., & Erlandsson, M. (2006). Bridging the gap between analysis and design: improving existing driver interfaces with tools from the framework of cognitive work analysis. *Cognition, Technology & Work*, 8(1), 41–49.
- Jenkins, D. P., Stanton, N. A., Salmon, P. M., Walker, G. H., & Young, M. S. (2008). Using cognitive work analysis to explore activity allocation within military domains. *Ergonomics*, 51(6), 798–815.
- Jiancaro, T., Jamieson, G. A., & Mihailidis, A. (2013). Twenty Years of Cognitive Work Analysis in Health Care: A Scoping Review. *Journal of Cognitive Engineering and Decision Making*, 8(1), 3–22. <http://doi.org/10.1177/1555343413488391>
- Kapur, N., Parand, A., Soukup, T., Reader, T., & Sevdalis, N. (2016). Aviation and healthcare: a comparative review with implications for patient safety. *JRSM Open*, 7(1), 2054270415616548.
- Kelders, S. M., Kok, R. N., Ossebaard, H. C., & Van Gemert-Pijnen, J. E. W. C. (2012). Persuasive system design does matter: a systematic review of adherence to web-based interventions. *Journal of Medical Internet Research*, 14(6), e152.
- Kelter, U., Wehren, J., & Niere, J. (2005). A Generic Difference Algorithm for UML Models. *Software Engineering*, 105–116. Retrieved from <http://balrok.com/assets/2protected6e113102/book/uml/generic.pdf>
- Kilgore, R. M., St-Cyr, O., & Jamieson, G. A. (2009). From Work Domains to Worker Competencies: A Five-Phase CWA. In *Applications of Cognitive Work Analysis* (1st ed., pp. 15–48). Boca Raton: CRC Press.

- Kirschenbaum, S. S., Trafton, J. G., & Pratt, E. (2007). Comparative cognitive task analysis. In *Expertise out of context: Proceedings of the Sixth International Conference on Naturalistic Decision Making* (pp. 327–336). Erlbaum New York.
- Kralj, B., & Kantarevic, J. (2012). Primary care in Ontario: reforms, investments and achievements. *Ontario Medical Review, 79*(2), 18–24.
- Larsen, I. K., Smastuen, M., Johannesen, T. B., Langmark, F., Parkin, D. M., Bray, F., & Moller, B. (2009). Data quality at the Cancer Registry of Norway: An overview of comparability, completeness, validity and timeliness. *European Journal of Cancer, 45*(7), 1218–1231. <http://doi.org/10.1016/j.ejca.2008.10.037>
- Lee, Y. W., Strong, D. M., Kahn, B. K., & Wang, R. Y. (2002). AIMQ: A methodology for information quality assessment. *Information and Management, 40*(2), 133–146. [http://doi.org/10.1016/S0378-7206\(02\)00043-5](http://doi.org/10.1016/S0378-7206(02)00043-5)
- Légaré, F., Stacey, D., Graham, I. D., Elwyn, G., Pluye, P., Gagnon, M.-P., Desroches, S. (2008). Advancing theories, models and measurement for an interprofessional approach to shared decision making in primary care: a study protocol. *BMC Health Services Research, 8*(1), 2. <http://doi.org/10.1186/1472-6963-8-2>
- Legare, F., & Witteman, H. O. (2013). Shared Decision Making: Examining Key Elements And Barriers To Adoption Into Routine Clinical Practice. *Health Affairs, 32*(2), 276–284. <http://doi.org/10.1377/hlthaff.2012.1078>
- Lehane, P., Toleman, M., & Benecke, J. (2000). Applying ecological interface design to experimental apparatus used to monitor a refrigeration plant. In *User Interface Conference, 2000. AUIC 2000. First Australasian* (pp. 41–48). IEEE.
- Lehto, T., & Oinas-Kukkonen, H. (2011). Persuasive features in web-based alcohol and smoking interventions: a systematic review of the literature. *Journal of Medical Internet Research, 13*(3), e46.
- Liaw, S.-T., Rahimi, A., Ray, P., Taggart, J., Dennis, S., de Lusignan, S., Talaei-Khoei, A. (2013). Towards an ontology for data quality in integrated chronic disease management: a realist review of the literature. *International Journal of Medical Informatics, 82*(1), 10–24.

- Lim, R. H. M., Anderson, J. E., & Buckle, P. W. (2015). Work Domain Analysis for understanding medication safety in care homes in England : an exploratory study Work Domain Analysis for understanding medication safety in care homes in England : an exploratory study. *Ergonomics*, 59(AUGUST), 15–26. <http://doi.org/10.1080/00140139.2015.1057542>
- Limbourg, Q., & Vanderdonckt, J. (2004). Comparing task models for user interface design. *The Handbook of Task Analysis for Human-Computer Interaction*, 6, 135–154.
- Lintern, G., Miller, D., & Baker, K. (2002). Work centered design of a USAF mission planning system. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* (Vol. 46, pp. 531–535). SAGE Publications.
- Lintern, G., & Naikar, N. (2000). The use of work domain analysis for the design of training systems. In *Proceedings of the human factors and ergonomics society annual meeting* (Vol. 44, pp. 198–201). SAGE Publications.
- Lockton, D., Harrison, D., & Stanton, N. A. (2010). The Design with Intent Method: A design tool for influencing user behaviour. *Applied Ergonomics*, 41(3), 382–392.
- Madnick, S. E., Wang, R. Y., Lee, Y. W., & Zhu, H. (2009). Overview and Framework for Data and Information Quality Research. *Journal of Data and Information Quality*, 1(1), 1–22. <http://doi.org/10.1145/1515693.1516680>
- Maguire, H. (2007). Book Review: Data Quality: Concepts, Methodologies and Techniques by C. Batini and M. Scannapieco. *International Journal of Information Quality*, 1(4), 444. <http://doi.org/10.1504/IJIQ.2007.016717>
- Majeed, A. (2004). Sources, uses, strengths and limitations of data collected in primary care in England. *Health Statistics Quarterly / Office for National Statistics*, (21), 5–14. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/15615148>
- Majeed, A., Car, J., & Sheikh, A. (2008). Accuracy and completeness of electronic patient records in primary care. *Family Practice*, 25(4), 213–4. <http://doi.org/10.1093/fampra/cmn047>
- Makoul, G., & Clayman, M. L. (2006). An integrative model of shared decision making in medical encounters. *Patient Education and Counseling*, 60(3), 301–312. <http://doi.org/10.1016/j.pec.2005.06.010>

- Mazaeva, N., & Bisantz, A. M. (2007). On the representation of automation using a work domain analysis. *Theoretical Issues in Ergonomics Science*, 8(6), 509–530.
- McIlroy, R. C., & Stanton, N. (2015). Ecological Interface Design Two Decades On: Whatever Happened to the SRK Taxonomy? *Human-Machine Systems, IEEE Transactions on*, 45(2), 145–163.
- Mikkonen, J., & Raphael, D. (2010). Social determinants of health: The Canadian facts. Retrieved from [http://www.thecanadianfacts.org/The\\_Canadian\\_Facts.pdf](http://www.thecanadianfacts.org/The_Canadian_Facts.pdf)
- Miller, A. (2004). A work domain analysis framework for modelling intensive care unit patients. *Cognition, Technology & Work*, 6(4), 207–222. <http://doi.org/10.1007/s10111-004-0151-5>
- Mioch, T., Mistrzyk, T., & Rister, F. (2010). Procedure design and validation by cognitive task model simulations. *19th Conference on Behavior Representation in Modeling and Simulation*, (October 2015), 232–239.
- Mohr, D. C., Schueller, S. M., Montague, E., Burns, M. N., & Rashidi, P. (2014). The behavioral intervention technology model: an integrated conceptual and technological framework for eHealth and mHealth interventions. *Journal of Medical Internet Research*, 16(6), e146.
- Momtahan, K., & Burns, C. (2004). Applications of ecological interface design in supporting the nursing process. *Journal of Healthcare Information Management : JHIM*, 18(4), 74–82. Retrieved from <http://europepmc.org/abstract/MED/15537138>
- Montgomery, D. C. (2001). Introduction to statistical process control. *John Wiley & Sons, New York, NY*.
- Münch, C., Gottschall, M., Hübsch, G., Köberlein-Neu, J., Schübel, J., Bergmann, A., & Voigt, K. (2016). Quality of documented diagnosis in primary care-An analysis using the example of thyroid disorders. *Zeitschrift Fur Evidenz, Fortbildung Und Qualitat Im Gesundheitswesen*, 115, 56–62.
- Naikar, N., Pearce, B., Drumm, D., & Sanderson, P. M. (2003). Designing teams for first-of-a-kind, complex systems using the initial phases of cognitive work analysis: Case study. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 45(2), 202–217.
- Naikar, N., & Sanderson, P. M. (2001). Evaluating design proposals for complex systems with work domain analysis. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 43(4), 529–542.

- Naikar, N., Sanderson, P. M., & Lintern, G. (1999). Work domain analysis for identification of training needs and training-system design. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* (Vol. 43, pp. 1128–1132). SAGE Publications.
- Naikar, N., & Saunders, A. (2003). Crossing the boundaries of safe operation: An approach for training technical skills in error management. *Cognition, Technology & Work*, 5(3), 171–180.
- Nemeth, C. (2006). Discovering Healthcare Cognition: The Use of Cognitive Artifacts to Reveal Cognitive Work. *Organization Studies*, 27(7), 1011–1035.  
<http://doi.org/10.1177/0170840606065708>
- Oinas-Kukkonen, H. (2010). Behavior Change Support Systems : The Next Frontier for Web Science. *Web Science Conference*, (April), 1–8.
- Oinas-Kukkonen, H., & Harjumaa, M. (2008). Towards Deeper Understanding of Persuasion in Software and Information Systems. *First International Conference on Advances in Computer-Human Interaction*, 200–205. <http://doi.org/10.1109/ACHI.2008.31>
- Oinas-Kukkonen, H., & Harjumaa, M. (2009). Key Issues , Process Model , and System Features Persuasive Systems Design : Key Issues , Process Model , and System Features. *Communications of the Association for Information Systems*, 24(28), 485–500.
- OntarioMD. (n.d.). EMR Practice Enhancement Program. Retrieved February 23, 2017, from [https://www.ontariomd.ca/portal/server.pt/community/emr\\_practice\\_enhancement\\_program/727](https://www.ontariomd.ca/portal/server.pt/community/emr_practice_enhancement_program/727)
- OntarioMD. (2015). EMR Funding Overview. Retrieved from [https://www.ontariomd.ca/portal/server.pt/community/emr\\_funding/](https://www.ontariomd.ca/portal/server.pt/community/emr_funding/)
- Pakhomov, S., Shah, N., Hanson, P., Balasubramaniam, S., Smith, S. a, & Smith, S. A. (2008). Automatic quality of life prediction using electronic medical records. *AMIA Annual Symposium Proceedings / AMIA Symposium*. *AMIA Symposium*, 545–9. Retrieved from <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2656045&tool=pmcentrez&rendertype=abstract>
- Papanicolas, I., & Smith, P. (2013). *Health system performance comparison: an agenda for policy, information and research: an agenda for policy, information and research*. McGraw-Hill Education (UK).

- Paternò, F., Mori, G., & Galiberti, R. (2001). CTTE : An Environment for Analysis and Development of Task Models of Cooperative Applications. *CHI'01 Extended Abstracts on Human*, 21–22.  
<http://doi.org/10.1145/634067.634084>
- Patterson, E. S., Schweikhart, S., Anders, S., Brungs, S., & Render, M. L. (2007, October). Towards a Functional Model of Quality Improvement Collaboratives. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* (Vol. 51, No. 11, pp. 707-711). SAGE Publications.
- Pearce, C. M., de Lusignan, S., Phillips, C., Hall, S., & Travaglia, J. (2013). The computerized medical record as a tool for clinical governance in Australian primary care. *Interactive Journal of Medical Research*, 2(2), e26. <http://doi.org/10.2196/ijmr.2700>
- Perera, G., Holbrook, A., Thabane, L., Foster, G., & Willison, D. J. (2011). Views on health information sharing and privacy from primary care practices using electronic medical records. *International Journal of Medical Informatics*, 80(2), 94–101. <http://doi.org/10.1016/j.ijmedinf.2010.11.005>
- Provost, L. P., & Murray, S. (2011). *The health care data guide: learning from data for improvement*. John Wiley & Sons.
- Rasmussen, J. (1974). *The human data processor as a system component. Bits and pieces of a model* (Risø-M).
- Rasmussen, J. (1983). Skills, rules, and knowledge; signals, signs, and symbols, and other distinctions in human performance models. *Systems, Man and Cybernetics, IEEE Transactions on, SMC-13*(3), 25–266. <http://doi.org/10.1109/TSMC.1983.6313160>
- Rezai, L. S., & Burns, C. M. (2014). Using Cognitive Work Analysis and a Persuasive Design Approach to Create Effective Blood Pressure Management Systems. *Proceedings of the International Symposium on Human Factors and Ergonomics in Health Care*, 3(1), 36–43.  
<http://doi.org/10.1177/2327857914031005>
- Saleem, J. J., Russ, A. L., Sanderson, P., Johnson, T. R., Zhang, J., & Sittig, D. F. (2009). Current challenges and opportunities for better integration of human factors research with development of clinical information systems. *Yearbook of Medical Informatics*, 48–58. Retrieved from <http://archive.itee.uq.edu.au/~cerg/publications/IMIAYearbook2009SittigEtAlHumanFactorsHealthInformatics.pdf>

- Salmon, P., Regan, M., Lenné, M. G., Stanton, N. A., & Young, K. (2007). Work domain analysis and intelligent transport systems: implications for vehicle design. *International Journal of Vehicle Design*, 45(3), 426–448.
- Sanderson, P., Li, X., Memisevic, R., Wong, W. B. L., & Choudhury, S. (2005). Evaluating functional displays for hydropower systems: Model-based guidance of scenario design. In *Proceedings of the European Annual Meeting on Decision Making and Control (EAM2005)*, Athens, Greece.
- Sarcevic, A., Lesk, M. E., Marsic, I., & Burd, R. S. (2010). Towards an efficient method for studying collaborative practices of emergency care teams.
- Sauro, J., & Lewis, J. R. (2016). *Quantifying the user experience: Practical statistics for user research*. Morgan Kaufmann.
- Shankaranarayanan, G., & Cai, Y. (2006). Supporting data quality management in decision-making. *Decision Support Systems*, 42(1), 302–317.
- Shepherd, A. (1998). HTA as a framework for task analysis. *Ergonomics*, 41(11), 1537–1552.  
<http://doi.org/10.1080/001401398186063>
- Sollie, A., Roskam, J., Sijmons, R. H., Numans, M. E., & Helsper, C. W. (2016). Do GPs know their patients with cancer? Assessing the quality of cancer registration in Dutch primary care: a cross-sectional validation study. *BMJ Open*, 6(9), e012669.
- Sollie, A., Sijmons, R. H., Helsper, C., & Numans, M. E. (2017). Reusability of coded data in the primary care electronic medical record: A dynamic cohort study concerning cancer diagnoses. *International Journal of Medical Informatics*, 99, 45–52.
- St-Maurice, J. (2011). Primary Care Data : Gold or Pyrite? A Literature Review. *Journal of Healthcare Information Management*, 25(2), 27–31. <http://doi.org/10.1267/METH03040416>
- St-Maurice, J., & Burns, C. M. (2014). User Perception of Data and Medical Record Personalities. *Proceedings of the International Symposium of Human Factors and Ergonomics in Healthcare*, 3(1), 15–22. <http://doi.org/10.1177/2327857914031002>
- St-Maurice, J., & Burns, C. M. (2015). Using Comparative Cognitive Work Analysis to Identify Design Priorities in Complex Socio-Technical Systems. In *Proceedings of the International Symposium on Human Factors and Ergonomics in Health Care* (Vol. 4, pp. 118–123). SAGE Publications.



- St-Maurice, J., & Burns, C. M. (in press). An Exploratory Case Study to Understand Primary Care Users and Their Data Quality Tradeoffs. *Journal of Data and Information Quality*. ACM Publications.
- St-Maurice, J., Kuo, M.-H., & Gooch, P. (2013). A proof of concept for assessing emergency room use with primary care data and natural language processing. *Methods of Information in Medicine*, 52(1), 33–42. <http://doi.org/10.3414/ME12-01-0012>
- Starfield, B. (1997). The future of primary care in a managed care era. *International Journal of Health Services*, 27(4), 687–696.
- Stone, M. a, Camosso-Stefinovic, J., Wilkinson, J., de Lusignan, S., Hattersley, A. T., & Khunti, K. (2010). Incorrect and incomplete coding and classification of diabetes: a systematic review. *Diabetic Medicine : A Journal of the British Diabetic Association*, 27(5), 491–7. <http://doi.org/10.1111/j.1464-5491.2009.02920.x>
- Strong, D. M., Lee, Y. W., & Wang, R. Y. (1997). Data quality in context. *Communications of the ACM*, 40(5), 103–110. <http://doi.org/10.1145/253769.253804>
- Suls, J., & Bruchmann, K. (2013). Social Comparison and Persuasion Processes in Health Communications, (October), 1–20. <http://doi.org/10.1093/oxfordhb/9780199795833.013.023>
- Sulmasy, L. S., López, A. M., & Horwitch, C. A. American College of Physicians Ethics, Professionalism and Human Rights Committee. (2017). Ethical Implications of the Electronic Health Record: In the Service of the Patient. *Journal of General Internal Medicine*, 1-5.
- Thiru, K., Hassey, A., & Sullivan, F. (2003). Systematic review of scope and quality of electronic patient record data in primary care. *British Medical Journal*, 326.
- Thompson-Leduc, P., Turcotte, S., Labrecque, M., & Légaré, F. (2016). Prevalence of clinically significant decisional conflict: an analysis of five studies on decision-making in primary care. *BMJ Open*, 6(6), e011490. <http://doi.org/10.1136/bmjopen-2016-011490>
- Thompson, V. D. (2015). *Health and health care delivery in Canada*. Elsevier Health Sciences.
- Thor, J., Lundberg, J., Ask, J., Olsson, J., Carli, C., Härenstam, K. P., & Brommels, M. (2007). Application of statistical process control in healthcare improvement: systematic review. *Quality and Safety in Health Care*, 16(5), 387–399.

