7-25-2022 6:00 PM

Developing artificial intelligence and machine learning to support primary care research and practice

Jacqueline K. Kueper, *The University of Western Ontario*

Supervisor: Lizotte, Daniel J, *The University of Western Ontario*

A thesis submitted in partial fulfillment of the requirements for the Doctor of Philosophy degree in Epidemiology and Biostatistics/Computer Science

Abstract

This thesis was motivated by the potential to use "everyday data", especially that collected in electronic health records (EHRs) as part of healthcare delivery, to improve primary care for clients facing complex clinical and/or social situations. Artificial intelligence (AI) techniques can identify patterns or make predictions with these data, producing information to learn about and inform care delivery. Our first objective was to understand and critique the body of literature on AI and primary care. This was achieved through a scoping review wherein we found the field was at an early stage of maturity, primarily focused on clinical decision support for chronic conditions in high-income countries, with low levels of primary care involvement and model evaluation in real-world settings.

Our second objective was to demonstrate how AI methods can be applied to problems in descriptive epidemiology. To achieve this, we collaborated with the Alliance for Healthier Communities, which provides team-based primary health care through Community Health Centres (CHCs) across Ontario to clients who experience barriers to regular care. We described sociodemographic, clinical, and healthcare use characteristics of their adult primary care population using EHR data from 2009-2019. We used both simple statistical and unsupervised learning techniques, applied with an epidemiological lens. In addition to substantive findings, we identified potential avenues for future learning initiatives, including the development of decision support tools, and methodological considerations therein.

Our third objective was to advance interpretable AI methodology that is well-suited

for heterogeneous data, and is applicable in clinical epidemiology as well as other settings. To achieve this, we developed a new hybrid feature- and similarity-based model for supervised learning. There are two versions, fit by convex optimization with a sparsity-inducing penalty on the kernel (similarity) portion of the model. We compared our hybrid models with solely feature- and similarity-based approaches using synthetic data and using CHC data to predict future loneliness or social isolation. We also proposed a new strategy for kernel construction with indicator-coded data.

Altogether, this thesis progressed AI for primary care in general and for a particular health care organization, while making research contributions to epidemiology and to computer science.

Key Words

Artificial Intelligence, Machine learning, Clinical Epidemiology, Descriptive Epidemiology, Primary Care, Primary Health Care, Decision Support, Learning Health Systems

Lay Summary

This thesis was motivated by the potential to use "everyday data", which is data generated through activities outside formal research settings, to improve primary care for clients facing complex clinical and/or social situations. Artificial intelligence (AI) and its subfield machine learning include techniques that can analyze these data and provide information to help guide care delivery, such as personalized treatment recommendations or risk estimates. In our first study we summarized the state of AI and primary care research, finding the field was at an early stage of maturity with knowledge gaps for how to best develop, implement, and evaluate AI for primary care.

Our second study was done in collaboration with the Alliance for Healthier Communities, which provides team-based primary health care through Community Health Centres (CHCs) across Ontario to clients who otherwise experience barriers to regular care. We performed a large-scale description of sociodemographic, clinical, and healthcare characteristics of their adult primary care clients from 2009 through 2019 to learn about this population and areas where AI and decision support tools may be useful. We additionally identified methodological considerations for AI to work well in primary care settings. To accomplish this we used both simple statistical techniques traditionally used in descriptive epidemiology and techniques from machine learning that can capture more complex patterns in the data. Our approach can be followed to improve population-level descriptions in other settings as well.

In our third study we developed new machine learning methods for analyzing large,

diverse datasets, such as electronic health records from CHCs. We combined two existing techniques, feature and kernel learning, into a single hybrid model. We demonstrated how to interpret our models and use them for prediction and for epidemiological studies, using synthetic data and in a case study to predict social isolation and loneliness for the Alliance population. We also proposed a new way to capture similarity between clients, for use in the kernel part of our model, in terms of deviations from population-level expectations.

Altogether this thesis advanced AI for primary care while making methodological contributions to the fields of epidemiology and computer science.

Combined PhD in Epidemiology and Computer Science

This thesis is for a combined doctoral degree in Epidemiology & Biostatistics (Epidemiology focus) and Computer Science. It is the first of its kind at Western University, made possible through an agreement between the Department of Epidemiology and Biostatistics, the Department of Computer Science, and the School of Graduate and Postdoctoral Studies.

Successful completion of the combined PhD requires satisfying program requirements of the two Western doctoral programs simultaneously within a single program structure. Students in a combined PhD program receive training and are assessed to ensure proficiency in each discipline separately (e.g., courses, comprehensive exams), and then bring knowledge and skills from the two disciplines together to engage in a single, interdisciplinary doctoral research project (this thesis). After successful defense of the PhD thesis against the standards of both disciplines, Western University will award the student a degree listing both programs.

As noted in the joint expectations for a combined PhD in Epidemiology and Computer Science, signed in memorandum in April 2020 (Appendix B), the present thesis is expected to contain content equivalent to at least three research papers and make contributions to both the fields of Epidemiology & Biostatistics and Computer Science.

Co-Authorship Statement

Jacqueline Kueper is the primary author and drafted all chapters of this thesis. Kueper's supervisor, Dr. Dan Lizotte, and advisory committee, Drs. Jennifer Rayner and Merrick Zwarenstein, participated in the conception, planning, and interpretation of research throughout the dissertation, as well as critical review of the drafted materials. Kueper completed all analyses under the direct supervision of Dr. Dan Lizotte.

Chapter 3 includes a published scoping review for which Dr. Amanda Terry is a co-author. She participated in screening and data extraction with Kueper and Dr. Lizotte, as well as interpretation of results and review of the manuscript in collaboration with all co-authors.

Chapters 4 and 5 rely on de-identified electronic health record data made available by the Alliance for Healthier Communities. Dr. Rayner is their Director of Research and Evaluation and facilitated collaboration with the Alliance, as well as provided guidance towards research questions and analyses to ensure relevance of the work to the Alliance.

Jacqueline Kueper is the first author on all manuscripts pertaining to this thesis; her contributions towards each are over 85%.

Acknowledgements

I would like to acknowledge the following people and institutions for their support towards my doctoral studies. Thank you:

To my supervisor, Dr. Dan Lizotte, for helping to identify and nourish my interest in methods research; I learned so much from you and feel incredibly fortunate to have ended up with skills in an area I love and believe in. Thank you for taking the risk to supervise me on this unpaved journey, and for your patience, support, and guidance along the way.