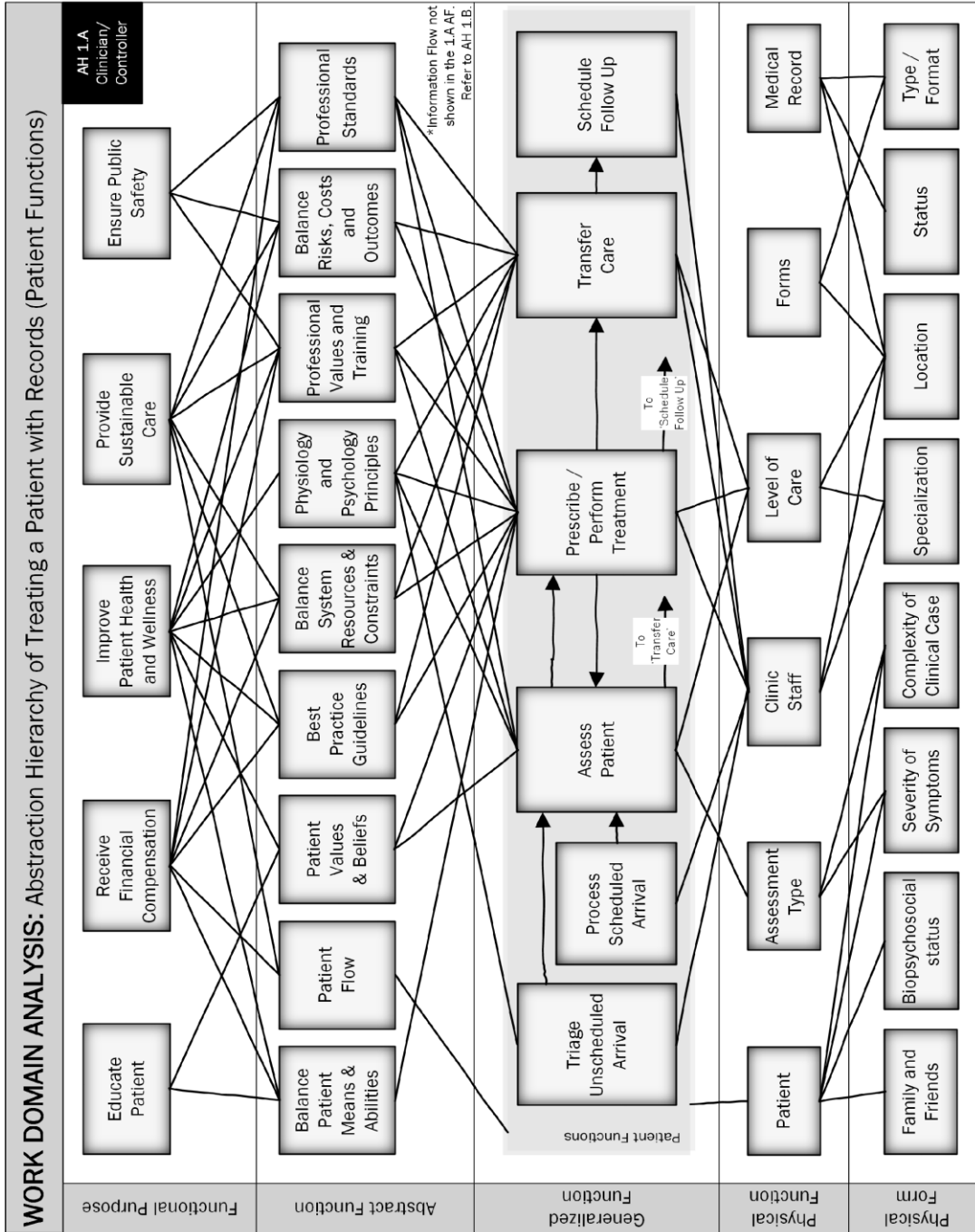
- Todoran, I., Lecornu, L., Khenchaf, A., & Caillec, J.-M. Le. (2015). A Methodology to Evaluate Important Dimensions of Information. *ACM Journal of Data and Information Quality*, 6(2–3, Article 11), 23. <http://doi.org/http://dx.doi.org/10.1145/2744205>
- Torning, K. (2013). A Review of Four Persuasive Design Models. *International Journal of Conceptual Structures and Smart Applications*, 1(2), 17–27. <http://doi.org/10.4018/ijcssa.2013070103>
- Upton, C., & Doherty, G. (2008). Extending Ecological Interface Design principles: A manufacturing case study. *International Journal of Human-Computer Studies*, 66(4), 271–286.
- van der Bij, S., Khan, N., ten Veen, P., de Bakker, D. H., & Verheij, R. A. (2016). Improving the quality of EHR recording in primary care: a data quality feedback tool. *Journal of the American Medical Informatics Association*, ocw054.
- Vicente, K. J. (1999). *Cognitive Work Analysis*. Mahwah, NJ: Lawrence Erlbaum Associates, Inc.
- Vicente, K. J., & Rasmussen, J. (1992). Ecological interface design: Theoretical foundations. *Systems, Man and Cybernetics*. Retrieved from [http://ieeexplore.ieee.org/xpls/abs\\_all.jsp?arnumber=156574](http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=156574)
- Volker, N., Davey, R. C., Cochrane, T., Williams, L. T., & Clancy, T. (2014). Improving the prevention of cardiovascular disease in primary health care: the model for prevention study protocol. *JMIR Research Protocols*, 3(3), e33. <http://doi.org/10.2196/resprot.2882>
- Wand, Y., & Wang, R. Y. (1996). Anchoring data quality dimensions in ontological foundations. *Communications of the ACM*, 39(11), 86–95. <http://doi.org/10.1145/240455.240479>
- Wang, R., Strong, D., & Guarascio, L. (1996). Beyond accuracy: What data quality means to data consumers. *J. of Management Information*, (October). Retrieved from <http://web.mit.edu/tdqm/www/papers/94/94-10.html>
- Watson, M. O., & Sanderson, P. (1998). Work domain analysis for the evaluation of human interaction with anaesthesia alarm systems. *Proceedings of the 8th Australasian Conference on Computer-Human Interaction (OzCHI '98)*, 228–235. <http://doi.org/10.1109/OZCHI.1998.732219>
- Watson, M. O., & Sanderson, P. M. (2007). Designing for Attention With Sound: Challenges and Extensions to Ecological Interface Design. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 49(2), 331–346. <http://doi.org/10.1518/001872007X312531>

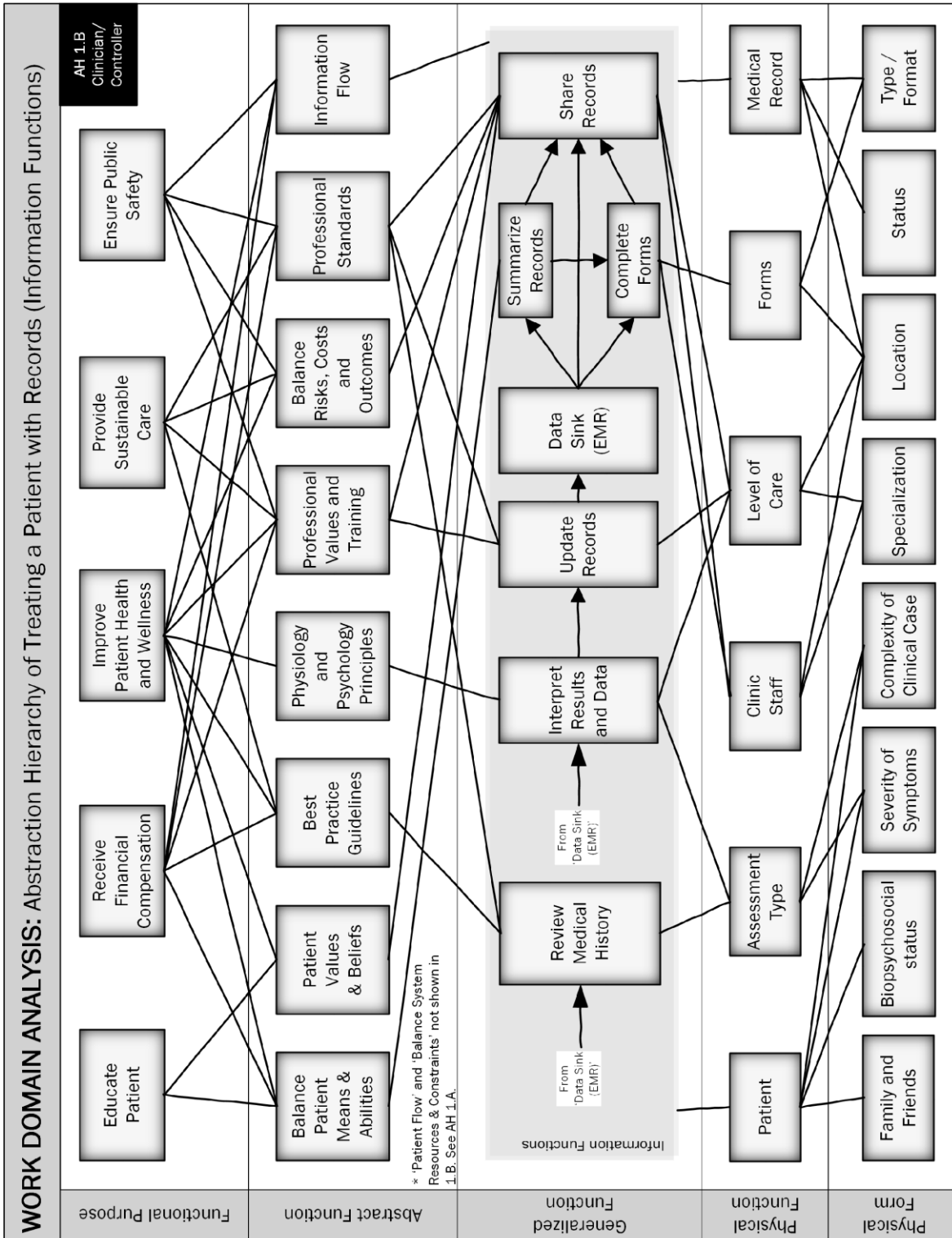
- Webster, P. C. (2012). After all the time and money invested, will e-health ever deliver on its promise? Retrieved from <http://www.theglobeandmail.com/report-on-business/rob-magazine/after-all-the-time-and-money-invested-will-e-health-ever-deliver-on-its-promise/article5796658/>
- Weilburg, J. B., Dreyer, K. J., Rosenthal, D. I., & Thrall, J. H. (2009). Effect of Computerized Order Entry with Integrated Decision Support on the Growth of Outpatient Procedure Volumes, *251*(1), 147–155.
- Weston, W. W. (2001). Informed and shared decision-making: the crux of patient-centered care. *CMAJ: Canadian Medical Association Journal*, *165*(4), 438–9. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/11531054>
- Wheeler, D. J. (2012). Lies, Damned Lies, and Teens Who Smoke While Driving. *Quality Digest Daily*, *January 4*. Retrieved from <http://www.spcpress.com/pdf/DJW237.pdf>
- Yu, C. H., Stacey, D., Sale, J., Hall, S., Kaplan, D. M., Ivers, N., Straus, S. E. (2014). Designing and evaluating an interprofessional shared decision-making and goal-setting decision aid for patients with diabetes in clinical care - systematic decision aid development and study protocol. *Implementation Science*, *9*(1), 16. <http://doi.org/10.1186/1748-5908-9-16>



# Appendix A

## Abstraction Hierarchies of Patient Treatment (AH 1.A & 1.B)

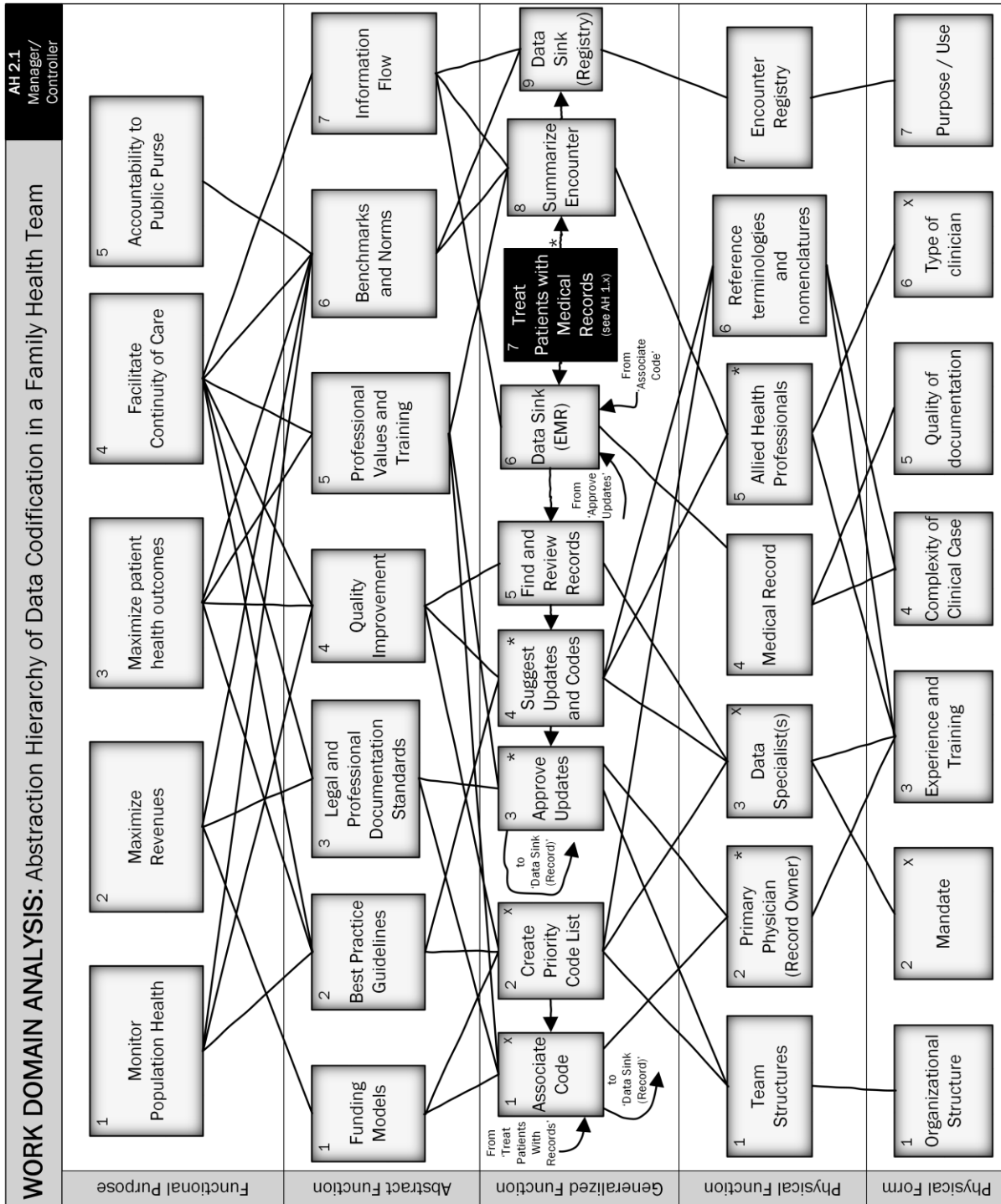




# Appendix B

## Work Domain Analysis of Data Codification

### Abstraction Hierarchy of Data Codification in a Family Health Team (AH 2.1)



### **AH 2.1 - Functional Purpose (FP)**

ID	Box Name in AH 2.1	Description & Rational for Functional Purpose Modeling Choices	Links in AH 2.1
			<i>How (AF)</i>
<b>FP1</b>	Monitor Population Health	Primary care users code data to measure and monitor population health within the community. Coded data can help identify patients with common chronic diseases. This is mediated by quality initiatives (e.g. programs that are funded to target specific patient populations) and professional values (e.g. an interest in running a preventative practice that priorities chronic disease management).	AF2 AF4 AF6
<b>FP2</b>	Maximize Revenues	Primary care users code data to maximize revenues. Codification can help physicians identify patients who are eligible for special billing, and may also be a requirement for billing the Government. Primary care physicians are able to bill the Government with special codes to receive bonus payments if data is coded properly. This is mediated by funding models, professional values and organizational benchmarks.	AF1 AF3 AF6
<b>FP3</b>	Maximize Patient Health Outcomes	Codification can allow clinicians to activate decision support tools, such as the drug interaction module of an EMR or facilitate the adoption of best practice guidelines. This helps maximize patient health outcomes. This is mediated by quality improvement initiatives, professional values and relative norms.	AF2 AF4 AF5 AF6
<b>FP4</b>	Facilitate Continuity of Care	Identifying patients and codifying data can support patients in their journey across the healthcare system. For example, there are initiatives to identify and track patients who have COPD. The purpose of these initiatives is to improve their care through identification, and to reduce their readmission rates in hospital. This is mediated by professional guidelines, quality improvement programs, training, relative norms and information flow.	AF2 AF3 AF4 AF5 AF6 AF7
<b>FP5</b>	Accountability to Public Purse	Within Family Health Teams, reports need to be generated to the Government to demonstrate accountability for funding. Coded data, which can be created within a registry, can be used to generate accurate reports. This is mediated by performance benchmarks.	AF6



### AH 2.1 - Abstract Function (AF)

ID	Box Name in AH 2.1	Description & Rational for Abstract Function Modeling Choices	Links in AH 2.1	
			Why (FP)	How (GF)
AF1	Funding Models	In primary care, there are potential bonuses associated with care. For example, a physician is eligible to bill an extra code to the Government every quarter for diabetic patients. Funding models allow clinicians to maximize revenues and incentivize certain coding practices. Priority codes are associated with funding models.	FP2	GF1 GF2
AF2	Best Practice Guidelines	Best practice guidelines can help clinicians provide the best care and maximize health outcomes. Coding records allows for the identification of patients to apply best practice guidelines. Guidelines will influence which codes are needed as a priority.	FP1 FP3 FP4	GF2 GF4
AF3	Legal and Professional Documentation Standards	Legal documentation standards moderate revenues and provide a record of diagnosis that rationalize and support codification. To facilitate the continuity of care, documentation standards must be adopted to support and maximize transfers of care.	FP2 FP4	GF1 GF3
AF4	Quality Improvement	Quality improvement initiatives aim to improve health care quality. Quality improvement initiatives also address documentation gaps and data quality problems. Quality improvement maximizes health outcomes, facilitates the continuity of care and allows a family health team to monitor population health. These programs guide improvement.	FP1 FP3 FP4	GF2 GF4 GF5
AF5	Professional Values and Training	Professional values and training play a role in whether or not a clinician will choose to code their data, or approve amendments suggested by allied health professionals. Training also influences how clinicians record information in reporting registries.	FP1 FP3 FP4	GF1 GF3 GF8
AF6	Benchmarks and Norms	Benchmarks and norms are used by the Government and health team administration to compare performance between employees and organizations. Benchmarks give users a reason to input data, as it enables them to compare themselves to benchmarks.	FP1 FP2 FP3 FP4 FP5	GF8 GF9
AF7	Information Flow	Information Flow is a representation of information that enters the system, is used and is stored in an EMR, registry or mental model.	FP4	GF6 GF8 GF9

### AH 2.1 - Generalized Function (GF)

ID	Box Name in AH 2.1	Description & Rational for Generalized Function Modeling Choices	Links in AH 2.1		
			Why (AF)	How (PF)	Within GF
<b>GF1</b>	Associate Code	As part of data quality programs, clinicians are encouraged to use common terminology and codes within the EMR. In this workflow, clinicians associate the recommended code to a note or diagnosis field. Easy to remember codes are taken from a shortlist.	AF1 AF3 AF5	PF2	To GF6, from AH1.A, from GF2
<b>GF2</b>	Create Priority Code List	As part of data quality programs, common codes are suggested as a shortlist for clinicians to adopt in their records.	AF1 AF2 AF4	PF1 PF3 PF6	To GF1
<b>GF3</b>	Approve Updates	The owner of the medical records (generally the family physician) must approve changes to their records, since they are the health information custodian.	AF3 AF5	PF1 PF2	From GF4, to GF6
<b>GF4</b>	Suggest Updates and Codes	Allied health professional and data quality specialists review records and suggest changes and codes to enhance clinical documentation.	AF2 AF4	PF3 PF5 PF6	To GF3, from GF5
<b>GF5</b>	Find and Review Records	Allied health professional and data quality specialists create searches and find records that meet specific criteria.	AF4	PF3	To GF4, from GF6
<b>GF6</b>	Data Sink (EMR)	Data comes from and goes into the EMR system after being processed through workflows.	AF7	PF4	To GF5
<b>AH 1.A + AH 1.B</b>	Treat Patients with Medical Records	Data is created during patient treatment. A separate abstraction hierarchy represents this domain. This generates data to the EMR and forms the basis for summary data and codification.	AH 1.x	AH 1.x	To GF6, To GF8, To GF1
<b>GF8</b>	Summarize Encounter	Clinicians who work for Family Health Teams must summarize their activities for reporting purposes. Data is collected in a registry, or inside the EMR.	AF5 AF6 AF7	PF5	From GF7, To GF9
<b>GF9</b>	Data Sink (Registry)	The data registry is used to create Government reports.	AF7	PF7	From GF8

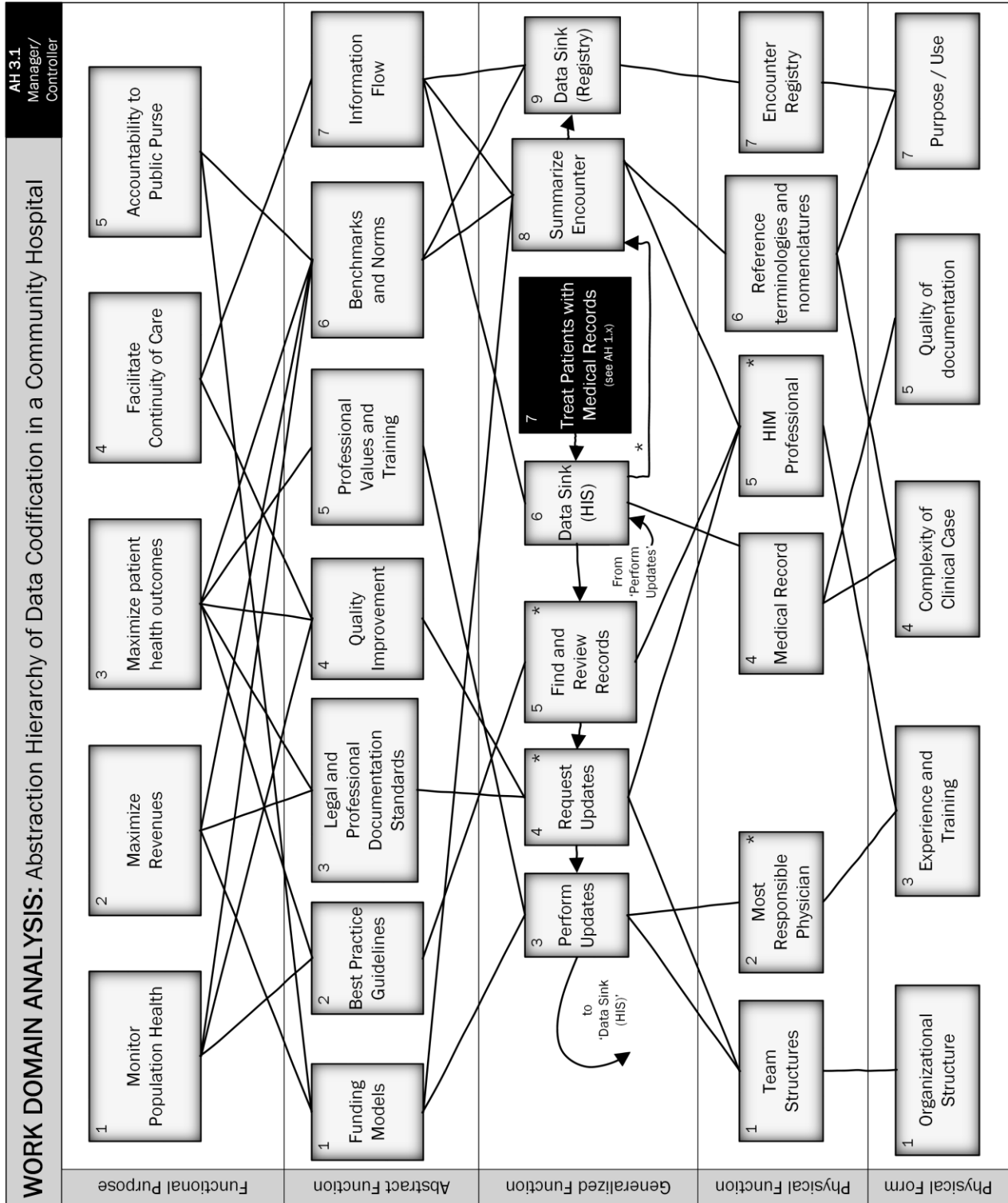
### AH 2.1 - Physical Function (PF)

ID	Box Name in AH 2.1	Description & Rational for Physical Function Modeling Choices	Links in AH 2.1	
			Why (GF)	How (PFo)
PF1	Team Structures	Team structures play an important role in workflows and influence the type of authority that allied health professionals have regarding information and codes within the medical record.	GF2 GF3	PFo1
PF2	Primary Physician (Record Owner)	Primary Physicians are the medical record owners. They are important actors that must approve any changes or amendments to the medical record.	GF1 GF3	PFo3
PF3	Data Specialist(s)	Data specialists (such as Quality Improvement Decision Support Specialists) are employed by Family Health Teams to work on health care quality improvement and data quality improvement projects. They provide expertise and workflow support.	GF2 GF4 GF5	PFo2 PFo3
PF4	Medical Record	All information that is coded needs to be reflective of the information contained within the medical record.	GF6	PFo4 PFo5
PF5	Allied Health Professionals	Allied Health Professionals are responsible for providing additional health care services in clinics. They have patient treatments roles, coding roles and support a variety of workflows.	GF4 GF8	PFo3 PFo6
PF6	Reference Terminologies and Nomenclatures	External reference terminology and nomenclatures (such as ICD-10-CA) are internationally recognized lists of codes that can be used to consistent codify data and ensure is it easier to tag data and find information when required.	GF2 GF4	PFo3 PFo4
PF7	Encounter Registry	The encounter registry at a Family Health Team is a database where information is stored and used for Government reporting purposes. This role can sometimes also be filled by reports from the EMR, if data is properly coded.	GF9	PFo7

### AH 2.1 - Physical Form (PFo)

ID	Box Name in AH 2.1	Description & Rational for Physical Form Modeling Choices	Links in AH 2.1
			<i>Why (PF)</i>
<b>PFo1</b>	Organizational Structure	The organization type and structure will influence how clinical teams are structured. Some family health team are governed by a community board, while others are governed by a board of physicians. Some clinics work with independent doctors, while others work in larger groups or networks.	PF1
<b>PFo2</b>	Mandate	Data specialists have varying mandates based on their experience and role at a Family Health Team. This is an important attribute in understanding how a data specialist will approach data workflows and support quality improvement.	PF3
<b>PFo3</b>	Training and Experience	Experience level is an important attribute for physicians (record owners), data specialists and allied health professionals. This has an impact on coding.	PF2 PF3 PF5 PF6
<b>PFo4</b>	Complexity of Clinical Case	The complexity of a clinical case and patient status will have an important impact on the medical record and may impact the ability to accurately code diagnoses.	PF4 PF6
<b>PFo5</b>	Quality of Documentation	The quality of records will have an important impact on the medical record and impact the ability for other clinicians to interpret data, properly code data and understand the clinical case.	PF4 PF6
<b>PFo6</b>	Type of Clinician	Allied Health Professionals can be different types of clinicians. This will impact their scope of practice and their ability to code different types of information. Some clinicians might have a data quality mandate, or a health care quality improvement role.	PF5
<b>PFo7</b>	Purpose / Use	The purpose of a registry will influence what data is entered into the system, and how that data will be used. This will impact which fields are required and how information is collected.	PF7

**Abstraction Hierarchy of Data Codification in a Community Hospital (AH 3.1)**



### **AH 3.1 - Functional Purpose (FP)**

ID	Box Name in AH 3.1	Description & Rational for Functional Purpose Modeling Choices	Links in AH 3.1
			<i>How (AF)</i>
<b>FP1</b>	Monitor Population Health	In community hospitals, data is coded and provided to Government agencies, such as the Canadian Institute of Health Information (CIHI), to monitor patient health and health delivery. CIHI generates annual reports based on this data.	AF2 AF4 AF6
<b>FP2</b>	Maximize Revenues	Coded data is used to calculate resource intensity scores for patient visits. New funding models, such as the Health System Research Fund (HSRF) program, rely on high-quality coded data to generate appropriate compensation for hospitals. Revenue generation is an increasingly important reason to code accurate data.	AF1 AF3 AF6
<b>FP3</b>	Maximize Patient Health Outcomes	The codification of data in community hospitals enables the identification of community trends, such as the prevalence of mental health diagnoses. Hospitals can use this information to create interventions aimed at helping specific patient populations. Coded data is not available in the Hospital Information System (HIS) and hospitals are only able to enhance the outcomes of future patients through audits of historic performance.	AF2 AF3 AF4 AF5 AF6
<b>FP4</b>	Facilitate Continuity of Care	Identifying patients and codifying data can the hospital better identify each patient’s journey across the healthcare system and better understand reasons for re-admissions. Coding data allows hospitals to relate data back to other sectors of the healthcare system, such as primary care. For example, there are initiatives to identify and track patients who have COPD. Coded data helps identify future opportunities to improve the health care system.	AF4 AF7
<b>FP5</b>	Accountability to Public Purse	In community hospitals, data is coded and provided to Government agencies, such as CIHI. CIHI’s annual reports compare regions and hospitals, and generate accountability within hospitals when benchmarks and delivery norms are not achieved. CIHI generates public report cards and hospital rankings based on coded data.	AF1 AF6

### **AH 3.1 - Abstract Function (AF)**

ID	Box Name in AH 3.1	Description & Rational for Abstract Function Modeling Choices	Links in AH 3.1	
			Why (FP)	How (GF)
<b>AF1</b>	Funding Models	In hospitals, funding is provided through several mechanisms. New funding models have increased the importance of high-quality data, such as the Health System Research Fund (HSRF) program. These models need to be taken into consideration when coding data and ensuring the captured data is complete and accurate. New funding models give hospitals increasing reasons to code data.	FP2 FP5	GF3 GF8
<b>AF2</b>	Best Practice Guidelines	Best practice guidelines can help clinicians provide the best care and maximize health outcomes. Coding records allows for the identification of patients to apply best practice guidelines. Since there is a delay in coding, best practice adherence is measured retrospectively.	FP1 FP3	GF5
<b>AF3</b>	Legal and Professional Documentation Standards	Legal documentation standards provide a record of diagnosis that rationalize and support codification. In hospitals, if data is not in a medical record, it cannot be coded. Documentation standards are also shaped by policies and procedures.	FP2 FP3	GF4
<b>AF4</b>	Quality Improvement	Quality improvement initiatives aim to improve health care quality. To support higher care quality, quality improvement initiatives also address documentation and data quality. Quality improvement maximizes health outcomes, facilitates the continuity of care and allows the hospital to monitor population health.	FP1 FP3 FP4	GF4 GF5
<b>AF5</b>	Professional Values and Training	Professional ethics dictate that if data is not in a medical record, it cannot be coded. Training and values also enable clinicians to fully document encounters in their notes to support coding.	FP3	GF3
<b>AF6</b>	Benchmarks and Relative Norms	Benchmarks and norms are used by the Government and hospital administration to compare performance and understand minimum workflow performance. Benchmarks support workflows and system goals.	FP1 FP2 FP3 FP5	GF8 GF9
<b>AF7</b>	Information Flow	Information Flow is a representation of information that enters the system, is used and is stored in a hospital information system, or abstract database.	FP4	GF6 GF8 GF9

### AH 3.1 - Generalized Function (GF)

ID	Box Name in AH 3.1	Description & Rational for Generalized Function Modeling Choices	Links in AH 3.1		
			Why (AF)	How (PF)	Within GF
<b>GF3</b>	Perform Updates	If Health Information Management professionals (e.g. Coders) detect deficiencies or errors within medical records, the most responsible physician must perform updates or risk losing hospital privileges.	AF1 AF5	PF1 PF2	From GF4, To GF6.
<b>GF4</b>	Request Updates	If coders detect deficiencies or errors within medical records, they will formally request an update or amendment. Records might be deficient because of quality, or because of minimum legal requirements.	AF3 AF4	PF1 PF5	To GF3, from GF5
<b>GF5</b>	Find and Review Records	As part of normal quality assurance processes during coding, Health Information Management professionals review medical records to ensure they are complete and error-free.	AF2 AF3	PF5	To GF4, from GF6
<b>GF6</b>	Data Sink (HIS)	Data comes from and goes into the Hospital Information System (HIS) and is also supported by paper records. The HIS feeds data abstracting and coding as the abstracters did not personally treat patients.	AF7	PF4	To GF5, to GF8
<b>AH 1.A + AH 1.B</b>	Treat Patients with Medical Records	Data is created based during patient treatment. A separate abstraction hierarchy represents this domain. This abstraction puts data into the hospital information system (HIS).	AH 1.x	AH 1.x	To GF6
<b>GF8</b>	Summarize Encounter	Health Information Management professionals review medical records and create abstracts by summarizing and coding data to represent the clinical encounter.	AF1 AF6 AF7	PF5 PF6	From GF6
<b>GF9</b>	Data Sink (Registry)	The abstraction software collects the summarized and coded information, which is eventually transformed into a single file and submitted to CIHI on a monthly basis.	AF6 AF7	PF7	From GF8



### AH 3.1 - Physical Function (PF)

ID	Box Name in AH 3.1	Description & Rational for Physical Function Modeling Choices	Links in AH 3.1	
			Why (GF)	How (PFo)
<b>PF1</b>	Team Structures	Team structures and dynamics play an important role in workflows and influence the type of authority that Health Information Management Professionals are given within the hospital and the influence they have on clinical documentation.	GF3 GF4	PFo1
<b>PF2</b>	Most Responsible Physician	Hospitals are the medical record owners. The most responsible physician must perform any changes or amendments to the medical record.	GF3	PFo3
<b>PF4</b>	Medical Record	All information that is coded needs to reflect what information is contained within the medical record.	GF6	PFo4
<b>PF5</b>	HIM Professionals	Health Information Management Professionals are responsible for providing for coding data, reviewing records for quality and supporting a variety of records-related workflows.	GF4 GF5 GF8	PFo3 PFo7
<b>PF6</b>	Reference Terminologies and Nomenclatures	External reference terminology and nomenclatures (such as ICD-10-CA) are internationally recognized codes that can be used to consistently codify data and ensure it is easier to tag data and find information when required.	GF8	PFo4 PFo5
<b>PF7</b>	Encounter Registry	The encounter registry with the hospital is a database where abstracted data is stored and used for reporting to Government agencies.	GF9	PFo7

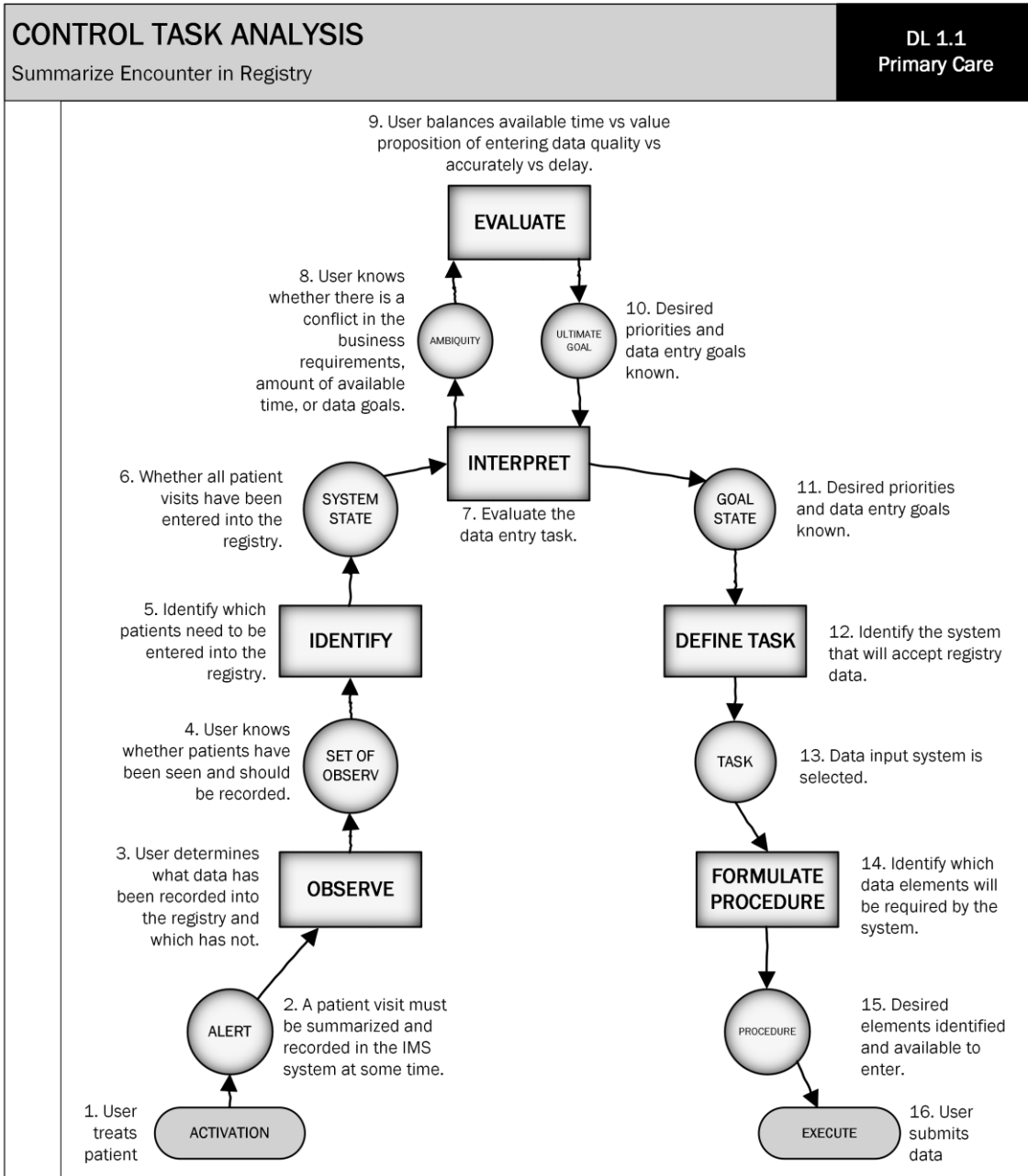
### AH 3.1 - Physical Form (PFo)

ID	Box Name in AH 3.1	Description & Rational for Physical Form Modeling Choices	Links in AH 3.1
			<i>Why (PF)</i>
<b>PFo1</b>	Organizational Structure	The organization type and structure will influence how teams are structured and supported. Teaching hospitals may be structured differently than community hospitals. The size of the community hospital will also impact its organizational structure.	PF1
<b>PFo3</b>	Experience and Training	Experience level is an important attribute for physicians and allied health professionals.	PF2 PF5
<b>PFo4</b>	Complexity of Clinical Case	The complexity of a clinical case and patient status will have an important impact on the medical record and may impact the ability to accurately code diagnoses.	PF4
<b>PFo5</b>	Quality of Documentation	The quality of documentation within medical records impacts the Health Information Management professional's ability to code the information.	PF4
<b>PFo7</b>	Purpose / Use	The purpose of a registry will influence what data is entered into the system, and how that data will be used. This will impact which fields are required and how information is collected. In hospitals, CIHI's registries (such as DADs and NACRS) are managed and mandated by law. The purpose of data will also influence which references terminologies are required to describe the clinical situation with codes.	PF6 PF7

## Appendix C

### Control Task Analysis Decision Ladders and Summary Tables

#### Summarize Encounter in Registry – Baseline (Control Task 1.1)



**Summarize Encounter in Registry – Baseline (Control Task 1.1)**

Related Abstraction Hierarchies: AH 1.A, AH 1.B, AH 2.1

Descendant Decision Ladders: DL 1.1.1 (Novice), DL 1.1.2 (Intermediate), DL 1.1.3 (Expert)

Step	Ladder Code	Text	Description
1	Activation	User treats patient.	Patients are seen daily by users, who are also clinicians.
2	Alert	A patient visit must be summarized and recorded in the reporting system at some time.	After each patient visit, a user must enter a record into the registry system.
3	Observe	User determines what data has been recorded into the registry and which has not.	The user must determine whether or not the visit has been recorded into the registry. In some situations, this is trivial. At other times, users delay their entry of the encounter by several weeks and it is not always clear which entries have been entered, and which have not.
4	Set of Observations	User knows whether patients have been seen and should be recorded.	The user is presented with, or has a mental model, of all the patients that have been seen and that should have an entry in the registry.
5	Identify	Identify which patients need to be entered into the registry.	Based on the list of patients seen and patient current in the registry, the user can identify which patient encounters must still be entered into the system and which have previously been documented.
6	System State	Whether all patient visits have been entered into the registry.	The user interprets the system state and knows which users must be entered, and which are already present.
7	Interpret	Evaluate the data entry task.	The user evaluates the scope of the task. The user may assess the amount of available time and whether the required information is available.
8	Ambiguity	User knows whether there is a conflict in the business requirements, amount of available time, or data goals.	The user may enter a state of ambiguity. There may conflicting business requirements such as seeing more patients, or spending time entering data into the registry.