To my thesis advisory committee, Drs. Jennifer Rayner and Merrick Zwarenstein. Dr. Rayner, you epitomize "real world impact" research, and I am so thankful I had the opportunity to learn from your experience and insights at the Alliance for Healthier Communities. Dr. Zwarenstein, your enthusiasm and ever-present brainstorming about AI and clinical practice has been both informative and inspiring.

To my TUTOR-PHC "family" and especially mentor, Dr. Amanda Terry. Thank you for introducing me to primary health care research and teaching me how to navigate interdisciplinary projects therein.

To those who made the combined PhD program possible. Vice-Provost Linda Miller, for responding to my "cold e-mail" proposal and setting the wheels in motion. Dr. Hanan Lutfiyya, for being a role model and advocate, believing it was not too late for me to learn computer science. To those within the Department of Epidemiology and Biostatistics, Department of Computer Science, and the School of Graduate and Postdoctoral Studies who had open minds and worked hard to make

the specifics of the program a reality.

To the Canadian Institutes of Health Research and INSPIRE-PHC, for supporting me financially.

To my family and friends, especially parents, Maggie and Bernie, brother Eric, and Mark and Anne, for your love and support. I feel so fortunate to have you all in my corner. Jason Black, for many pre-pandemic late-nights at Kresge, and for your friendship and enthusiastic encouragement. Jordan Edwards, for inspiring collaboration and hard work in both research and biking during those early days.

And to my partner, Alyssa Gagne, for loving and supporting me during highs and lows, encouraging personal growth, and making sure I occasionally got out of both the apartment and my head (oftentimes on a bike). I am excited to see what adventures are next.

Contents

Abstract ii

Lay Summary iv

Combined PhD in Epidemiology and Computer Science vi

Co-Authorship Statement vii

Acknowledgements viii

1 Introduction 1

1.1 Motivation 1

1.2 Objectives 3

1.2.1 Research contributions 4

1.3 Thesis Organization 5

2 Background 6

2.1 Epidemiology 6

2.1.1 Clinical epidemiology 7

2.1.2 Social epidemiology 7

2.1.3 Health services research 8

2.2 Computer Science 9

2.2.1 Artificial intelligence 9

2.2.2 Machine learning 10

2.2.3 Explainable artificial intelligence 11
2.2.4 Model 12
2.3 Primary Health Care 13
2.3.1 Electronic health records 14
2.4 Learning Health System 15

3 Artificial Intelligence and Primary Care Research: A Scoping Review 16

3.1 Technical Background 16
3.2 Introduction 17
3.3 Methods 18
3.3.1 Search strategy 19
3.3.2 Study selection 19
3.3.3 Data extraction and synthesis 20
3.4 Results 21
3.4.1 Searches 21
3.4.2 Study purpose 22
3.4.3 Author appointment 23
3.4.4 Primary care function 24
3.4.5 Reported target end user 25
3.4.6 Health condition 26
3.4.7 Geographic location 26
3.4.8 AI subfield 27
3.5 Discussion 27
3.5.1 Key findings 27
3.5.2 Strengths and limitations 28
3.5.3 Future research 29
3.5.4 Conclusions 30

4 Describing a Complex Primary Health Care Population in a Learning Health System to Support Future Decision Support and Artifi-

cial Intelligence Initiatives **31**

4.1 Technical Background 32

4.1.1 Epidemiology 32

4.1.2 Unsupervised machine learning 33

4.2 Introduction 35

4.3 Methods 37

4.3.1 Study population and data source 37

4.3.2 General analysis plan 37

4.3.3 Sociodemographic characteristics 38

4.3.4 Clinical characteristics 38

4.3.5 Healthcare use characteristics 39

4.4 Results 40

4.4.1 Sociodemographic characteristics 41

4.4.2 Clinical characteristics 44

4.4.3 Healthcare use characteristics 48

4.5 Discussion 51

4.5.1 Sociodemographic characteristics 52

4.5.2 Clinical characteristics 53

4.5.3 Healthcare use characteristics 54

4.5.4 Strengths and limitations 55

4.5.5 Conclusions 55

5 Hybrid Feature- and Similarity-Based Models for Prediction and Interpretation on Large-Scale Observational Data **57**

5.1 Technical Background 58

5.1.1 Feature-based learning 58

5.1.2 Similarity-based learning 59

5.1.3 Model selection and performance 62

5.1.4 Standardization 64

5.2 Introduction 64

5.2.1 Generalizable insights about machine learning in the context of healthcare . . . 66
5.3 Related Work . . . 66
5.3.1 Integrating multiple types of data . . . 67
5.3.2 Combining model types . . . 67
5.3.3 Kernel functions for clinical data . . . 68
5.4 Methods . . . 69
5.4.1 The hybrid feature- and similarity-based model . . . 70
5.5 Evaluation . . . 74
5.5.1 Simulation study . . . 74
5.5.2 Methods for clinical case study . . . 77
5.5.3 Results for clinical case study . . . 82
5.6 Discussion . . . 84
5.6.1 Hybrid model methodology . . . 84
5.6.2 Clinical case study . . . 85
5.6.3 Limitations . . . 87
5.7 Conclusion . . . 88

6 Discussion **89**
6.1 Summary of Major Themes . . . 90
6.1.1 Artificial intelligence for primary care is at an early stage of maturity, but progressing . . . 90
6.1.2 Methodology needs to account for primary care complexity . . . 91
6.1.3 Epidemiology and computer science are complementary fields . . . 97
6.2 Avenues for Further Study . . . 99
6.2.1 Artificial intelligence for primary care . . . 99
6.2.2 The Alliance for Healthier Communities . . . 100
6.2.3 Epidemiology and computer science methods research . . . 102
6.3 Conclusions . . . 103
6.4 References . . . 104

APPENDICES **127**

A List of Abbreviations **128**

B Combined PhD Agreement **129**

C Objective 1 Extended Information **136**

C.1 PRISMA-ScR Checklist 137

C.2 Search Strategies 139

C.3 Additional Methods and Results 146

C.4 References 157

D Objective 2 Extended Information **185**

D.1 RECORD Checklist 186

D.2 Extra Figures 193

D.3 Extra Tables 202

D.4 Extra Technical Details 214

E Objective 3 Extended Information **215**

E.1 Illustrative Example: Hybrid Feature- and Similarity- Based Model . 216

E.2 Hybrid Model Code 217

E.3 Simulation Study Details 220

E.3.1 RBF kernel sigma selection 220

E.3.2 Selected hyperparameters and model coefficients 221

E.4 Clinical Case Study Details 223

E.4.1 Cohort overview 223

E.4.2 Application 1: prediction 226

E.4.3 Application 2: inference/interpretation 227

F Frequent Visitor Analyses **231**

F.1 Methods 232

F.1.1 All-time frequent visitor characteristic comparison 232

F.1.2 Risk factors for a period of frequent visits 232

F.2 Results 233
F.2.1 All-time frequent visitor characteristic comparison 233
F.2.2 Risk factors for a period of frequent visits 236
F.3 Discussion 238
F.4 References 239