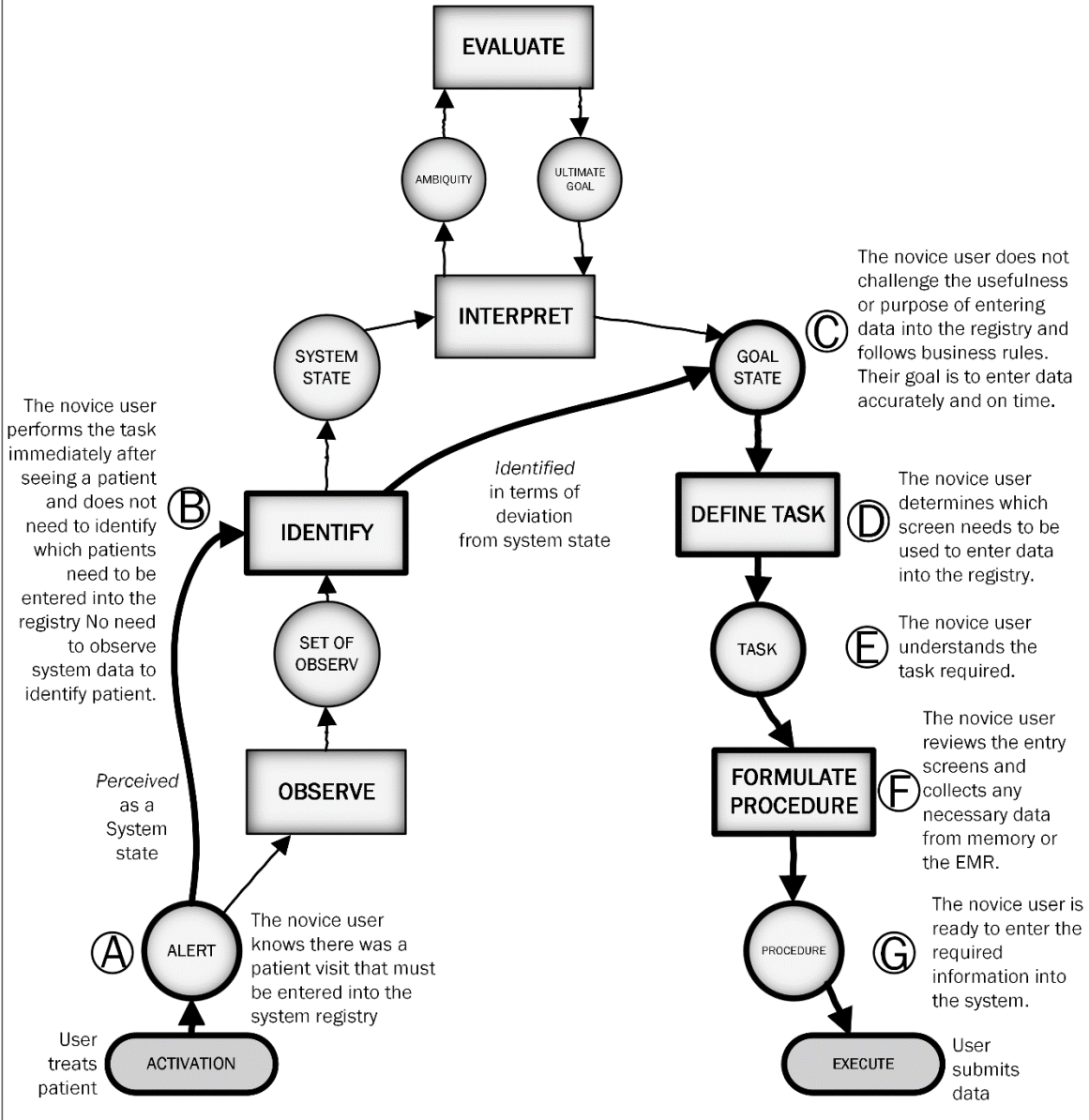
CONTROL TASK ANALYSIS Decision Ladder Summary			DLSUM 1.1 Primary Care
9	Evaluate	User balances available time vs value proposition of entering data quality vs accurately vs delay.	The user will balance their time and priorities. Users will determine whether they should invest more time to record data accurately, or delay entry tasks, or provide high levels of detail.
10	Ultimate Goal	Desired priorities and data entry goals understood.	The user understands their ultimate goal, as it relates to balance of priorities. They will perform their task accordingly.
11	Goal State	Desired priorities and data entry goals selected.	
12	Define Task	Identify the system that will accept registry data.	The user identifies the system required to enter data into the registry. The user may also require information from other systems, such as the clinic's electronic medical record or paper record system.
13	Task	Data input system is selected.	The user selects and opens required tools.
14	Formulate Procedure	Identify which data elements will be required by the system.	The user identifies which data elements will be required to enter data into the registry.
15	Procedure	Desired elements identified and available to enter.	The user knows which values belong to which fields in the registry data entry tool.
16	Execute	User submits data.	The user enters all values and submits data.

**Summarize Encounter in Registry – Novice User (Control Task 1.1.1)**

**CONTROL TASK ANALYSIS**

Summarize Encounter in Registry

**DL 1.1.1**  
**Primary Care**  
**Novice User**



**Summarize Encounter in Registry – Novice User (Control Task 1.1.1)**

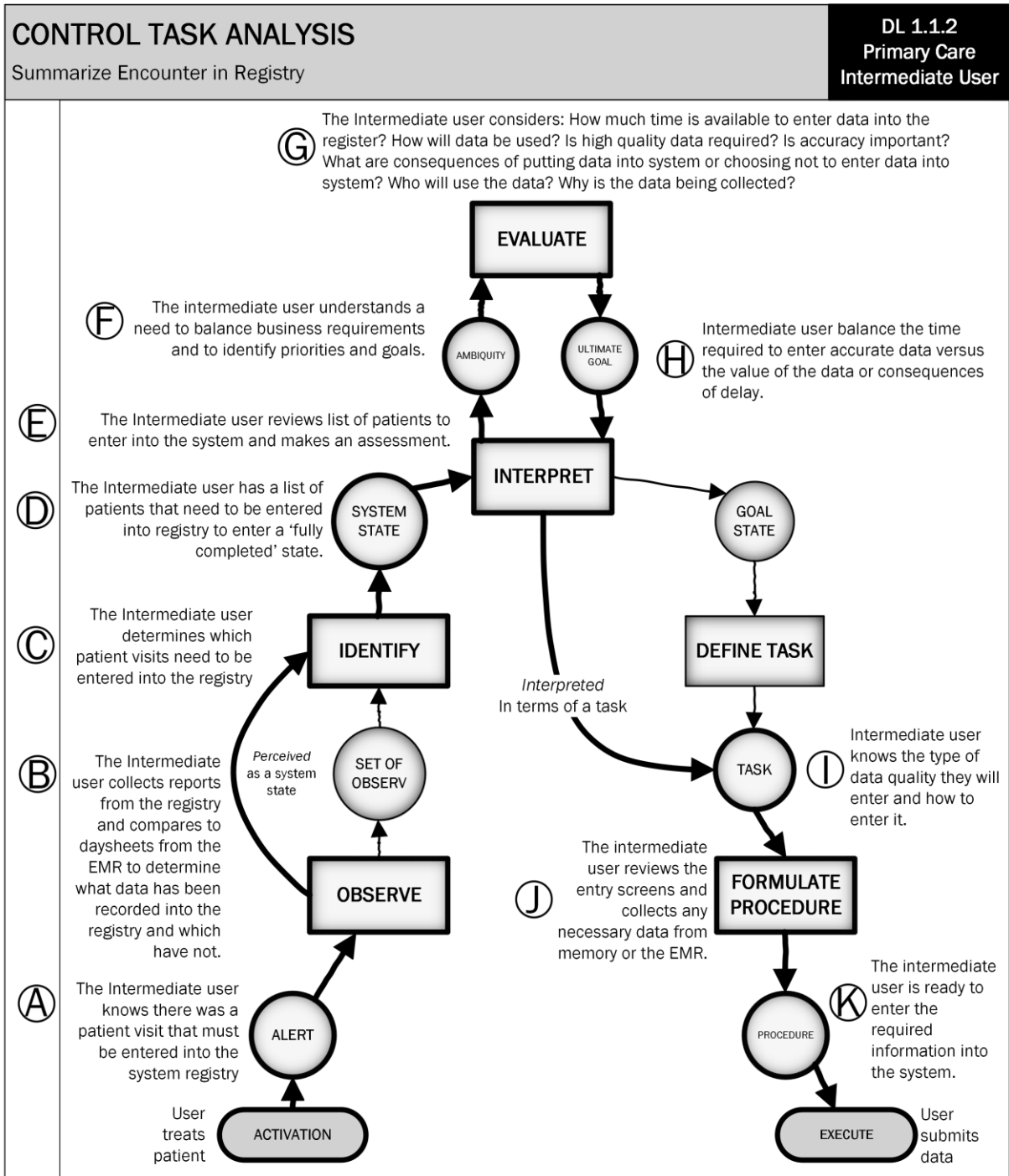
Related Abstraction Hierarchies: AH 1.A, AH 1.B, AH 2.1  
 Related Decision Ladders: DL 1.1, DL 1.1.2 (intermediate), DL 1.1.3 (Expert)

Step	Novice Approach	Type	Ladder Code	Abstraction Level
Begin	User treats patient.	Event	Activation	AH 1.A, Generalized Functions, Assess Patient
A	The novice user knows there was a patient visit that must be entered into the system registry and knows what patient data must be entered.	Knowledge State	Alert	AH 2.1, Generalized Function, Summarize Encounter
-	The user <i>perceives</i> the situation as a System state; data is not entered and must be entered.	Shortcut		-
B	The novice user performs the task immediately after seeing a patient and does not need to identify which patients need to be entered into the registry. No need to observe system or look at reports to identify patients.	Information Processing Activity	Identify	AH 2.1, Abstract Functions, Benchmarks, Norms and Standards
-	User <i>Identifies</i> the situation in terms of deviation from system state.	Shortcut		-
C	The novice user does not challenge the usefulness or purpose of entering data into the registry and follows business rules. Their goal is to enter data accurately and on time.	Knowledge State	Goal State	AH 2.1, Abstract Functions, Benchmarks, Norms and Standards
D	The novice user determines which tools are required to enter data into the registry.	Information Processing Activity	Define Task	AH 2.1, Abstract Functions, Information Flow
E	The novice user understands the tools required to perform the task.	Knowledge State	Task	AH 2.1, Generalized Function, Summarize Encounter

CONTROL TASK ANALYSIS Decision Ladder Summary (Novice User)				DLSUM 1.1.1 Primary Care
F	The novice user reviews the entry screens and collects any necessary data from memory or the EMR.	Information Processing Activity	Formulate Procedure	AH 2.1, Abstract Functions, Information Flow
G	The novice user is ready to enter the required information into the system.	Knowledge State	Procedure	AH 2.1, Generalized Function, Summarize Encounter
End	Novice user submits data.	Event	Execute	AH 2.1, Generalized Function, Data Sink



## Summarize Encounter in Registry - Intermediate User (Control Task 1.1.2)



**Summarize Encounter in Registry - Intermediate User (Control Task 1.1.2)**

Related Abstraction Hierarchies: AH 1.A, AH 1.B, AH 2.1

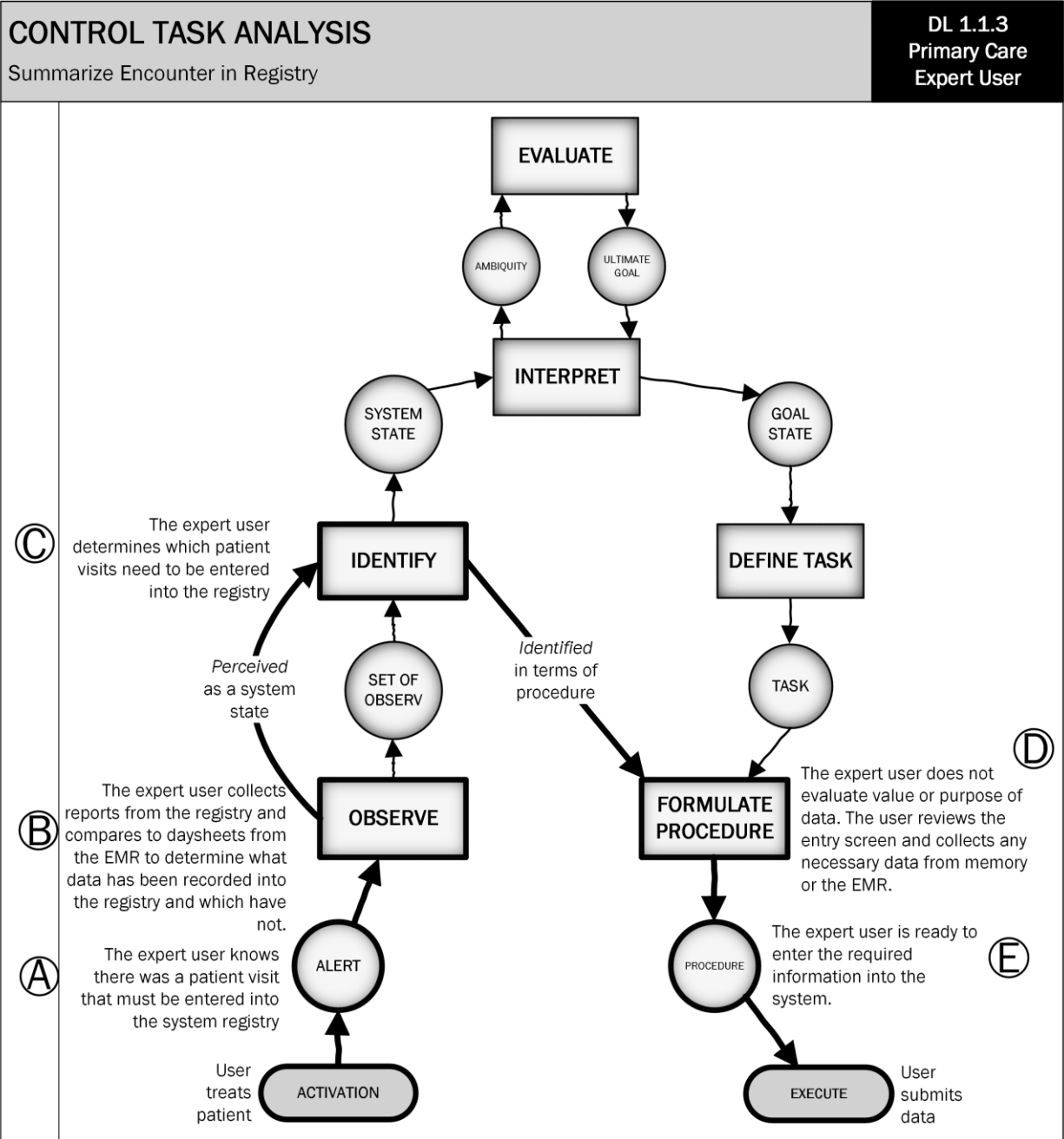
Related Decision Ladders: DL 1.1, DL 1.1.1 (Novice), DL 1.1.3 (Expert)

Step	Intermediate Approach	Type	Ladder Code	Abstraction Level
Begin	Intermediate user treats patient.	Event	Activation	AH 1.A, Generalized Functions, Assess Patient
A	The intermediate user knows there was a patient visit that must be entered into the system registry	Knowledge State	Alert	AH 2.1, Generalized Function, Summarize Encounter
B	The intermediate user collects reports from the registry and compares to daysheets from the EMR to determine what data has been recorded into the registry and which have not.	Information Processing Activity	Observe	AH 2.1, Generalized Function, Find and Review Records
-	The intermediate user <i>perceives</i> the situation as a system state. Data is not entered and must be entered.	Shortcut		-
C	The intermediate user determines which patient visits need to be entered into the registry after observing data from reports and scheduling systems.	Information Processing Activity	Identify	AH 2.1, Generalized Function, Find and Review Records, & AH 2.1, Abstract Functions, Information Flow
D	The intermediate user has a list of patients that need to be entered into registry to enter a 'fully completed' state.	Knowledge State	System State	AH 2.1, Generalized Function, Summarize Encounter
E	The intermediate user reviews list of patients to enter into the system and makes an assessment regarding current priorities and available time.	Information Processing Activity	Interpret	AH 2.1, Generalized Function, Summarize Encounter, Professional Training

CONTROL TASK ANALYSIS Decision Ladder Summary				DLSUM 1.1.2 Primary Care
F	The intermediate user understands a need to balance business requirements and to identify priorities and goals.	Knowledge State	Ambiguity	AH 2.1, Functional Purposes
G	The intermediate user evaluates several questions, such as: How much time is available to enter data into the register? How will data be used? Is high-quality data required? Is accuracy important? What are consequences of putting data into system or choosing not to enter data into system? Who will use the data? Why is the data being collected?	Information Processing Activity	Evaluate	AH 2.1, Abstract Functions, Professional Training, AH 2.1, Generalized Function, Summarize Encounter & AH 2.1, Functional Purposes
H	Intermediate user determines goal: choice to balance the time required to enter accurate data versus the value of the data. Identifies if timeliness of data is important vis-à-vis available time and capacity to see more patients.	Knowledge State	Ultimate Goal	AH 2.1, Abstract Functions, Organizational Efficiencies, Relative Norms, Professional Training
-	Intermediate user <i>interprets</i> situation in terms of a task. The task itself is understood.	Shortcut		-
I	The intermediate user understands the tools required to perform task.	Knowledge State	Task	AH 2.1, Abstract Functions, Information Flow
J	The intermediate user reviews the entry screens and collects any necessary data from memory or the EMR.	Information Processing Activity	Formulate Procedure	AH 2.1, Abstract Functions, Information Flow
K	The intermediate user is ready to enter the required information into the system.	Knowledge State	Procedure	AH 2.1, Generalized Function, Summarize Encounter

CONTROL TASK ANALYSIS Decision Ladder Summary				DLSUM 1.1.2 Primary Care
End	Intermediate user submits data.	Event	Execute	AH 2.1, Generalized Function, Data Sink

**Summarize Encounter in Registry – Expert User (Control Task 1.1.3)**



CONTROL TASK ANALYSIS				DLSUM 1.1.3
Decision Ladder Summary				Primary Care
<b>Summarize Encounter in Registry – Expert User (Control Task 1.1.3)</b>				
Related Abstraction Hierarchies: AH 1.A, AH 1.B, AH 2.1				
Related Decision Ladders: DL 1.1, DL 1.1.2 (Intermediate), DL 1.1.3 (Expert)				
Step	Expert Approach	Type	Ladder Code	Abstraction Level
Begin	Expert user treats patient.	Event	Activation	AH 1.A, Generalized Functions, Assess Patient
A	The expert user knows there was a patient visit that must be entered into the system registry	Knowledge State	Alert	AH 2.1, Generalized Function, Summarize Encounter
B	The expert user collects reports from the registry and compares to daysheets from the EMR to determine what data has been recorded into the registry and which have not.	Information Processing Activity	Observe	AH 2.1, Generalized Function, Find and Review Records
-	The expert user <i>perceives</i> the situation as a system state. Data is not entered and must be entered.	Shortcut		-
C	The expert user determines which patient visits need to be entered into the registry after observing data from reports and scheduling systems.	Knowledge State	Identify	AH 2.1, Generalized Function, Find and Review Records, & AH 2.1, Abstract Functions, Information Flow
-	The expert user <i>interprets</i> the situation in terms of a task	Shortcut		-
D	The expert user does not evaluate value or purpose of data; they have already decided this information in the past and this bias moves forward. The user reviews the entry screen and collects any necessary data from memory or the EMR.	Information Processing Activity	Formulate Procedure	AH 2.1, Abstract Functions, Benchmarks, Norms and Standards, Professional Training

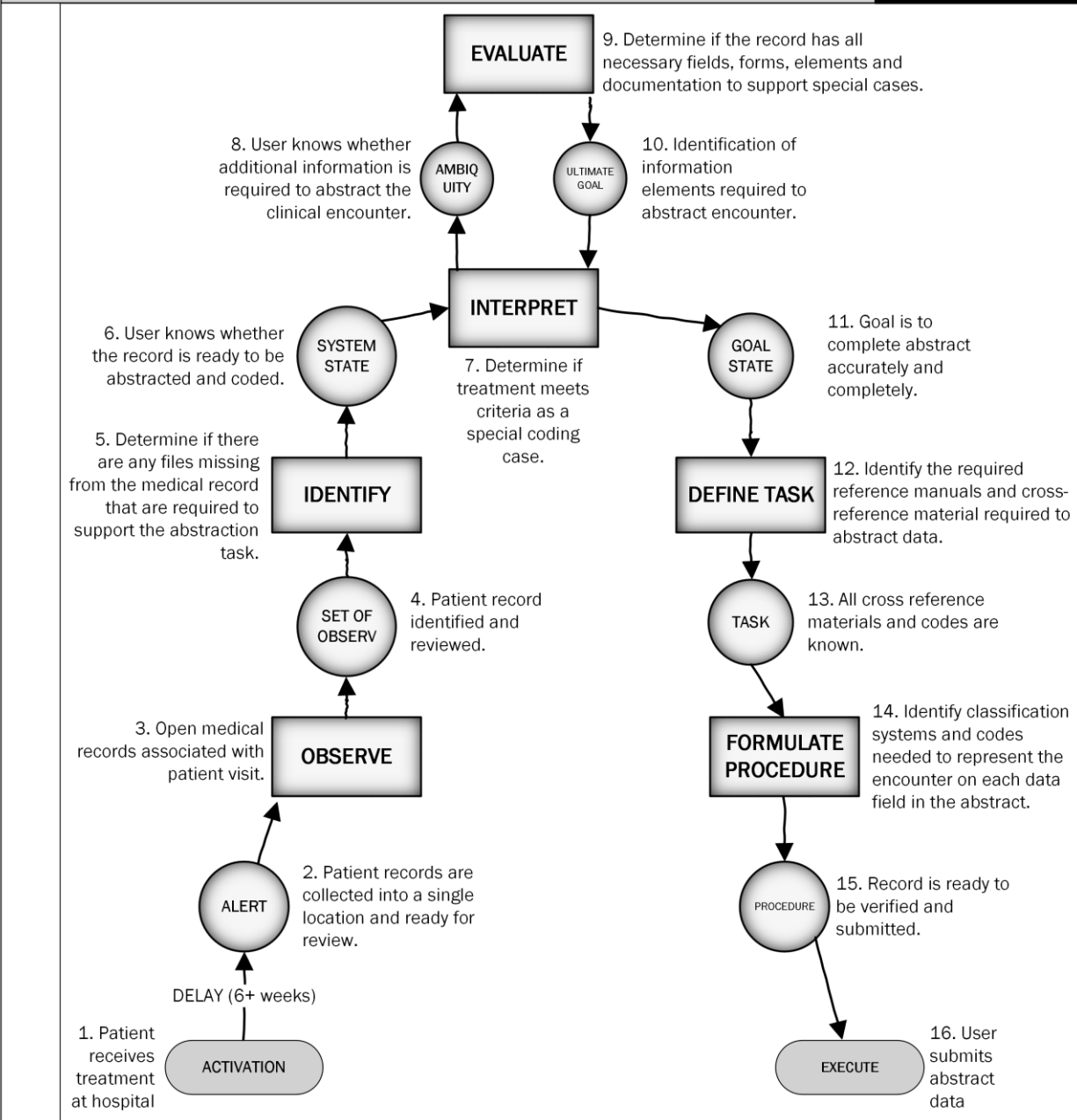
CONTROL TASK ANALYSIS Decision Ladder Summary				DLSUM 1.1.3 Primary Care
E	The user is ready to enter the required information into the system.	Knowledge State	Procedure	AH 2.1, Generalized Function, Summarize Encounter
End	User submits Data	Event	Execute	AH 2.1, Generalized Function, Data Sink