G Curriculum Vitae **240**

List of Tables

3.1 Appointments of study authors. 24

4.1 Sociodemographic characteristics. 41
4.2 Eleven-year period prevalence. 44

5.1 Jaccard similarity score example. 60
5.2 Jaccard similarity score example with common code. 61
5.3 AUROCs and coefficients for interpretation example. 73
5.4 Synthetic data study results. 77
5.5 Clinical case sudy predictive performance results. 82

6.1 Machine learning or decision support tool problem selection challenges due to the breadth of team-based primary care practice. 94
6.2 Methodological challenges for prediction of future events related to providing care across the life course. 95

C.1 Search terms for health sciences databases. 142
C.2 Search terms for computer science databases. 144
C.3 Data extraction field characterizations. 147
C.4 Complete author appointment counts. 152
C.5 Detailed breakdown of location. 155

D.1 Characteristic variable definitions. 203
D.2 Sociodemographic characteristics sub-strata. 207
D.3 Health care use characteristics. 210

D.4 Provider type counts. 211
D.5 Time series clustering of care access frequency. 212

E.1 Illustrative example feature coefficients 216
E.2 Synthetic data scenario 1: selected hyperparameters 221
E.3 Synthetic data scenario 1: model interpretation 221
E.4 Synthetic data scenario 2: selected hyperparameters 222
E.5 Synthetic data scenario 2: model interpretation 222
E.6 Synthetic data scenario 3: selected hyperparameters 222
E.7 Synthetic data scenario 3: model interpretation 222
E.8 Clinical case study baseline features. 225
E.9 Clinical case study selected hyperparameters. 226
E.10 Feature coefficients for models re-trained on all data 227
E.11 Client characteristics stratified by kernel coefficient 228

F.1 Frequent visitor univariate characteristic comparisons. 235
F.2 All-time frequent visitor univariate statistical comparisons. 236
F.3 Frequent visitor risk factor analyses. 237

List of Figures

3.1 PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) flow diagram. 22
3.2 Overall purpose of studies. 23
3.3 Primary care functions to be supported with artificial intelligence. . . 25
3.4 Health conditions studied. 26

4.1 Non-negative matrix factorization example. 34
4.2 Example observation-based period prevalence and cumulative incidence plots. 46
4.3 Condition co-occurrence patterns: heatmap representing the results of the Ising model. 47
4.4 Common care provider teams. 49
4.5 Care frequency clusters. 51

5.1 Data generating mechanisms used to contrast sequential and simultaneous hybrid model optimization. 72

C.1 Development of search strategies. 139
C.2 Detailed breakdown of primary care functions. 152
C.3 Author reported end user total counts. 153
C.4 Detailed breakdown of author reported intended end user combinations by study. 153
C.5 Detailed breakdown of health conditions. 154

C.6 Most frequent locations of data source or intended implementation with per capita rates. 154
C.7 Most frequent subfields of artificial intelligence with median year of publication. 156
C.8 Detailed breakdown of artificial intelligence subfields with median year of publication. 157

D.1 Cohort size by calendar- and observation-based time. 193
D.2 Observation-based period prevalence. 198
D.3 Cumulative incidence plots by days of observation since the first recorded event. 202

E.1 Heatmap demonstrating similarity of RBF kernel with various σ on a random sample. 220
E.2 Clinical case study cohort flow diagram 224

F.1 Frequent visitor cut-off for all-time access. 234
F.2 Frequent visitor cut-off for quarter-year. 237

Chapter 1

Introduction

1.1 Motivation

Increasing amounts and types of "everyday data" are being collected as a by-product of activities happening every second: combined with advancements in computational resources and methods to analyze these data, there are unprecedented opportunities to inform and improve everyday activities or associated human decisions. Of particular interest is "everyday data" from healthcare, such as data collected in electronic health records (EHRs) as a result of clinical contact, and the analysis of these data to improve care and by extension population health. This could happen at any scale. An entire healthcare system, specific organization, or single clinic can harness their care-derived data to better understand health related needs and characteristics of the population they serve, and to inform or develop tools that will support and improve care delivery for that population.[1–4]

The applied setting of interest throughout this thesis is primary care, where the potential for benefit under the above paradigm is particularly motivating for clients with complex medical and social conditions. In Canada, primary care is first-contact care provided in a community setting over the life course, serving as the foundation and entry point to the rest of the healthcare system.[5,6] Hence, primary care providers are

responsible for tasks related to primary prevention and screening, as well as treatment of acute conditions and management of chronic conditions to prevent or slow progression to health states where secondary or tertiary care is needed.[5,7] For clients experiencing complex health challenges, such as those with two or more chronic conditions (multimorbidity) or who experience social and structural barriers to health, primary care plays an important role in healthcare access and in coordinating multiple, often competing, care regimes.[7–10] Amid this complexity there is a lack of evidence around prognosis, diagnosis, and treatment, especially that is appropriately tailored to a client's social determinants of health.[8,11–26]

Methods from artificial intelligence (AI) and machine learning, a subfield of AI, may be applied to health data and fill gaps in complex client care, such as to learn patterns and make personalized predictions to help support care decisions. AI-driven decision support tools that are appropriately developed, rigorously evaluated, and used in informed ways have potential to improve care; however, there are also risks of making things worse, such as through poor-quality tools that reduce client safety or are biased against minority populations.[27–36] Understanding possibilities surrounding AI for healthcare and methodological subtleties underpinning specific scenarios can help shift the balance away from harm and towards more potential benefits. Historically, advancements in AI for health have focused on specialty and acute care settings moreso than primary care[37,38]; there are unanswered questions about how a technology that is expected to revolutionize healthcare will impact the foundation of these systems, and a need for technical research that is tailored to the unique attributes of primary care. Both epidemiology and computer science are needed to support progress.

Understanding and trust of AI models is key for primary care as well as other sectors, like law or finance, where model outputs are intended to support human decision-making.[39–41] Particularly since the advent of deep learning, the black box nature of some AI has come under scrutiny, especially for settings where "success" extends beyond technical performance to also require equity, safety, and human action.[39,40,42,43]

Although there have been advances in explanation, these are not yet sufficient. There is a need for advancements in machine learning methods that allow for incorporating expert prior knowledge and constraints in an interpretable way.[39–41] For tasks where causality is important, such as treatment effect estimation, an inherently interpretable model can be informed or evaluated with epidemiological and subject matter expertise in ways not possible with post hoc explanation of black box models.

This thesis lays foundations for the field of AI for primary care in general, and for future AI-related work at a particular health care organization that focuses on care for clients experiencing complex health challenges across Ontario. It then proposes new explainable machine learning methodology to target gaps in primary care and other high-risk decision making types of settings.

1.2 Objectives

We addressed three **objectives** motivated by the need to better understand how AI can be developed to support primary care:

1) To identify and summarize existing research that involves AI and primary care.
2) To demonstrate how a combination of simple statistical and more complex unsupervised learning techniques can be used to describe sociodemographic, clinical, and healthcare use characteristics of a complex primary care population for the purpose of supporting future initiatives, including the development of decision support tools.
3) To develop and evaluate an interpretable hybrid feature- and similarity-based model for supervised learning that takes advantage of rich but heterogeneous observational data sources and can be used for prediction and for investigation of causal relationships.

Objectives 2 and 3 were achieved in collaboration with the Alliance for Healthier Communities, which provides team-based primary health care through Community Health Centres (CHC) across Ontario to clients who otherwise experience barriers

to receiving regular care.[44] The population served through CHCs is heterogeneous and care decisions are often complex and challenging. By vote of their executive leaders in October 2020, the Alliance committed to using their data to improve care by adopting a learning health system (LHS) model,[45,46] making them one of the first documented primary care LHSs in North America.[4] Combined with their collection of rich sociodemographic data and motivation for health equity and social justice, the Alliance is a unique primary health care system in Canada with large potential to take advantage of their EHRs and various data analysis methods, including AI. An LHS is formally defined in Chapter 2.

1.2.1 Research contributions

The present thesis is for a combined degree in epidemiology and computer science. For clarity, the research objectives can be re-organized in terms of contributions to the two disciplines of interest:

To computer science and more specifically the subfield of machine learning, we contribute 1. hybrid feature- and similarity-based supervised learning methods (Chapter 5), 2. a new framework for thinking about similarity in kernel functions (Chapter 5), and 3. demonstration of how techniques from epidemiology can be used to inform machine learning projects, including the development of decision support tools (Chapters 3-6).

To epidemiology and more specifically the subfield of clinical epidemiology, we contribute 4. the first comprehensive review of AI for primary care research (Chapter 3), 5. the first large-scale description of adult primary care clients served by CHCs across Ontario (Chapter 4), and 6. demonstration of how techniques from computer science can be used to aid in population-level descriptive studies and in investigation of causal relationships (Chapters 4-6).

1.3 Thesis Organization

This thesis is in an integrated-article format. **Chapter 2** provides general background on key concepts and terms that are needed to understand the body of work. **Chapters 3, 4, and 5** include integrated articles that address Objectives 1, 2, and 3, respectively. Each of these chapters begins with an extra technical background section that defines discipline-specific terms and how they are used within the associated article. The general background in **Chapter 2** and technical background sections in **Chapters 3-5** are provided to increase accessibility of the work to readers from different disciplines; some or all of these background sections can be skipped by readers already familiar with the contents. **Chapter 6** concludes the thesis with an overarching discussion. There are several **appendices** that contain supporting materials, such as extra information on methods and results. Due to the nature of an integrated article format, there is some repetition between introductory sections.

Chapter 2

Background

This thesis relies on and makes contributions to both epidemiology and computer science, and takes place within the applied health setting of primary care. To support readership from multiple disciplines, brief explanations of key terms and how they are used throughout the body of work are provided below. While this Chapter provides explanations at a high-level to support conceptual understanding of the research, additional technical background is presented at the beginning of each integrated article chapter to review terms that are not used universally between all fields and are needed to fully understand each specific research study.

2.1 Epidemiology

Greenland and Rothman (2008) define epidemiology as "the study of the distribution of health-related states and events in populations. With this definition we intend to capture not only disease and illness, but physiologic states such as blood pressure, psychologic measures such as depression score, and positive outcomes such as disease immunity."[47] Two major types of epidemiology are descriptive and analytic. **Descriptive epidemiology** provides measures to understand health conditions within a population, such as the prevalence of hypertension. **Analytic epidemiology** focuses on identifying contributing or protective factors for a health state, for example

to obtain a valid and precise estimate of the effect of a potential treatment for slowing the progression of hypertension. Although much of analytic epidemiology is causal, there are some questions that may not require explicit causal relationships to be useful, such as exploratory risk factor analyses.

Our research included and made contributions towards both descriptive and analytic epidemiology. Epidemiology can also be broken down by application field; the two most relevant to our research are clinical and social epidemiology.

We describe core concepts in epidemiology, including cohorts, prevalence, and incidence, in the technical background for Chapter 4. In Chapter 5 we additionally introduce standardization as a way to assess whether outcome rates in a population under study match expectations based on a reference population.

2.1.1 Clinical epidemiology

Baron (2001) defines clinical epidemiology as "the application of epidemiologic and biostatistic techniques to clinical problems. In contrast to chronic disease epidemiology, which focuses on the discovery of the determinants of disease on a population level, clinical epidemiology aims to help clinicians conduct the daily work of caring for individual patients."[48] This thesis was motivated by the potential to use the collection and analysis of clinical data to provide information to support further care decisions.

2.1.2 Social epidemiology

Kaufman (2008) defines social epidemiology as "the study of relations between social factors and disease in populations... social epidemiology is characterized by explicit inclusion of social, economic, or cultural quantities in the exposure definition or the analytic model, or by explicit reference to social science theory in the interpretation."[49]

Our research was informed by social epidemiology and motivated by the potential for artificial intelligence (AI) to help identify the best care decisions in the context of an individual's sociodemographic and clinical characteristics. It is crucial to understand

how to incorporate factors such as social determinants of health into AI methods both because they are expected to improve performance and generalizability and because they are expected to reduce the risk of bias.

2.1.2.1 Social determinants of health

The Canadian Public Health Association defines social determinants of health as, "the social and economic factors that influence people's health. These are apparent in the living and working conditions that people experience every day. The social determinants of health influence health in many positive and negative ways".[50]

Health services are a key determinant of health. Other examples include income and income distribution, education, unemployment and job security, employment and working conditions, early childhood development, food insecurity, housing, social exclusion, social safety network, Indigenous status, gender, race, and disability. Social determinants of health are an explicit component of the Alliance for Healthier Communities care model,[51] and information representing several determinants is collected in their electronic health records (EHRs). Provision of health and social services can counteract some of the negative social determinants of health and promote health equity.