## Abstract Encounter to Registry - Baseline (Control Task 2.1)

### CONTROL TASK ANALYSIS

Abstract Encounter to Registry

DL 2.1  
Hospital

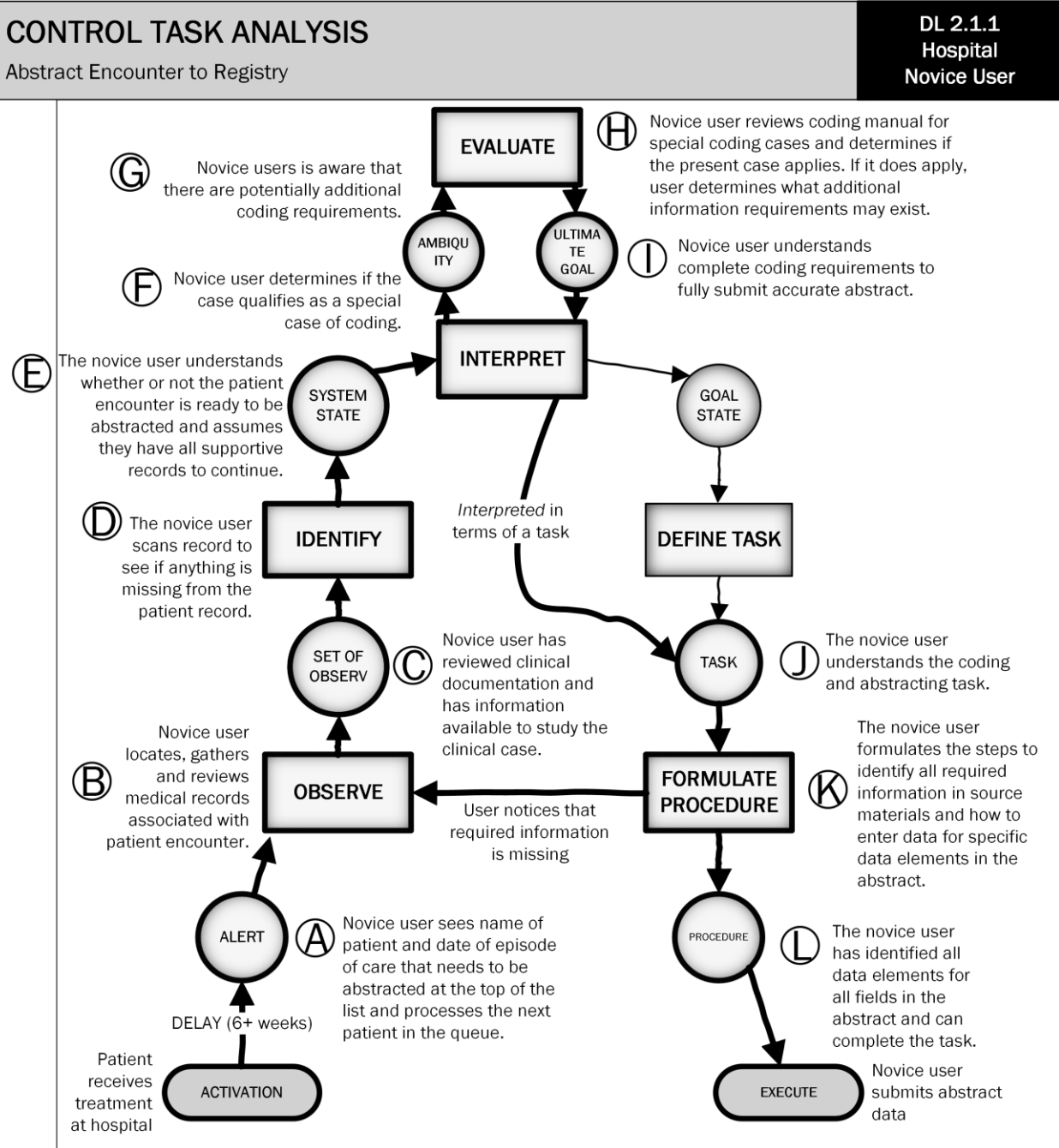




CONTROL TASK ANALYSIS Decision Ladder Summary			DLSUM 2.1 Hospital
<p><b><u>Abstract Encounter to Registry - Baseline (Control Task 2.1)</u></b>            Related Abstraction Hierarchies: AH 1.A, AH 1.B, AH 3.1            Descendant Decision Ladders: DL 2.1.1 (Novice), DL 2.1.2 (Expert)</p>			
Step	Ladder Code	Text	Description
1	Activation	Patient receives treatment at hospital	A patient is seen at the hospital. The patient is <b>NOT</b> seen by the user.
2	Alert	Patient records are collected into a single location and ready for review.	After a 6 week delay (+/-), the patient encounter reaches the top of the queue of the abstractor in the hospital. All patient records are collected into a single place (either digital copies or paper copies) and are ready for review.
3	Observe	Open medical records associated with patient visit.	The user opens the medical records and reviews the content.
4	Set of Observations	Patient record identified and reviewed.	Patient medical records are available and have been reviewed.
5	Identify	Determine if there are any files missing from the medical record that are required to support the abstraction task.	User identifies if there are missing components in the medical record, such as a discharge summary and dictation notes.
6	System State	User knows whether the record is ready to be abstracted and coded.	The record needs to be abstracted. The user knows that the record can be abstracted with available information.
7	Interpret	Determine if treatment meets criteria as a special coding case.	User interprets the clinical case to determine if there are special coding requirements associated with the case. For example, a stroke would require additional coding to capture information about the onset of symptoms and the administration of specific medications.
8	Ambiguity	User knows whether additional information is required to abstract the clinical encounter.	User understands if there is any ambiguity about the clinical case and whether or not there are special coding requirements.

CONTROL TASK ANALYSIS Decision Ladder Summary			DLSUM 2.1 Hospital
9	Evaluate	Determine if the record has all necessary fields, forms, elements and documentation to support special cases.	User evaluates the clinical case and ensures that all required documentation, forms, and coding manuals are available to support the task. Are there special coding tables or rules to understand about a special clinical case?
10	Ultimate Goal	Identification of information elements required to abstract encounter.	The user understands the scope of coding work and what elements must be included in the abstract.
11	Goal State	Goal is to complete abstract accurately and completely.	The user's goal is to complete the abstract with all the required fields and without error.
12	Define Task	Identify the required reference manuals and cross-reference material required to abstract data.	The user identifies the scope of the work and collects references manuals and lists to complete task.
13	Task	All cross reference materials and codes are known.	User has identified and collected all required materials to complete task.
14	Formulate Procedure	Identify classification systems and codes needed to represent the encounter on each data field in the abstract.	User identifies specific codes that will be required for various fields. The user determines how they will represent the clinical case in the abstract.
15	Procedure	Record is ready to be verified and submitted.	The user collects all data and is ready to verify and submit abstract.
16	Execute	User submits abstract data	User submit the abstract data into the registry.

**Abstract Encounter to Registry – Novice User (Control Task 2.1.1)**



CONTROL TASK ANALYSIS				DLSUM 2.1.1
Decision Ladder Summary (Novice User)				Hospital
<b><u>Abstract Encounter to Registry – Novice User (Control Task 2.1.1)</u></b>				
Related Abstraction Hierarchies: AH 1.A, AH 1.B, AH 3.1				
Related Decision Ladders: DL 2.1, DL 2.1.2 (Expert)				
Step	Novice Approach	Type	Ladder Code	Abstraction Level
Begin	Patient receives treatment at hospital	Event	Activation	AH 1.A, Generalized Functions, Treat Patient
A	Novice user sees name of patient and date of episode of care that needs to be abstracted at the top of the list and processes the next patient in the queue.	Knowledge State	Alert	AH 3.1, Generalized Functions, Summarize Encounter
B	Novice user locates, gathers and reviews medical records associated with patient encounter.	Information Processing Activity	Observe	AH 3.1, Generalized Functions, Find and Review Records
C	Novice user has reviewed clinical documentation and has information available to study the clinical case.	Knowledge State	Set of Observation	AH 3.1, Abstract Functions, Information Flow
D	The novice user scans record to see if anything is missing from the patient record.	Information Processing Activity	Identify	AH 3.1, Generalized Functions, Find and Review Records
E	The novice user understands whether or not the patient encounter is ready to be abstracted and assumes all supportive records are present to continue.	Knowledge State	System State	AH 3.1, Abstract Functions, Information Flow
F	Users interprets clinical situation and determines whether additional information is required to abstract the clinical encounter as a special case.	Information Processing Activity	Interpret	AH 3.1, Abstract Functions, Information Flow

CONTROL TASK ANALYSIS Decision Ladder Summary (Novice User)				DLSUM 2.1.1 Hospital
G	Novice user is aware that there are potentially additional coding requirements.	Knowledge State	Ambiguity	AH 3.1, Functional Purposes, Maximize Revenue + Accountability + Monitoring
H	Novice user reviews coding manual for special coding cases and determines if the present case applies. If it does apply, user determines what additional information requirements may exist.	Information Processing Activity	Evaluate	AH 3.1, Functional Purposes, Maximize Revenue + Accountability + Monitoring
I	Novice user understands complete coding requirements to fully submit accurate abstract. Is aware of all special case requirements and extra abstracting needs.	Knowledge State	Ultimate Goal	AH 3.1, Abstract Functions, Information Flow + Benchmarks + Training
-	Novice user <i>interprets</i> work in terms of a task	Shortcut		AH 3.1, Abstract Functions, Training
J	The novice user understands the coding and abstracting task.	Knowledge State	Task	AH 3.1, Abstract Functions, Training
K	The novice user formulates the steps to identify all required information in source materials and how to enter data for specific data elements in the abstract.	Information Processing Activity	Formulate Procedure	AH 3.1, Abstract Functions, Training + Information Flow
-	The novice user may be required to <i>Loopback</i> to re-observe the dataset to find data elements that may be missing. User might spend multiple time circling to identify if information is missing from record or was simply missed in first review.	Loopback		AH 3.1, Abstract Functions, Information Flow

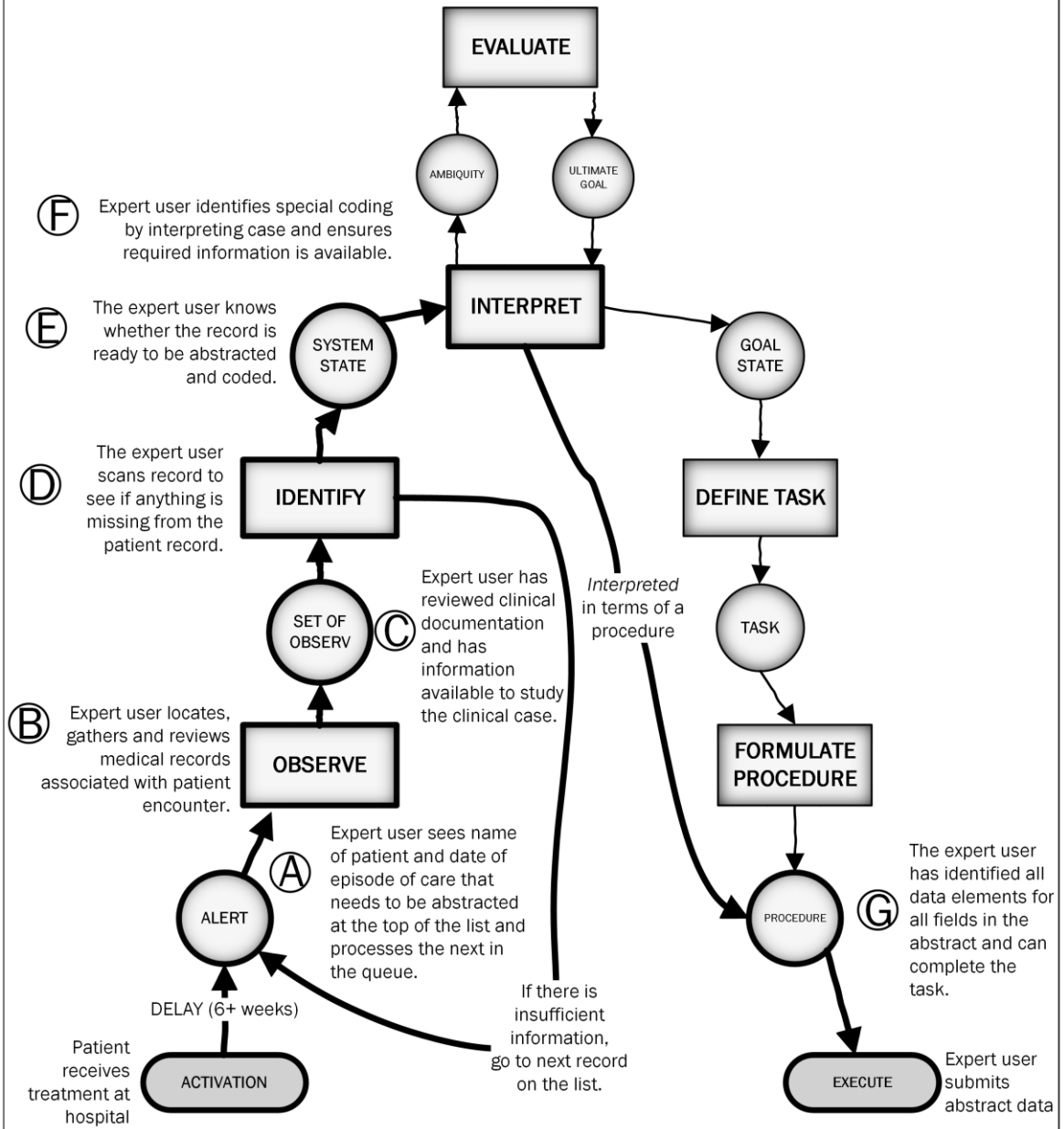
CONTROL TASK ANALYSIS Decision Ladder Summary (Novice User)				DLSUM 2.1.1 Hospital
L	The novice user has identified all data elements for all fields in the abstract and can complete the task.	Knowledge State	Procedure	AH 3.1, Abstract Functions, Training + Information Flow
End	Novice user submits abstract data	Event	Execute	AH 3.1, Generalized Functions, Summarize Encounter

**Abstract Encounter to Registry – Expert User (Control Task 2.1.2)**

**CONTROL TASK ANALYSIS**

Abstract Encounter to Registry

**DL 2.1.2  
Hospital  
Expert User**



CONTROL TASK ANALYSIS Decision Ladder Summary (Expert User)				DLSUM 2.1.2 Hospital
<b><u>Abstract Encounter to Registry – Expert User (Control Task 2.1.2)</u></b> Related Abstraction Hierarchies: AH 1.A, AH 1.B, AH 3.1 Related Decision Ladders: DL 2.1, DL 2.1.2				
Step	Expert Approach	Type	Ladder Code	Abstraction Level
Begin	Patient receives treatment at hospital	Event	Activation	AH 1.A, Generalized Functions, Treat Patient
A	Expert user sees name of patient and date of episode of care that needs to be abstracted at the top of the list and processes the next in the queue.	Knowledge State	Alert	AH 3.1, Generalized Functions, Summarize Encounter
B	Expert user locates, gathers and reviews medical records associated with patient encounter.	Information Processing Activity	Observe	AH 3.1, Generalized Functions, Find and Review Records
C	Expert user has reviewed clinical documentation and has information available to study the clinical case.	Knowledge State	Set of Observations	AH 3.1, Abstract Functions, Information Flow
D	The expert user scans record to see if anything is missing from the patient record.	Information Processing Activity	Identify	AH 3.1, Generalized Functions, Find and Review Records
-	The expert user may initiate a <i>Loopback</i> if they determine there is not sufficient information to properly abstract the clinical encounter, based on their expertise. They will move to next encounter.	Loopback		AH 3.1, Generalized Functions, Request Updates + Abstract Functions, Training
E	The expert user understands whether or not the patient encounter is ready to be abstracted and has all supportive records to continue.	Knowledge State	System State	AH 3.1, Abstract Functions, Information Flow
F	Expert user identifies special coding by interpreting case and ensures required information is available.	Information Processing Activity	Interpret	AH 3.1, Abstract Functions, Information Flow + Training

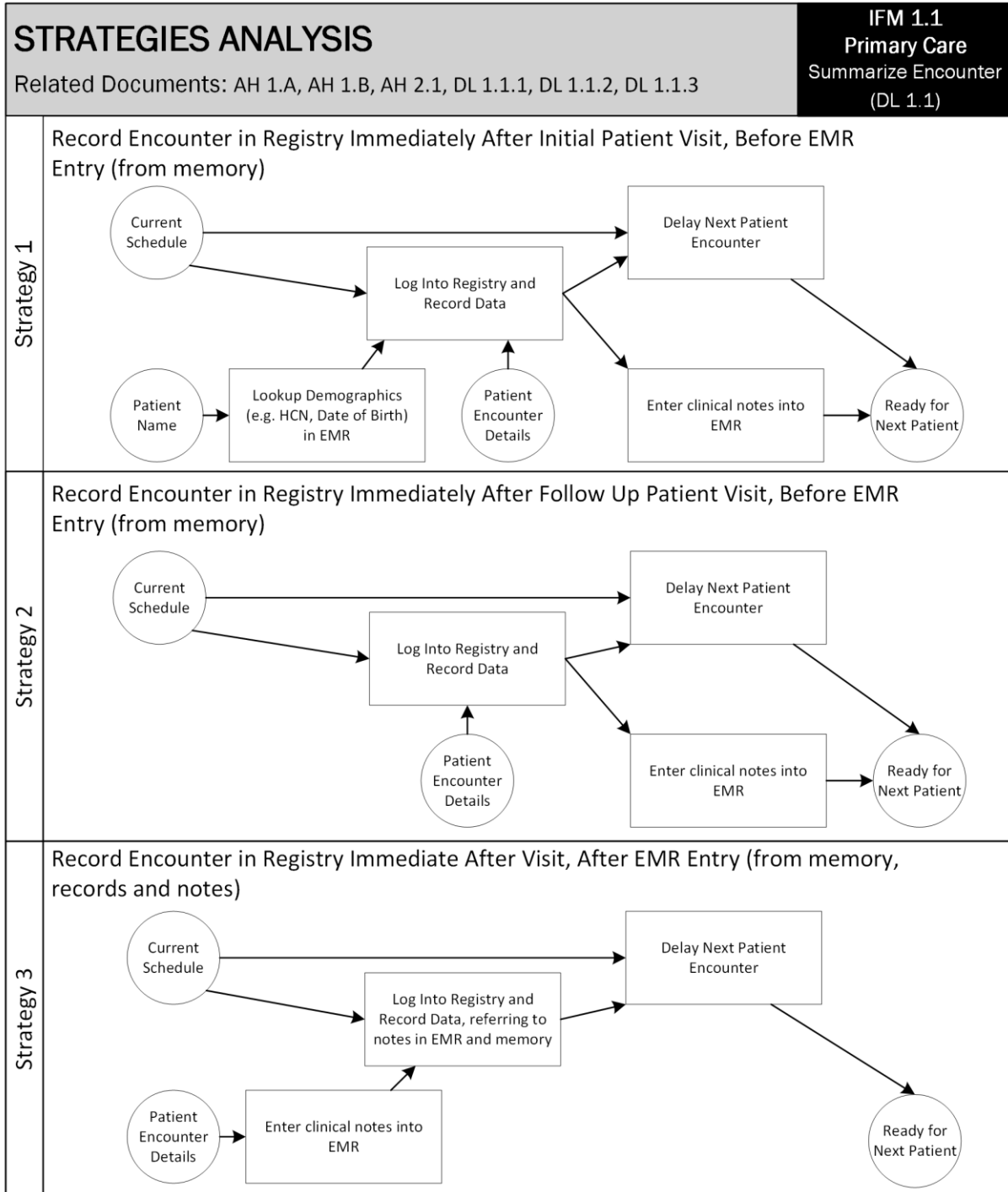


CONTROL TASK ANALYSIS Decision Ladder Summary (Expert User)				DLSUM 2.1.2 Hospital
-	Expert user <i>interprets</i> the work in terms of a procedure.	Shortcut		AH 3.1, Abstract Functions, Training
G	The expert user has identified all data elements for all fields in the abstract and can complete the task.	Knowledge State	Procedure	AH 3.1, Abstract Functions, Training + Information Flow
End	Expert user submits abstract data	Event	Execute	AH 3.1, Generalized Functions, Summarize Encounter