2.1.3 Health services research

The Canadian Institutes of Health Research (2019) defines health services research as "research with the goal of improving the efficiency and effectiveness of health professionals and the health care system, through changes to practice and policy. Health services research is a multidisciplinary field of scientific investigation that studies how social factors, financing systems, organizational structures and processes, health technologies, and personal behaviours affect access to health care, the quality and cost of health care, and, ultimately, Canadians' health and well-being."[52] This thesis includes analyses that measure and explore health service access at the Alliance for Healthier Communities, to identify general patterns and those that differ across

client subpopulations.

2.2 Computer Science

The ACM Task Force on the Core of Computer Science define "the discipline of computing [as] the systematic study of algorithmic processes that describe and transform information, their theory, analysis, design, efficiency, implementation, and application."[53] The focus area of computer science in this thesis is the use of AI to process digital data.

2.2.1 Artificial intelligence

AI is a rapidly growing area with no single, well-defined definition. The following brief history provides contextual background and motivations for AI in general; more concrete descriptions of the subfields used in this thesis are below[54]:

> *From its inception in the 1950s, AI was primarily concerned with processes by which computers might achieve 'intelligence' comparable to that of humans, and how we might recognize such intelligence.[55] Turing's (1950) seminal paper, 'Computing Machinery and Intelligence', was concerned more with the latter, but the work sparked a rich diversity of research activities.[55,56] The field of AI now encompasses a wide variety of methodology, much of which falls into two broad categories: rule-centred and data-centred. Rule-centred methods came from the study of logical reasoning, and are intended to capture intelligence by explicitly writing down the rules that govern it and then deploying that intelligence to carry out different tasks.[55] Data-centric methods like machine learning have focused more on learning to perform specific tasks using previously collected data rather than explicitly provided rules.[55]*

Readers interested in learning more about the general types of "intelligent tasks" that AI methods are able to perform can read our primer on AI for primary care.[57] This

thesis most heavily deals with machine learning, which is a subfield of AI.

2.2.2 Machine learning

Machine learning can be broken down into three (non mutually-exclusive) categories.

2.2.2.1 1) Supervised machine learning

Supervised machine learning models learn to associate labels with observations. In the context of health, the label is often an outcome, e.g., presence of hypertension, and the observations are often client characteristics, e.g., EHR history. Labels may be numeric (in regression problems) or categorical (in classification problems). Supervised learning uses existing labeled data, which contains a collection of observations with "true" labels, to learn how to predict the label for new, previously unseen observations. Supervised learning methods are able to capture complex (e.g., non-linear, additive) relationships between inputs and outputs, which is a strength for settings such as healthcare where a multitude of different factors contribute in different, potentially interacting ways to health states and outcomes. Example supervised machine learning techniques include Support Vector Machines, K-Nearest Neighbours, Naïve Bayes Classifier, and Random Forest Decision Trees.[55]

Chapter 5 presents advances to supervised machine learning methodology. This research included feature-based and similarity-based approaches, with kernel methods being the similarity-based approach of focus. An introduction to these concepts is provided in the Chapter 5 technical background section alongside an overview of model selection and performance assessment.

2.2.2.2 2) Unsupervised machine learning

Unsupervised machine learning models learn patterns from unlabeled data such as an EHR database with no predefined outcome. Common unsupervised machine learning techniques include clustering, topic modelling, and association mining to identify observations that tend to occur together.[55] Unsupervised machine learning techniques

are closely related to and overlap with statistical modelling approaches used in epidemiology, including latent class and latent factor models.[58,59]

Chapter 4 applies unsupervised machine learning techniques with EHR data. The technical background section includes an overview of unsupervised methods with the three techniques we applied: Ising models, non-negative matrix factorization, and K-medoids time series clustering.

2.2.2.3 3) Reinforcement learning

Reinforcement learning involves learning a series of actions to optimize rewards or punishments.[55] In the context of health care, these rewards or punishments could be benefits or side-effects of medications, and the goal could be to learn an optimal treatment regime using longitudinal data whereby different types or doses of treatment are administered. Reinforcement learning for tertiary health care settings (e.g., intensive care unit) is a rapidly developing area and the techniques are a plausible extension of the hybrid model research in Chapter 5; however, more foundational methodological work in supervised machine learning for primary care data is needed first. Thus, reinforcement learning is an opportunity for future work.

2.2.3 Explainable artificial intelligence

Machine learning has been criticized as being "black box", which may be problematic especially in the context of decision making where a predictive model is intended to augment decision making of a clinician and client. Explainable AI (XAI) emerged to combat this, such as by providing information alongside an output (e.g., prediction) that facilitates understanding about how the model is functioning.[40,60–62] There is an active research community around XAI, including for health care specifically, and for which published reviews and guidelines for researchers exist.[39,60–62] We developed our proposed machine learning methods in Chapter 5 under an XAI paradigm such that there is an opportunity for explainability.

2.2.4 Model

The term "model" has widely different definitions used both within and across disciplines. We use model in the context of data analysis: a mathematical or statistical model that relates inputs (i.e. observed data also referred to as features or independent variables) to outputs (e.g., probability of a pre-specified outcome).

2.2.4.1 Risk prediction models

Risk prediction models can estimate a client's risk for a future outcome (e.g., disease development) based on the client's observed characteristics.[63] Analogous terms include "prognostic predictive models" and "risk engines". Example clinical action arising from a risk prediction model for a chronic disease would be to order screening tests for a client in response to high predicted risk of the disease, or to not order screening tests for a client in response to low predicted risk of the disease. Chapter 5 develops and tests a series of risk prediction models.

2.2.4.2 Treatment effect models

Throughout this thesis "treatment" is interpreted broadly to mean any action or intervention taken to try and alter the future state of a client, e.g., medications, social interventions, and behavioural interventions. Whereas the focus for a risk prediction model is the absolute risk estimate, the focus for a treatment effect model is the expected benefit or harm (which could be a change in estimated risk) associated with starting or changing a treatment. Treatment effect models explicitly aim to be causal, meaning that changing the treatment will result in a change in the outcome[64,65]; misinterpreting a non-causal estimate as causal can lead to harm in high-risks decision making settings such as healthcare. We suspect that under an XAI framework, end users may interpret a risk prediction model in a causal way as an "upstream treatment effect model", e.g., "hypertension is the largest contributor to this patient's diabetes risk, therefore I want to intervene on blood pressure to reduce diabetes risk". This could be problematic if this is a biased (untrue) relationship between hypertension

and the outcome and it obscures identification of the next best care decision. We further discuss these ideas in Chapter 5 with respect to the proposed hybrid models.