## Appendix D

### Strategies Analysis Flow Maps and Summary Tables

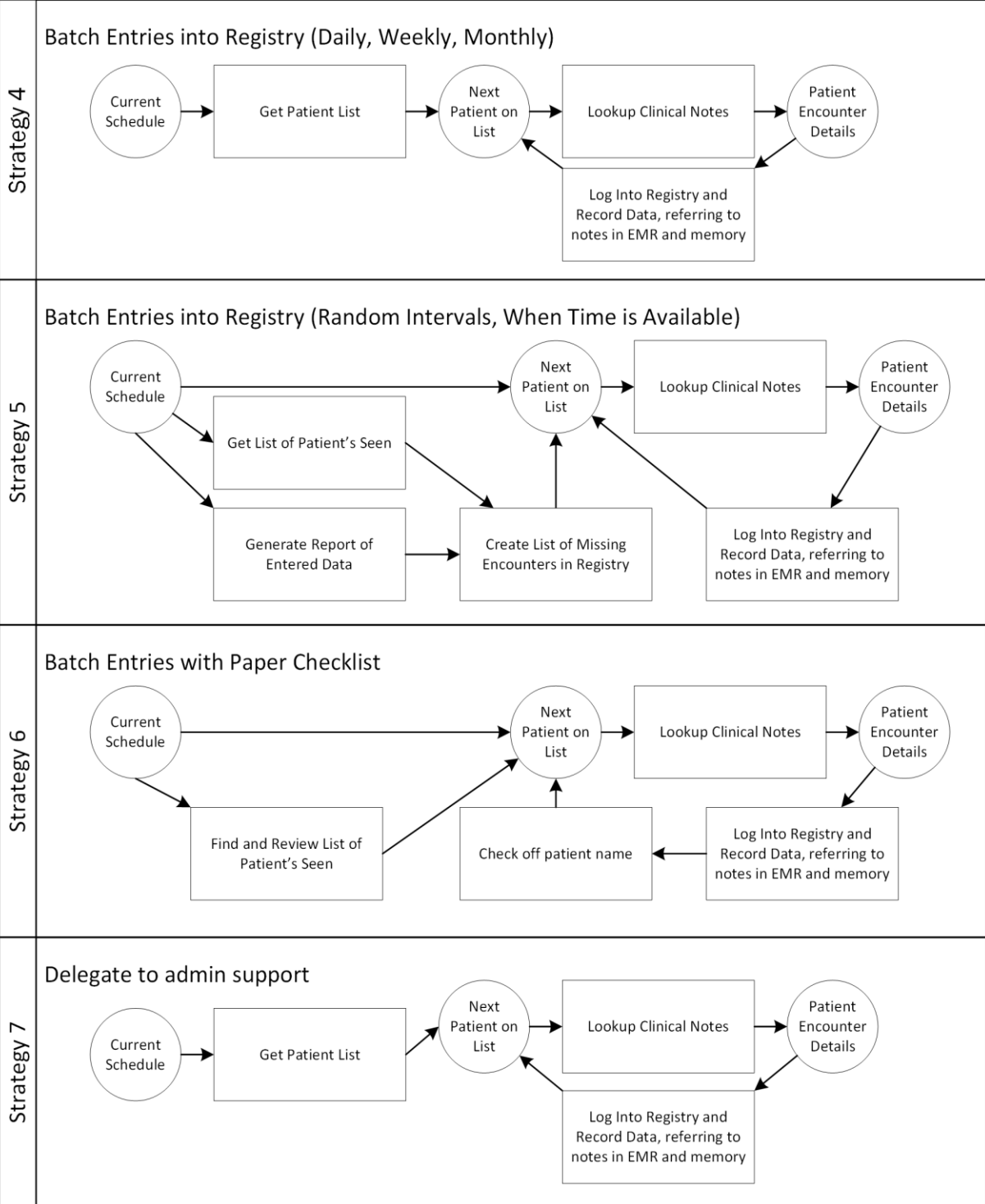
#### Summarize Encounter in Registry (Control Task 1.1)



# STRATEGIES ANALYSIS

Related Documents: AH 1.A, AH 1.B, AH 2.1, DL 1.1.1, DL 1.1.2, DL 1.1.3

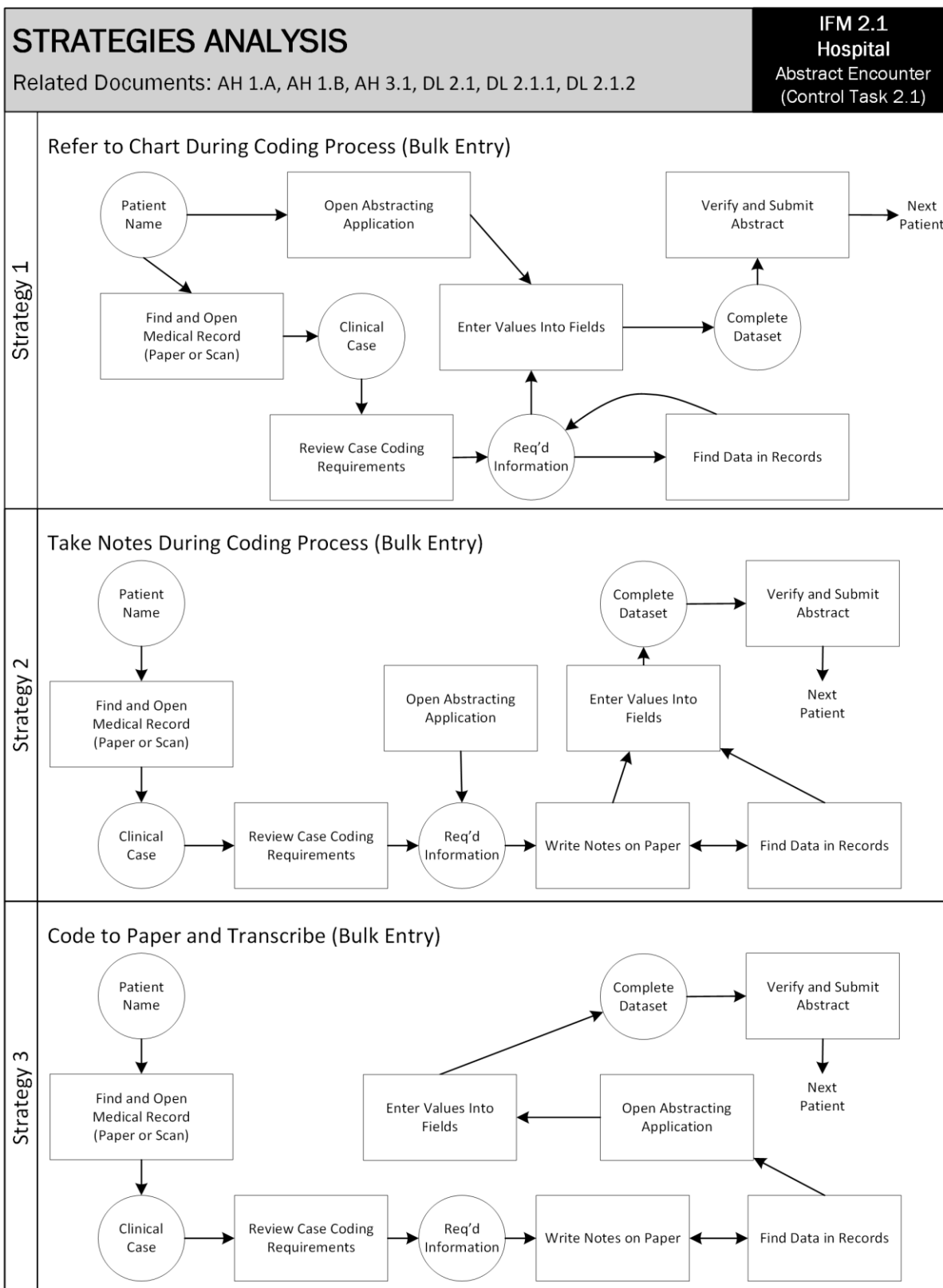
**IFM 1.1**  
**Primary Care**  
 Summarize Encounter  
 (DL 1.1)



STRATEGIES ANALYSIS Information Flow Map Summary			IFSUM 1.1 Primary Care
<b><u>Summarize Encounter in Registry – Baseline (Control Task 1.1)</u></b>			
Abstraction Hierarchies: AH 1.A, AH 1.B, AH 2.1 Decision Ladders: DLSUM 1.1, DLSUM 1.1.1, DLSUM 1.1.2, DLSUM 1.1.3			
ID	Strategy Name	Description	Triggers
1	Record Encounter in Registry Immediately After Initial Patient Visit, Before EMR Entry (from memory)	The user immediately inputs data into the registry after a patient visit. In the situation of an initial visit, certain information needs to be transcribed from the EMR into the registry (e.g. healthcare number) and requires a lookup in the EMR. The user must return to the EMR afterwards to enter clinical notes. With this strategy, subsequent appointments may be delayed to allow the user sufficient time to complete the task. The patient encounter details are in short term memory. This strategy requires using the EMR system twice, before and after entering data into the registry.	<ol style="list-style-type: none"> <li>1. Organizational Policy.</li> <li>2. User personality and values.</li> <li>3. Length of employment.</li> </ol>
2	Record Encounter in Registry Immediately After Follow Up Patient Visit, Before EMR Entry (from memory)	For a follow up visit, the patient name and demographics are already in the registry and do not require a lookup in the EMR system. The user must open the EMR to enter clinical notes after registering the encounter in the registry. Subsequent appointments are delayed. The patient encounter details are in short term memory. This strategy only requires use of the EMR once, after the data is entered into the registry.	<ol style="list-style-type: none"> <li>1. Organizational Policy.</li> <li>2. User personality and values.</li> <li>3. Length of employment.</li> </ol>
3	Record Encounter in Registry Immediate After Visit, After EMR Entry (from memory, records and notes)	Using this strategy, a user records their clinical notes into the EMR first. While it is open, the user enters data into the EMR registry and delays the following patient appointment until the task is complete. The patient encounter details are in short term memory. This strategy requires fewer steps, but requires the user to focus on the clinical documentation task before entering data into the registry.	<ol style="list-style-type: none"> <li>1. Organizational Policy.</li> <li>2. User personality and values.</li> <li>3. Length of employment.</li> </ol>
4	Batch Entries into Registry (Daily, Weekly, Monthly)	By using a batch process, users are able to complete the task without delaying patient encounters. With this strategy, users reserve specific periods of time in their calendar to enter data into the registry. The patient encounter details are remember through the clinical notes and are not in short term memory.	<ol style="list-style-type: none"> <li>1. Workload and schedule constraints.</li> <li>2. Increased experience.</li> </ol>

STRATEGIES ANALYSIS Information Flow Map Summary			IFSUM 1.1 Primary Care
5	Batch Entries into Registry (Random Intervals, When Time is Available)	Similar to the previous batch strategy, but without regular time intervals. In this strategy, the user does not know which encounters have been recorded into the registry and must cross reference entered data (via reports) against the list of patients that have been seen to determine what data is missing. This strategy has an added cognitive load, as the user must generate lists and cross reference information before being able to input data into the registry.	<ol style="list-style-type: none"> <li>1. Workload and schedule constraints.</li> <li>2. Increased experience.</li> </ol>
6	Batch Entries with Paper Checklist	Similar to the previous strategy, except the user keeps a paper list of the patients they have seen and checks names off a printed scheduled to keep track of what patients have been entered into the registry.	<ol style="list-style-type: none"> <li>1. Workload and schedule constraints.</li> <li>2. Increased experience.</li> <li>3. Lack of technical abilities.</li> </ol>
7	Delegate to Administrative Support	The user delegates the data entry task to a clerical team member. The clerical team member reviews medical records and schedules, and records information into the registry on behalf of the clinician.	<ol style="list-style-type: none"> <li>1. Workload and schedule constraints.</li> <li>2. Clinical workflows.</li> <li>3. Team structures.</li> <li>4. Lack of technical abilities.</li> </ol>

## Abstract Encounter to Registry (Control Task 2.1)



**Abstract Encounter to Registry (Control Task 2.1)**  
 Abstraction Hierarchies: AH 1.A, AH 1.B, AH 3.1  
 Decision Ladders: DLSUM 2.1, DLSUM 2.1.1, DLSUM 2.1.2

ID	Strategy Name	Description	Triggers
1	Refer to Chart During Coding Process (Bulk Entry)	The Health Information Management Professional opens the abstracting tool and reviews medical records (paper and electronic) to identify the values required for each field as they complete the abstract. Medical records are reviewed and coded in tandem.	<ol style="list-style-type: none"> <li>1. Department policy.</li> <li>2. Inexperience.</li> </ol>
2	Take Notes During Coding Process (Bulk Entry)	The Health Information Management Professional reviews the medical record before beginning the abstract and takes notes (written or mental) based on anticipated pieces of information that will be required to complete the abstract. Once the abstract entry process begins, the user has pre-identified all data and is able to quickly enter all fields in the abstract without a significant amount of back and forth to the medical records.	<ol style="list-style-type: none"> <li>1. Department policy.</li> <li>2. Experience.</li> </ol>
3	Code to Paper and Transcribe (Bulk Entry)	The Health Information Management Professional reviews the medical records and completes a paper-based abstract instead of taking notes (e.g. uses pre-electronic forms to capture required fields). Once the paper abstract is complete, the user transcribes the information into the software, field by field.	<ol style="list-style-type: none"> <li>1. Department policy.</li> <li>2. Lack of technical abilities.</li> <li>3. Use of historic paper processes.</li> </ol>

## Appendix E

### Worker Competencies Analysis Tables

WORKER COMPETENCIES ANALYSIS				SRK 1.1
SRK Inventory				Primary Care
<b><u>Summarize Encounter in Registry – Baseline (Control Task 1.1)</u></b>				
Related Documents: AH 1.A, AH 1.B, AH 2.1, DL 1.1, DL 1.1.1, DL 1.1.2, DL 1.1.3, IFM 1.1				
Information Processing Step	Result Knowledge State	Skill-Based Behavior	Rule-Based Behavior	Knowledge-Based Behavior
1. User treats patient	2. A patient visit must be summarized and recorded in the reporting system at some time.	Professional training in patient care and treatment.	Apply company policies regarding the entry of data into the registry.	Reason what type of patient encounter constitutes a visit (e.g. phone call vs direct encounter)
3. A patient visit must be summarized and recorded in the reporting system at some time.	4. User knows whether patients have been seen and should be recorded.	Generate list of patients from the EMR schedule.	Apply company policies regarding the maximum amount of delay.	Recall of patients that were treated (knowledge of services provided)
5. Identify which patients need to be entered into the registry.	6. Whether all patient visits have been entered into the registry.	Generate list of patients who have been entered into the registry and compare to the list of patients seen.	Use heuristics to cross-validate two lists of patient names.	Recall patients whose data was entered into the registry.
7. Evaluate the data entry task.	8. User knows whether there is a conflict in the business requirements, amount of available time, or data goals.	Perceive the amount of available time and the amount of time required to perform the task.	Interpret company policies and procedures.	Estimate how the data entered will impact reports and benchmarks. Review current schedule and priorities.



<b>Information Processing Step</b>	<b>Result Knowledge State</b>	<b>Skill-Based Behavior</b>	<b>Rule-Based Behavior</b>	<b>Knowledge-Based Behavior</b>
9. User balances available time vs value proposition of entering data quality vs accurately vs delay.	10. Desired priorities and data entry goals known.	Perceive the value of entering data accurately versus the amount of time it will take. Manage and prioritize time.	Interpret company policies and procedures.	Reason based on experiences with benchmarking and comparisons.
11. Identify the system that will accept registry data.	12. Data input system is selected.	Open browser and open appropriate tool	Follow organizational requirements.	Identify the registry system and URL described during training.
13. Identify which data elements will be required by the system.	14. Desired elements identified and available to enter.	Navigate screen and select appropriate values for fields.	Apply web-form input rules and mandatory field requirements.	Recall details and knowledge from patient encounter.

**Abstract Encounter to Registry - Baseline (Control Task 2.1)**

Related Documents: AH 1.A, AH 1.B, AH 3.1, DL 2.1, DL 2.1.1, DL 2.1.2, IFM 2.1

Information Processing Step	Result Knowledge State	Skill-Based Behavior	Rule-Based Behavior	Knowledge-Based Behavior
1. Patient receives treatment at hospital	2. Patient records are collected into a single location and ready for review.	Retrieving records and identifying lists of work.	Hospital health information management policies and procedures.	Locate files and understand record workflows in hospitals.
3. Open medical records associated with patient visit.	4. Patient record identified and reviewed.	Searching medical records and identifying important sections.	Workflows and standard record processing procedures.	Understand the organizations of the medical record and its content. Understand medical terminology.
5. Determine if there are any files missing from the medical record that are required to support the abstraction task.	6. User knows whether the record is ready to be abstracted and coded.	Review medical record in paper / digital format and look for required content.	Follow hospital and national coding requirement documents.	Understand the elements required in clinical document to support coding task.
7. Determine if treatment meets criteria as a special coding case.	8. User knows whether additional information is required to abstract the clinical encounter.	Review medical record to identify specific fields and their values.	Follow project criteria definitions provided by hospital, province and CIHI.	Understand medical terminology and physiology.
9. Determine if the record has all necessary fields, forms, elements and documentation to support special cases.	10. Identification of information elements required to abstract encounter.	Review medical record in paper / digital format and look for required content.	Follow hospital and national coding requirement documents.	Understand the elements required in clinical document to support coding task.

<b>Information Processing Step</b>	<b>Result Knowledge State</b>	<b>Skill-Based Behavior</b>	<b>Rule-Based Behavior</b>	<b>Knowledge-Based Behavior</b>
12. Identify the required reference manuals and cross-reference material required to abstract data.	13. All cross reference materials and codes are known.	Review list of codes for services, institutions, etc.	Follow hospital and national coding requirement documents.	Understand meaning of special codes and their implication for reporting and funding.
14. Identify classification systems and codes needed to represent the encounter on each data field in the abstract.	15. Record is ready to be verified and submitted.	Open Folio product and lookup ICD-10 codes. Open and navigate abstracting software. Select appropriate values for fields.	Follow user interface input rules and complete mandatory field requirements.	Understand medical terminology and principles for input into Folio tool. Understand CMG calculation concepts and codes. Use professional ethic guidelines.

## Appendix F

### Qualitative Results of Data Quality Measure Interviews

DQ DIMENSION	ORGANIZATIONAL CONCERNS	MEASUREMENT	FEASIBILITY (per analysts)
Consistent Representation and Interpretability	There is a difference between the language used by “IT” and the language used by the management team. In this context, is the data reflective of reality? Does data properly allow users to differentiate between various situations? If a user selects an item from a list, what does it mean? The definition of each data point in the drop down menus can be interpreted by users in different ways.	Not seen as feasible.	Out of scope. No way to analyse historic trends from dataset or measure user tradeoffs.
Accuracy, Integrity, and Validity	How does the management team know if the data is actually answering the question being asked? Is it reliable? Is the user inputting data correctly? Accuracy and validity problems are likely negatively associated with completeness, since forcing users to enter “something” will force them to enter “anything”.	Measuring non-logical entries (e.g. a visit date being recorded before it occurred, or a date prior to system inception) would be a good reflection of data accuracy, validity and attention to detail.	Possible to program SQL to identify non-logical entries and review historic data.
Completeness	There have been issues with incomplete data, specifically with referrals. This is because the field was not always mandatory. Previously, it was difficult to report on a per-doctor basis without this referral data being entered and reporting was distorted. This caused a loss in trust in the data and reports and it was ultimately caused by user’s entering incomplete data.	Measuring how many records have data in all fields would be a good way to measure completeness. This would be as simple as seeing how many fields are missing and therefore incomplete. A High level of completeness shows higher attention to detail.	Possible to program SQL to identify incomplete entries and review historic data.

DQ DIMENSION	ORGANIZATIONAL CONCERNS	MEASUREMENT	FEASIBILITY (per analysts)
Timeliness	The issue of timeliness is important. It is a good reflection of how the system is being used and how much time it takes to enter data. A delay in entering data is a good reflection of data entry habits and ‘data hygiene’. As the delay in time of recording data increases, details get lost and the data is more general. Delays in data entry probably impact data quality negatively.	The difference between the entry date and the actual encounter day is a reasonable way to measure timeliness. If delay goes up, data quality and specificity likely go down.	Possible to program SQL to calculate this delay and review historic data.
Believability	Sometimes there is a trust issue with the data and the data needs to be double checked with information in EMR system. Generally, though, users accept the measures from the system to assess patient volume.	Not seen as feasible.	Out of scope. No way to analyse historic data or tradeoffs.
Use, Usefulness, Usability, and Accessibility	Sometimes it is not clear why data is being measured and why it is collected. Is anyone using it? Allowing users to view their own data would help improve data quality. Some users do not seem to know this is possible and think that the system is a ‘manager thing’ that consumes data and provides no benefit to them.	Measuring how many times people generate reports from the system might be a good way of measuring usefulness.	Possible to see how often a user opens the “Reports” feature of the system.

# Appendix G

## Registry System Screenshots

### RECORD INDIVIDUAL ENCOUNTERS

**\* Denotes a required field**

Appointment Date\*

Clinic

Type of Encounter\*

Referral Date  I know the date the referral was received

Referral Source\*

#### ARRIVAL STATUS

On Time  Late  Cancellation  Late Cancellation  No Show

#### PATIENT INFORMATION

Patient\*

Presenting Problem\*

Show list of other topics discussed

#### APPOINTMENT OUTCOME

Visit Outcome\*

Show list of secondary outcomes

Duration Of Encounter(minutes)\*

Other Patient Specific Time (minutes)

Comments

## Appendix H

### Additional Quantitative Data

#### Data Grouped by Delay

DELAY CATEGORY	# RECORDS
Same Day	99,878 (50.7%)
7 days (1 week)	42,152 (21.4%)
14 days (2 weeks)	13,432 (06.8%)
28 days (4 weeks)	18,477 (09.4%)
56 days (8 weeks)	14,446 (07.3%)
84 days (12 weeks)	5,041 (02.6%)
120 days (17 weeks / 4 months)	1,313 (00.7%)
Erroneous (NULL)	2,228 (01.1%)

#### Data Grouped by Month

MONTH	# RECORDS
January	16,910 (8.6%)
February	15,437 (7.8%)
March	17,285 (8.8%)
April	17,596 (8.9%)
May	17,303 (8.8%)
June	17,190 (8.7%)
July	15,791 (8.0%)
August	13,760 (7.0%)
September	17,013 (8.6%)
October	18,065 (9.2%)
November	17,087 (8.7%)
December	13,530 (6.9%)

**Data Grouped by Year**

<b>YEAR</b>	<b># RECORDS</b>
2012	43,040 (21.9%)
2013	50,288 (25.5%)
2014	52,267 (26.5%)
2015	51,372 (26.1%)

**Data Grouped by Program Area**

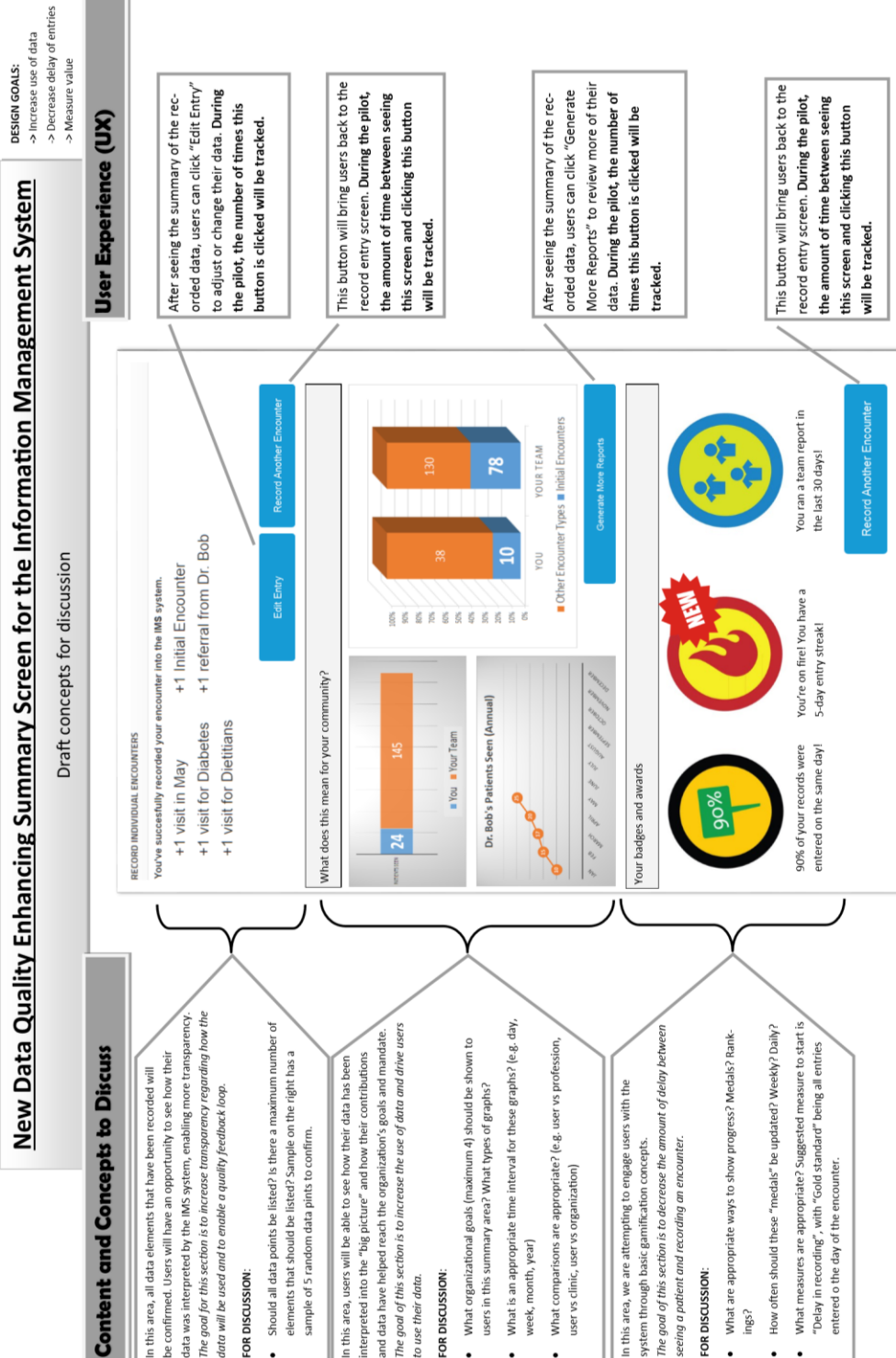
<b>PROGRAM AREA</b>	<b># RECORDS</b>
Counselling Services	51,604 (26.2%)
Diabetes Care Services	55,599 (28.2%)
Health Education and Nursing Support	42,869 (21.8%)
Health Promotion Activities	10,131 (5.1%)
Homecare Services	9,842 (5.0%)
Nutrition Consultations	17,070 (8.7%)
Pharmacy Consultations	9,852 (5.0%)



# Appendix I

## Data Quality Design Worksheets

1st Draft – Initial Proposed Design - August 2016



**RECORD INDIVIDUAL ENCOUNTERS**

You've successfully recorded your encounter into the IMS system.

+ 1 visit in [[current month]]  
 + 1 initial encounter [[OR]] +1 follow up encounter = 5 visits this week

You have created ## care plans this week

Edit Entry
Record Another Encounter

**How did this change your current reporting statistics?** (help)

No Shows  
34%

Visits  
66%

Initial Visits  
19%

Follow Up  
Visits  
81%

Visits Over Last 10 Business Days

Visits vs No Show & Cancellations

Generate More Reports

**Badges and awards** (help)

90% of your records were entered on the same day!

You're on fire! You have a 5-day entry streak!

You ran a report in the last 60 days!

Record Another Encounter

**50.123% of users entered their data on the same day this quarter. Let's improve!**

Summarize data that was inputted

Text summary.

**\*\*\* PROGRAMMING NOTE \*\*\* Not clear on what to summarize. Should discuss.**

Show Visits vs No Shows as pie chart

Show Follow Ups vs Initial Visits as pie chart

Show summary of last 10 business days. Show no-shows & visits as stacked column graph.

**\*\*\* PROGRAMMING NOTE \*\*\* To use to time tracker data to understand what someone's last 10 days were for non FTE, vacation, sick, etc. This would form the X axis.**

Different badges for 40%, 50%, 80%, 99% "Entered on the same day". Time interval should be this quarter.

Different badges for 2 day, 4 day, 8 day, 10 day+, streaks.

Different badges for 1 report, 2 report, 3 report, 4 report, 5 report

**\*\*\* PROGRAMMING NOTE \*\*\* Need to figure out some mechanics. What if user doesn't have a badge? Placeholder? What would place holder say? How does users know they want to try and achieve this badges? Perhaps a small popup when you click the (help) link ?**

# Appendix J

## Data Quality Summary Screenshot

✓ Encounter recorded successfully.

### Your Updated Data Based On Your Entry

+1 Visit in April

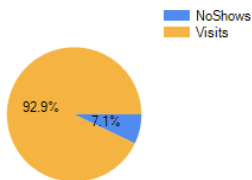
+1 Follow Up

You have recorded 7 visits this month

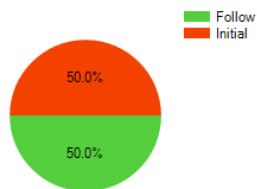
[EDIT ENTRY](#)

### How did this change your current reporting statistics?

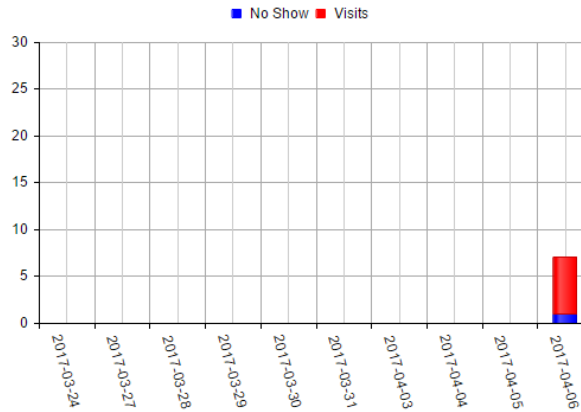
Visits vs No shows  
(Past 3 Months)



Follow Up vs Initial  
(Past 3 Months)



Visits Over Last 10 Business Days



[REVIEW MY STATS](#)

### Badges and Awards



You have earned the 90% badge!

90% or more of your records were entered on the same day for the past two weeks!

Your percentage: **100%**

Did you know? 79.8% of all FHT encounters were entered on the same day for the past two weeks.

[RECORD ANOTHER ENCOUNTER](#)

## **Appendix K**

### **Survey Questions**

**NOTE: Questions below were slightly modified to mask the identity of the FHT.**

#### **User Data (page 1)**

1. Over the last 2 months, how often have you signed into reporting system to enter data?
2. How long have you used the reporting system?

#### **Usability Questions (Likert scale, 1 to 5) – based on SUS – page 2**

1. I like using the reporting system.
2. I find the reporting system unnecessarily complex
3. I find the reporting system easy to use
4. I find that I need the support of a technical person to be able to use the system
5. I find the various functions (data entry, reporting, etc.) in this system are well integrated
6. I think there was too much inconsistency in the reporting system
7. I would imagine that most people would learn to use the reporting system very quickly
8. I find the system very cumbersome to use
9. I feel very confident using the system
10. I needed to learn a lot of things before I could get going with the reporting system
11. I feel motivated to put information into the reporting system.
12. I feel like the data I am entering into reporting is collected and used for a meaningful purpose.
13. I think the system captures high quality data.

#### **New Workflow Questions – page 3**

The FHT wants to help you understand encourage you to use your data on a more regular basis and enter data on a more regular basis. Over the last two months, a new summary screen was deployed. The summary screen is shown below [screenshot].

1. I like the new summary screen.
2. I find the new summary screen unnecessarily complex.
3. I find the new summary screen easy to understand.
4. I find that I need the support of a technical person to explain what I'm seeing on the new summary screen.
5. I find this new summary screen integrates well with the rest of the system.
6. I think there was too much inconsistency with the new summary screen.
7. I would imagine that most people would understand the new summary screen very quickly.
8. I find the new summary screen to be cumbersome.
9. I feel very confident interpreting the new summary screen.
10. I needed to learn a lot before I could understand the new summary screen.
11. The new summary screen had a noticeable impact on my overall workload.
12. It took more time to enter data into the system because of this screen.

13. The new screen clarifies how my entry is recorded into the system.
14. The new screen motivated me to enter data more carefully.
15. The new screen motivated me to enter my data on the same day as I saw a patient.
16. The new screen helps me understand how the data I enter will be used.

Which graph / element do you find the most interesting?

Which graph / element do you find the least interesting?

#### **Comments and Suggestions – Page 4**

Do you have any comments about the new summary screen?

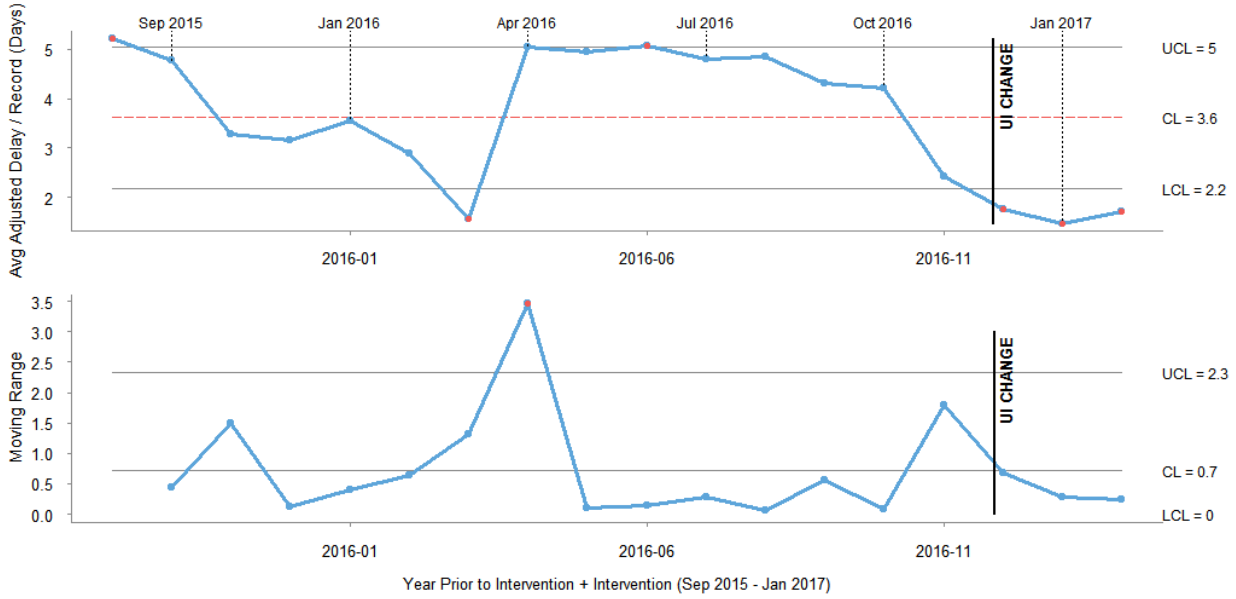
Do you have any comments about the reporting system?

How could you be motivated to enter accurate, complete and timely data into the reporting system?

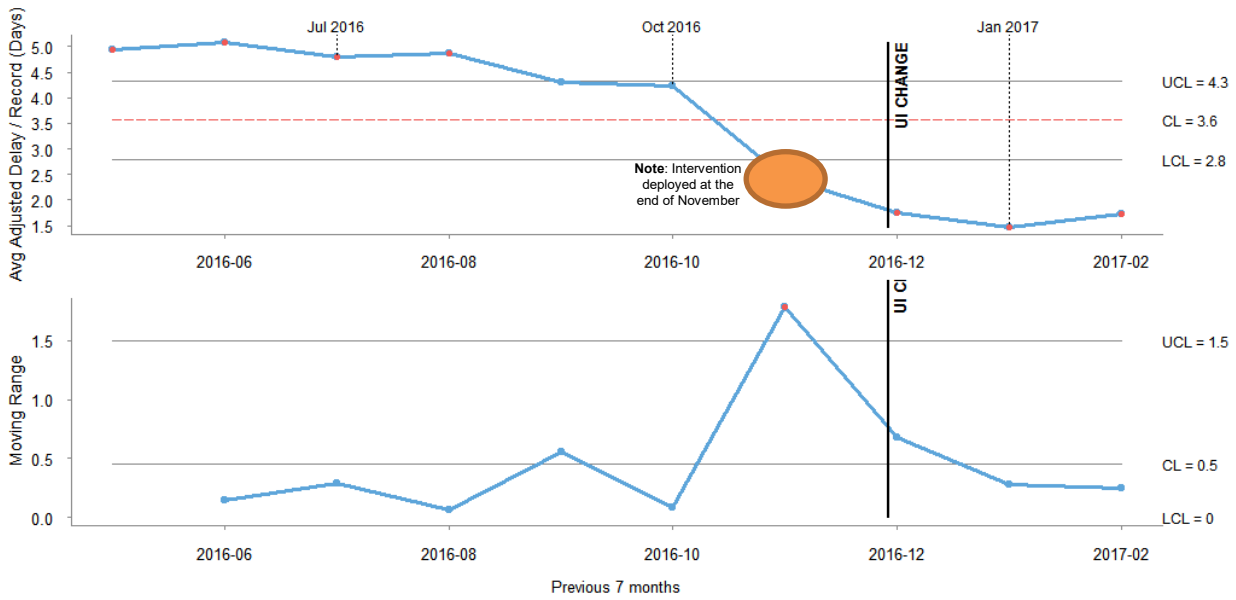
# Appendix L

## XmR Charts by Month

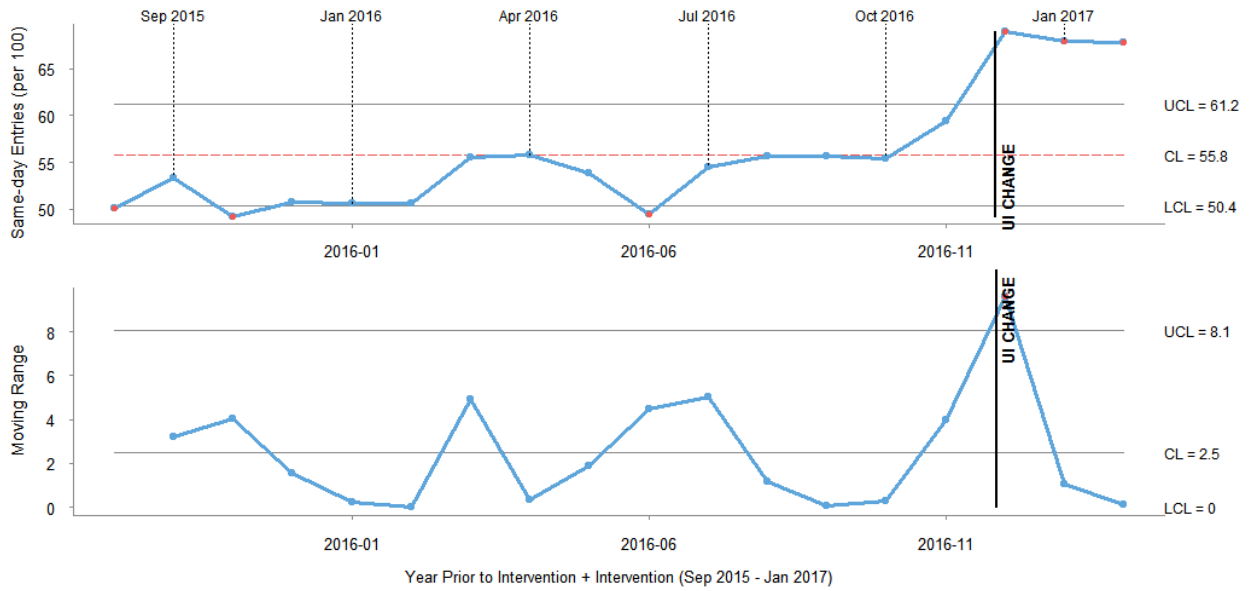
**XmR Control Chart of Average User Adjusted Recording Delay (Timeliness)**