2.3 Primary Health Care

Primary health care is "the level of a health service system that provides entry into the system for all new needs and problems, provides person-focused (as opposed to disease-oriented) care over time, provides care for all but very uncommon or unusual conditions, and coordinates or integrates care provided elsewhere or by others.[5(pp8-9)] **Primary care** is a subfield of primary health care that typically focuses on "family medicine" and includes family physician, nurse, and nurse practitioner care providers; primary *health* care additionally includes providers such as social workers, dietitians, and physiotherapists.[66] Chapters 4 and 5 include research with a primary care cohort situated within the context of a primary health care setting, such that eligible clients must have received primary care, but their access to broader primary health care services and providers was considered in analyses. The use of both primary care and primary health care terminology throughout this thesis is intentional and provides distinction about the impact or relevance of the associated information.

Clinical problems can roughly be divided into those that are simple, complicated, and complex.[67] Simple problems, such as recording a blood pressure reading, are those that may require technique and terminology refinement but once solved can be addressed in a standardized fashion to produce good results each time.[67,68] The main opportunities for technology intervention will include automation that may not require AI. Complicated problems may include subsets of simple problems but are more than a collection of simple problems that can be independently solved; they are challenging due to scale and the need for coordination or specialized expertise, but formulae to solve them and achieve high certainty of outcomes is possible.[68] Complicated problems in healthcare, such as treatment of advanced cancer, are often addressed by specialty physicians and may be amenable to the development of clinical practice guidelines.[67] Opportunities for technology here include AI, such as to improve

the efficiency or accuracy of identifying the parameters of a complicated problem (e.g., exact diagnosis). While both of the above appear in primary care settings, there are also a large number of complex problems, such as how to care for older adults with multiple chronic conditions and limited monetary resources.[67] These types of situations do not lend themselves to straightforward protocols as the intervention(s) and outcome(s) of interest vary on a person-by-person basis and over time. Complex problems may include simple and complicated aspects, but multiple outcomes may need to be taken into account and solutions may not generalize between clients.[68] Some of the most hopeful applications of AI-based tools in primary care are to provide decision support for complex problems, such as by providing additional information to augment clinical decision making like personalized risk estimates or treatment suggestions that take into account individual client scenarios and preferences.[69,70]

2.3.1 Electronic health records

An EHR is "a secure, integrated collection of a person's encounters with the health care system; it provides a comprehensive digital view of a client's health history".[71] EHRs contain historical data on clients over time, including information about sociodemographic characteristics, diagnoses, care provided, and other health outcomes.

EHRs are designed to support clinical care; research is a secondary purpose and challenges not seen in data collected for research purposes can arise. For example, clients may be observed at irregular time intervals and what is or is not entered into a client's record may be impacted by behavioural or political factors. Advantages include having data from all clients who received care, which may mitigate selection biases, and the opportunity to develop models that work with "everyday data" already present in clinical encounters. Chapters 4 and 5 include research that used EHR data from the Alliance for Healthier Communities.

2.4 Learning Health System

The term "learning health system" (LHS) was defined in 2006 by the Institute of Medicine as, "science, informatics, incentives, and culture are aligned for continuous improvement and innovation, with best practices seamlessly embedded in the delivery process and new knowledge captured as an integral by-product of the delivery experience".[72,73]

A hallmark of LHSs is a commitment to using data to inform or improve care delivery, e.g., through research studies, quality improvement initiatives, or development of decision support tools with EHR data.[1,2,4,74,75] This thesis includes research done in collaboration with one of the first primary health care, LHSs in Canada. An important aspect of this LHS is their commitment to equity and care for complex or historically marginalized subpopulations; analyses done with their data are intended to align with these mandates, such as through careful consideration of social determinants of health. In Chapter 4 we introduce this health care system and the clients they serve in more depth.

Chapter 3

Artificial Intelligence and Primary Care Research: A Scoping Review

Scoping reviews are recognized as an important intellectual contribution in epidemiology and in health research. This Chapter contains the first scoping review on artificial intelligence (AI) and primary care research,[a] which included reviewing thousands of potentially relevant documents to provide the first comprehensive overview and synthesis of the field. It laid a foundation both for the remaining chapters of this thesis and for the field more generally. The study was published in the top primary care journal in North America, the *Annals for Family Medicine*, was presented in several contexts, and was used to inform work at the College of Family Physicians of Canada[76,77] and at the American Board of Family Medicine.[78] The discussion in Chapter 6 will highlight key research developments since the time of this review.

3.1 Technical Background

Scoping review: A scoping review uses a rigorous and systematic search strategy to identify relevant literature on a topic, and then synthesizes or summarizes the located

[a]A version of this chapter has been published: Kueper JK, Terry AL, Zwarenstein M, Lizotte DJ. Artificial intelligence and primary care research: a scoping review. *Annals of Family Medicine.* 2020;18(3):250-258. doi:10.1370/afm.2518

literature to answer a research question about that topic.[79–81] Scoping reviews are related to but distinct from systematic reviews, where studies focusing on a specific research question are collected usually with the goal of conducting a meta-analysis and obtaining a global effect estimate for a specified exposure or treatment.[82] A scoping review typically captures literature addressing a variety of research questions related to a topic-area, and the summary of this literature targets objectives such as to identify the size and type of existing evidence on a topic, to identify any evidence gaps, and/or to inform future research or related initiatives.[79] Scoping reviews serve as a basis to comment on research practices and the field as a whole.

3.2 Introduction

AI research began in the 1950s, and public, professional, and commercial recognition of its potential for adoption in health care settings is growing.[29,83–88] This application includes primary care,[89–91] defined by Barbara Starfield as "The level of a health service system that provides entry into the system for all new needs and problems, provides person-focused (as opposed to disease-oriented) care over time, provides care for all but very uncommon or unusual conditions, and coordinates or integrates care provided elsewhere or by others."(pp8-9)[5] Given the recent surge in uptake of electronic health records (EHRs) and thus availability of data,[92,93] there is potential for AI to benefit both primary care practice and research, especially in light of the breadth of practice and rapidly increasing amounts of information that humans cannot meaningfully condense and comprehend.[5,27–29,31,83,85–91,94–97]

AI's immediate usefulness is not guaranteed, however: EHRs were predicted to transform primary care for the better, but led to unanticipated outcomes and encountered barriers to adoption.[92,98–100] AI could also harm, for example, by exaggerating racial, class, or sex biases if models are built with biased data or used with new populations for whom performance may be poor. Liability, trust, and disrupted workflow are further concerns.[86]

AI initially focused on how computers might achieve humanlike intelligence and how we might recognize this.[55,56] Two approaches emerged, *rule centric* and *data centric*. Rule-centric methods capture intelligence by explicitly writing down rules that govern intelligent decision making, whereas data-centric methods learn specific tasks using previously collected data.[55] Examples of health applications are presented below.

MYCIN was the first rule-based AI system for health care, developed in the 1970s to diagnose blood infections using more than 450 rules derived from experts, textbooks, and case reports.[55,101] Although met with initial enthusiasm, rule-centric methods faltered when faced with increasing complexity. As availability of EHRs increased, AI shifted toward data-centric, machine learning methods designed to automatically capture complex relationships within health data. Machine learning methods are now used in health research to predict diabetes and cancer from health records,[96,102–104] and together with computer vision have been applied to skin cancer diagnosis based on skin lesion images.[105,106] Machine learning and natural language processing methods extract structured information from unstructured text data,[95] which could potentially remove some of the EHR-associated burden from clinicians.[87,107,108]

These examples predominantly come from referral care settings, not from primary care, where the spectrum of illness is wider, and clinicians have fewer diagnostic instruments or tests available. Despite optimism for using AI to benefit primary care, there is no comprehensive review of what contribution AI has made so far, and thus little guidance on how best to proceed with research. To address this gap, our objective was to identify and assess the nature and extent of the body of research involving AI and primary care.

3.3 Methods

We performed a scoping review according to published guidelines whereby a systematic search strategy identifies literature on a topic, data are extracted from relevant documents, and findings are synthesized.[80,81,109] We followed the Preferred Reporting

Items for Systematic Reviews and Meta-Analyses for Scoping Reviews (PRISMA-ScR) Checklist (Appendix C),[79] and registered our protocol with the Open Science Framework (osf.io/w3n2b).

3.3.1 Search strategy

We developed our search strategies iteratively and in collaboration with a medical sciences librarian for health sciences, computer science, and interdisciplinary databases. Strategies included key words and, where possible, subject headings around the concepts of AI and primary care. Terms were identified through searches of the National Library of Medicine MeSH Tree Structures and by discipline experts on our review team. Appendix C.2 contains an overview of the search strategy development process and final strategies for the 11 published or gray literature databases: Medline-OVID, EMBASE, CINAHL, Cochrane Library, Web of Science, Scopus, Institute of Electrical and Electronics Engineers Xplore, Association for Computing Machinery Digital Library, MathSciNet, Association for the Advancement of Artificial Intelligence, and arXiv. Retrieved references were uploaded into Covidence.[110] Where possible, English-language limits were set; to estimate the amount of literature missed, searches were rerun for a subset of the databases (Medline-OVID, CINAHL, Web of Science) with language limits reset to accept all non-English languages. Each search retrieved fewer than 10 documents. We used Covidence[110] to remove duplicate results and facilitate the screening process.

3.3.2 Study selection

3.3.2.1 Title and abstract screening

For preliminary screening, two reviewers (JKK, DJL) independently rated document titles and abstracts as to whether they met our eligibility criteria: (1) reported on research, (2) mentioned or alluded to AI, and (3) mentioned primary care data source, setting, or personnel. We pilot-tested the first 25 and next 100 documents, discussing disagreements to ensure mutual understanding of the eligibility criteria and capture of

relevant literature. A third reviewer (ALT) resolved remaining initial disagreements. If two reviewers rated a document as meeting the above criteria, the document progressed to full-text screening. A large number of documents on computerized cognitive behavioral therapy (37 documents) were excluded because underlying methods were often unclear and reviews on these systems already exist.[111–115]

3.3.2.2 Full-text screening

For our full-text screening, two reviewers (JKK, DJL) independently reviewed the full text of each document for the following eligibility criteria: (1) was a research study, (2) developed or used AI (Appendix Table C.3 contains subfield definitions), (3) used primary care data and/or study was conducted in a primary care setting and/or explicitly mentioned study applicability to primary care. Documents were excluded if they were narratives or editorials, did not apply to primary care, or were not accessible in English language full text. As for title and abstract screening, we performed pilot-testing and refined the eligibility criteria. Disagreements were resolved by discussion until consensus was reached.

A notable challenge arose from authors' use of terminology that overlaps with AI when the methods used are not considered AI; we excluded these studies. For example, one study referred to simple string matching as natural language processing.[116] We also excluded 34 studies because there was insufficient information to determine whether AI was involved, even after consulting references cited in methods.

3.3.3 Data extraction and synthesis

We developed the data extraction sheet iteratively to ensure relevant and consistent information capture, performing pilot-testing and revisions for 3 and then 5 randomly selected articles.[106,117–123] Remaining documents were split alphabetically and extracted independently (100 by ALT, 50 by DJL, 250 by JKK). We extracted the following information: publication details, study purpose(s), author appointment(s), primary care function(s), author-intended target end user(s), target health condi-

tion(s), location of data source(s) (if any), AI subfield(s), the reviewer who performed extraction, and any reviewer notes. We agreed on definitions for each data extraction field (Appendix Table C.3). For fields except publication details, author appointments, and additional notes, we predefined categories based on the pilot testing and on content knowledge; studies could belong to multiple categories. An "other" category captured specifics of studies that did not fit into a predefined category, and an "unknown" category was used if not enough information was provided for category selection. We summarized results as categorical variables for seven data extraction fields and performed selected cross-tabulations.

3.4 Results

3.4.1 Searches

We retrieved 5,515 nonduplicate documents for title and abstract screening; 727 met the eligibility criteria for full-text screening and 405 met the final criteria as shown in Figure 3.1. Appendix C.4 contains a list of the 405 references. The AI and primary care study with the earliest date of publication, 1986, developed a supervised machine learning method to support abdominal pain diagnoses.[124] Studies are summarized below according to the seven key data extraction categories mentioned above.