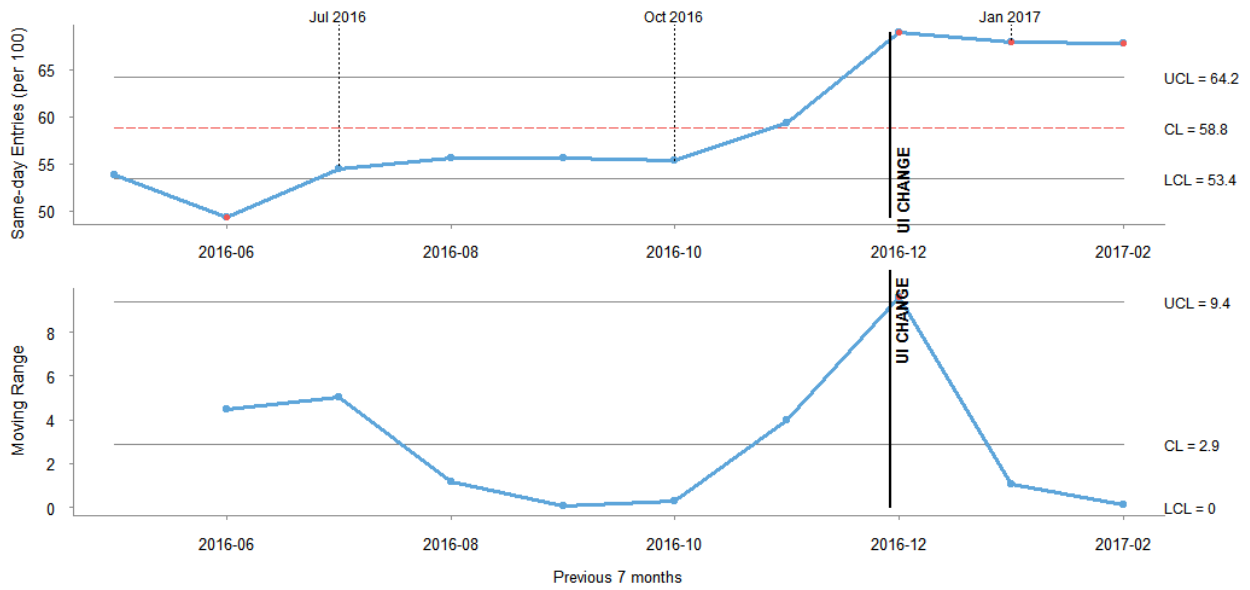
**XmR Control Chart of Average User Adjusted Recording Delay (Timeliness)**



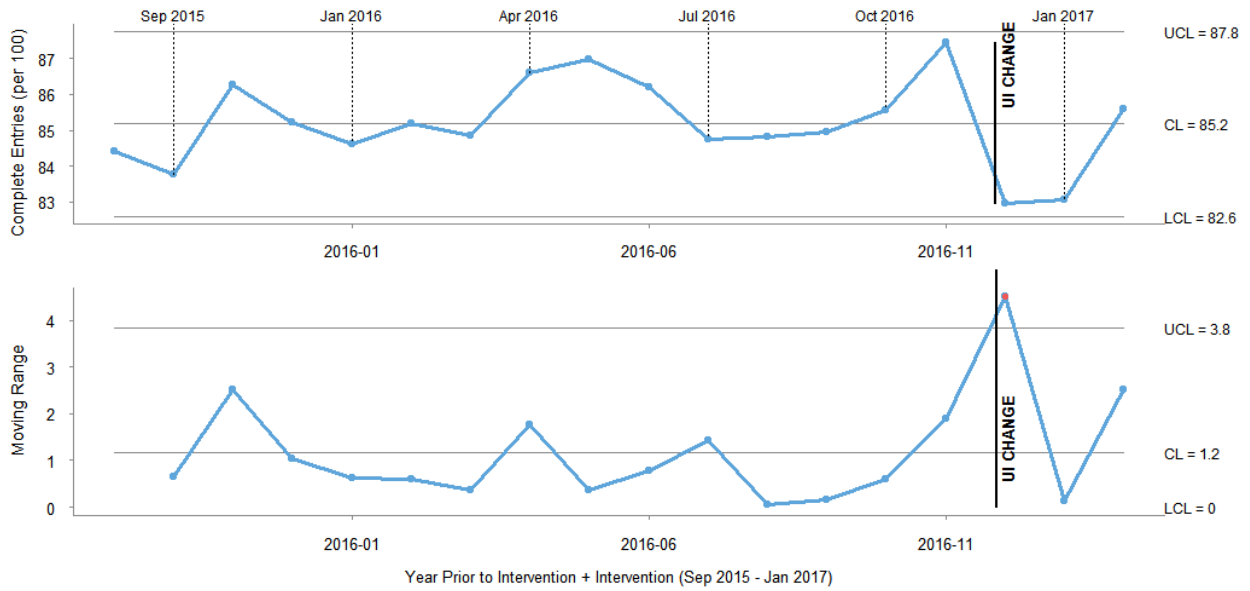
**XmR Control Chart of User Same-day-Entry Measure (Timeliness)**



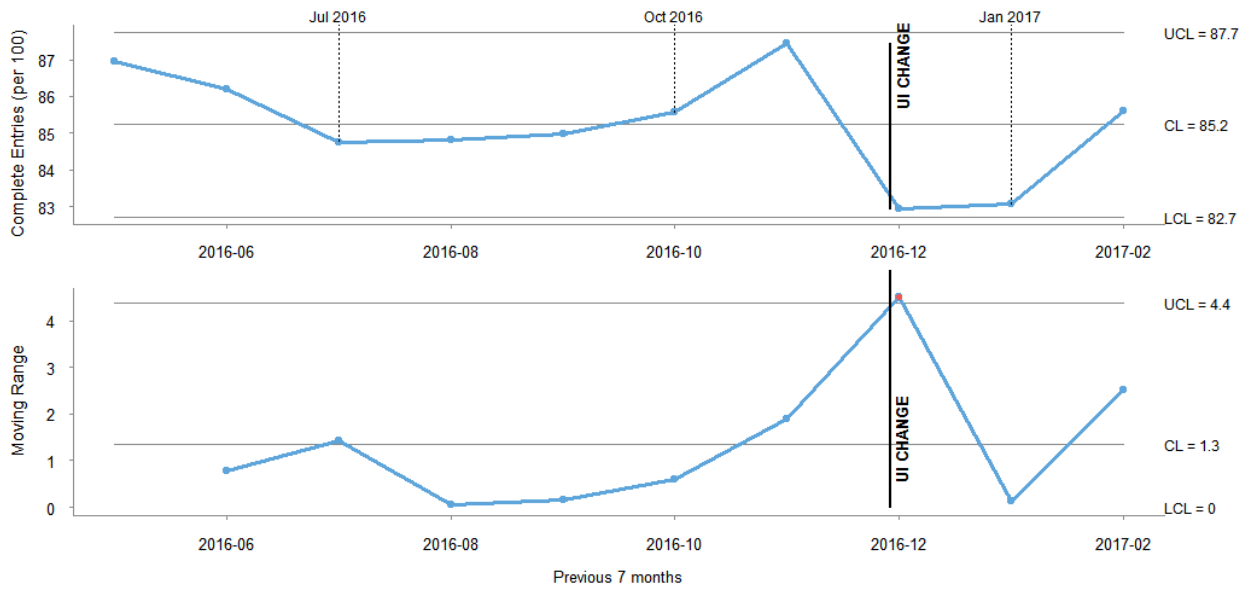
**XmR Control Chart of User Same-day-Entry Measure (Timeliness)**



**XmR Control Chart of User Complete-Entry Measure (Completeness)**

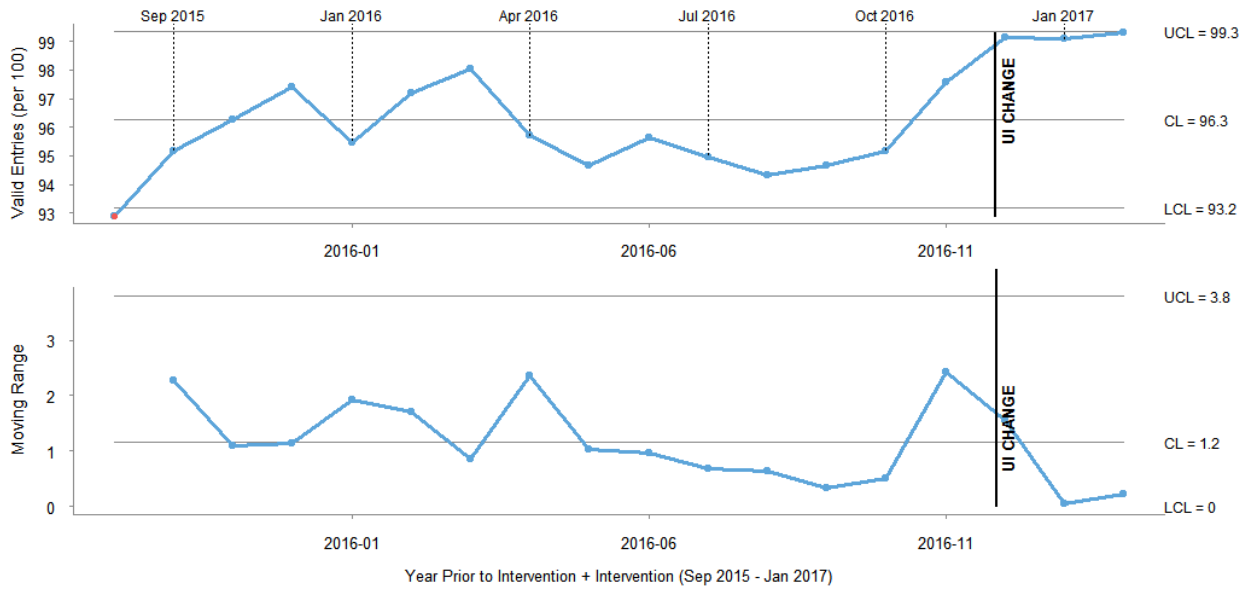


**XmR Control Chart of User Complete-Entry Measure (Completeness)**

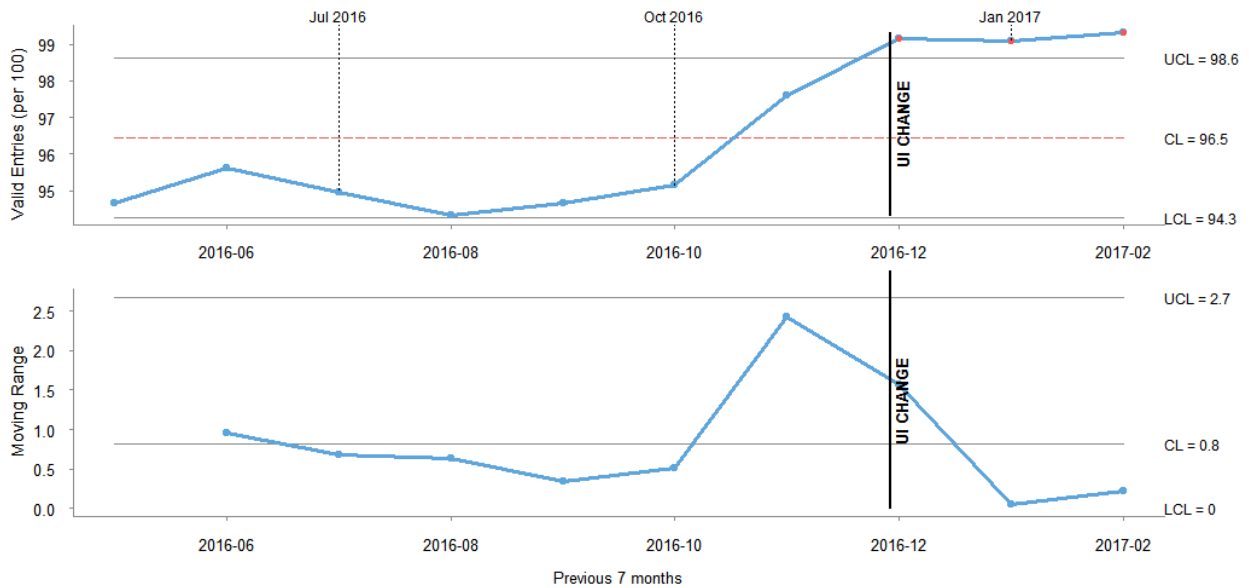




**XmR Control Chart of User Valid-Entry Measure (Validity)**



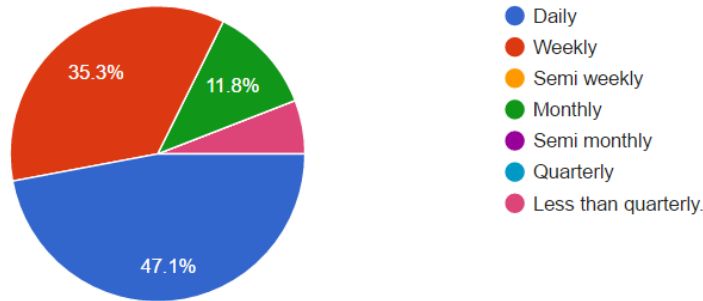
**XmR Control Chart of User Valid-Entry Measure (Validity)**



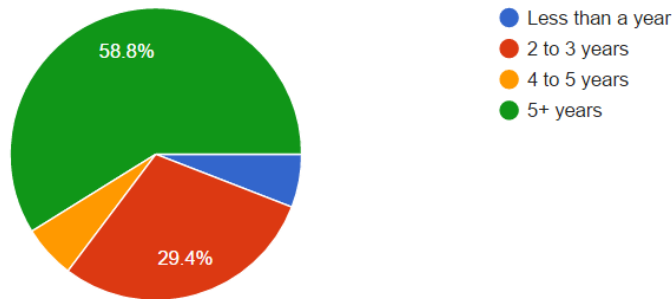
## Appendix M

### Survey Results

How often do you sign into the reporting system to enter data? (17 responses)



How long have you been a user of the reporting system? (17 responses)



Please answer the questions below regarding the REPORTING system.

Question (1 = Strongly Disagree, 5 = Strongly Agree)	Average [95% CI LL-UL]	Median	Mode
[SUS1] I like using the reporting system.	3.18 [2.69 - 3.67]	3	3
[SUS2] I find the reporting system unnecessarily complex	2.53 [1.98 - 3.08]	2	2
[SUS3] I find the reporting system easy to use	3.59 [3.18 - 4.00]	4	4
[SUS4] I find that I need the support of a technical person to be able to use the system	1.65 [1.34 - 1.96]	2	2
[SUS5] I find the various functions (data entry, reporting, etc.) in this system are well integrated	3.00 [2.49 - 3.51]	3	4
[SUS6] I think there was too much inconsistency in the reporting system	2.53 [2.08 - 2.98]	2	2
[SUS7] I would imagine that most people would learn to use the reporting system very quickly	3.71 [3.35 - 4.06]	4	4
[SUS8] I find the system very cumbersome to use	2.81 [2.29 - 3.34]	2.5	2
[SUS9] I feel very confident using the system	3.94 [3.53 - 4.35]	4	4
[SUS10] I needed to learn a lot of things before I could get going with the reporting system	2.24 [1.81 - 2.66]	2	2
I feel motivated to put information into the reporting system.	2.88 [2.37 - 3.39]	3	4

Question (1 = Strongly Disagree, 5 = Strongly Agree)	Average [95% CI LL-UL]	Median	Mode
I feel like the data I am entering into reporting is collected and used for a meaningful purpose.	3.29 [2.73 - 3.86]	4	4
I think the system captures high quality data.	2.94 [2.52 - 3.37]	3	3

**In November, a new summary screen was introduced into the reporting system. Please answer the following questions about the new feature.**

Question (1 = Strongly Disagree, 5 = Strongly Agree)	Average [95% CI LL-UL]	Median	Mode
[SUS1] I like the new summary screen.	3.18 [2.62 - 3.73]	4	4
[SUS2] I find the new summary screen unnecessarily complex.	2.47 [2.06 - 2.88]	2	2
[SUS3] I find the new summary screen easy to understand.	3.82 [3.62 - 4.03]	4	4
[SUS4] I find that I need the support of a technical person to explain what I'm seeing on the new summary screen.	2.06 [1.63 - 2.48]	2	2
[SUS5] I find this new summary screen integrates well with the rest of the system.	3.06 [2.63 - 3.48]	3	3
[SUS6] I think there was too much inconsistency with the new summary screen.	2.47 [2.15 - 2.79]	2	2
[SUS7] I would imagine that most people would understand the new summary screen very quickly.	3.59 [3.18 - 4.00]	4	4
[SUS8] I find the new summary screen to be cumbersome.	3.35 [2.84 - 3.87]	4	4
[SUS9] I feel very confident interpreting the new summary screen.	3.82 [3.55 - 4.10]	4	4
[SUS10] I needed to learn a lot before I could understand the new summary screen.	1.82 [1.45 - 2.20]	2	2
The new summary screen had a noticeable impact on my overall workload.	2.94 [2.48 - 3.40]	3	3
It took more time to enter data into the system because of this screen.	3.29 [2.53 - 4.06]	4	4
The new screen clarifies how my entry is recorded into the system.	2.94 [2.41 - 3.47]	3	3
The new screen motivated me to enter data more carefully.	2.47 [1.92 - 3.02]	3	3
The new screen motivated me to enter my data on the same day as I saw a patient.	2.53 [1.95 - 3.11]	3	3
The new screen helps me understand how the data I enter will be used.	2.59 [2.07 - 3.10]	3	3

**Do you have any comments about the reporting system?**

<b>Respondent</b>	<b>Comment</b>
1	Do not like the time out feature
2	The reporting system captures information on patient's that we see. When we use the log form for phone calls, admin information, training others, etc. it doesn't always capture what we do and could be improved. For patient's that are a no show - could we enter what they were coming to see us for? It seems like a common theme for smoking, COPD etc. Can this be captured?
5	It feels slower to load pages and move between data entry to the new summary screen.
6	could better capture information. Seems that some of the categories are too simplistic. I do not like that late cancellation, cancellations, and no shows are lumped as no shows in the stats.
9	Many of my patient encounters are phone calls - I don't really feel they are captured appropriately because all though they are grouped as phone calls in "type of encounter", the "arrival status" portion doesn't really make sense for a phone call.
10	It would save considerable time if the date selected (the very first step) would remain the same each time (until changed by the user) rather than constantly defaulting to the current date.
13	I do like the summary screen, but it is annoying that I have to take an extra step to get back into entering individual encounters. I work from 3 different locations, and every time I have to reenter into individual encounter I need to re-select my location. It doesn't sound like much, but when you have 8-12 encounters a day it adds up.
14	It's tedious and the stats that are most important for my job are not captured or used.
16	the date needs to stay the same until I change it, instead of entering it with every pt encounter please

**Do you have any comments about the new summary screen?**

Respondent	Comment
1	Busy screen. Too much information to look at while entering stats
2	it is great, but takes a lot to add another same day entry.
5	It makes me feel anxious and unhappy to see a lot of no shows.
6	The new screen data seems to put more unnecessary pressure on data entry
7	Looks good
8	Even though it only take a few seconds for the new screen to load and then a few more seconds to click "record encounters" and for that screen to load, it really adds up! Stats seem to take way longer now.
9	I like seeing the graphs - I'm a visual person and this helps to summarize what I view as important info about my practice
10	Please remove - adds time to data entry and doesn't change practice.
11	I would prefer to see the summary screen once only when i start to enter data, not have to bypass it each time i go to "encounter"
14	Seems unnecessary
15	I don't need to see my percentages page after entering each client encounter. Could be used as a summary page of day/week/month. Easy to read and understand.
16	the summary needs only to come up when I have completed all entries, not after every pt encounter, takes too much time
17	The new summary screen added lag time to inputting stats, and made the process cumbersome.

**How could you be motivated to enter accurate, complete and timely data into the reporting system?  
Did the summary screen help?**

Respondent	Comment
2	if it was a shorter process to chart each patient. It seems to take lots of time to enter each specific patient and what they are here for.
5	I always have entered my data on the same day. Summary screen just makes me anxious (it makes me want to advocate for admin support to have reminder calls).
6	I have usually recorded data on the same day. The new screens seems to discourage that.
7	I personally did not see any difference
8	It helped slightly. I find I am now entering stats every 7-9 days instead of every 9-14 days.
9	I'm not sure it provided extra motivation; I'm a pretty organized person so have always wanted to keep on top of doing stats.
10	Monetary rewards. No, the summary screen did not help - sorry. :(
11	at first it helped somewhat; now i again rely on my own motivation to keep up to date, which ebbs and flows with the demands of my schedule.
13	A bit. I know reporting is necessary but it is also very time consuming. Ultimately, it would be magical if it were integrated right into PSS!
14	Make it easier and simpler, less time consuming.
15	I was already entering data on daily basis, but I do feel it could act as a motivator to those who have not in recent past.
16	I know I need to enter my work into [the reporting tool] but I am not particularly motivated to do so, not sure what would motivate me me.
17	It was nice to see incentives on the screen of reaching goals and receiving badges, but other incentives would likely help motivate