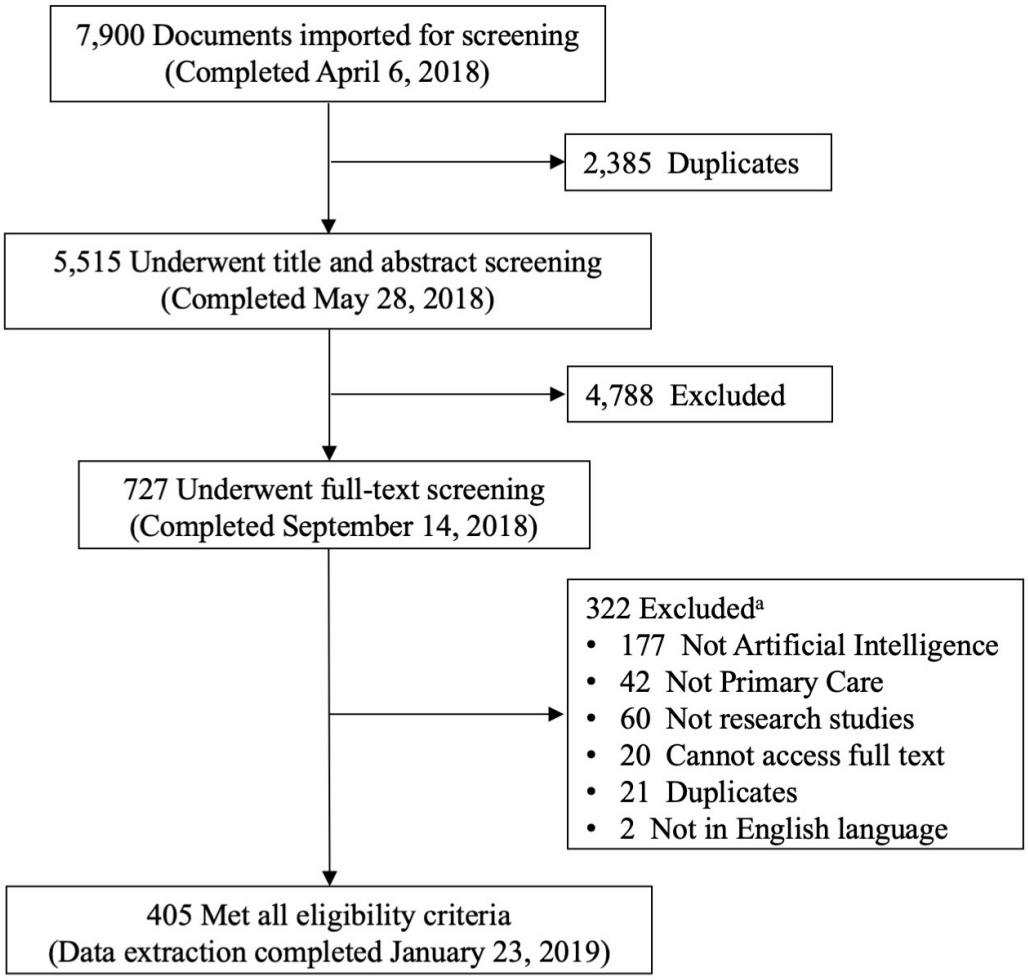


Figure 3.1: PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) flow diagram.

a: "Not primary care" use as exclusion when multiple criteria applied

3.4.2 Study purpose

The majority of studies (270 studies, 66.7%) developed new or adapted existing AI methods using secondary data. The second most common study purpose (86 studies, 21.2%) was analyzing data using AI techniques, such as eliciting patterns from health data to facilitate research. Few (28 studies, 6.9%) evaluated AI application in a real-world setting.

Some series of studies reported on multiple stages of a project, from AI development to pilot-testing; these projects included intended end users located in a primary care

setting.[125–132] A small minority of studies (21 studies, 5.2%) had multiple purposes. Figure 3.2 presents all combinations.

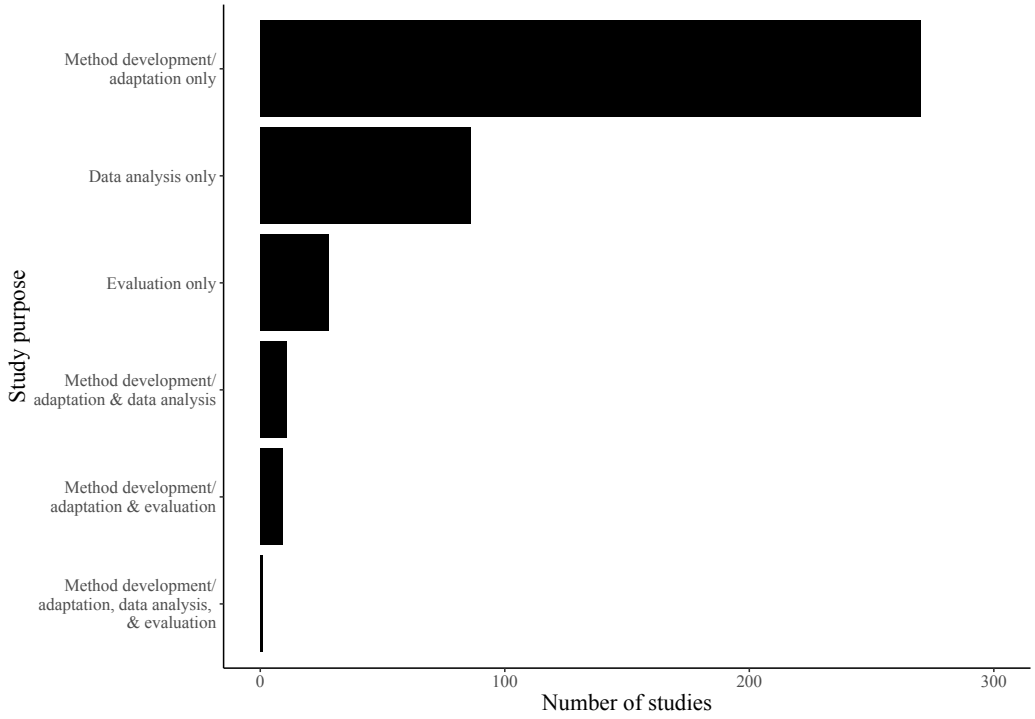


Figure 3.2: Overall purpose of studies.

3.4.3 Author appointment

We categorized author appointments into 4 categories: (1) technology, engineering, and math (TEM) discipline, meaning an author appointed in a department of mathematics, engineering, computer science, informatics, and/or statistics; (2) primary care discipline, meaning an author appointed in a department of family medicine, primary care, community health, and/or other analogous term; (3) nursing discipline; and (4) other. Authors were predominantly from TEM disciplines with 214 studies (52.8%) having at least one author with a TEM appointment compared with just 57 studies (14.1%) having at least one author with a primary care appointment. Twenty-three studies (5.7%) had a primary care–appointed author listed first and 27 (6.7%) had one listed last. These patterns remained when unspecified or general medical appointments (i.e. nonspecialist) were counted as primary care appointments. Four studies

had authors with nursing appointments. Cross-tabulations between study purpose and author appointment categories did not suggest that author appointment types differed by study purpose. Table 3.1 presents a summary of the body of literature broken into primary care and TEM author disciplines; Appendix Table C.4 breaks down author appointments into 16 categories.

Table 3.1: Appointments of study authors.

Author Appointment Category	n (%) of Studies
Primary care and TEM	27 (6.67)
Primary care and no TEM	30 (7.41)
TEM and no primary care	187 (46.17)
Neither TEM nor primary care	161 (39.75)

Note: To be included in a row count, a study must have had at least one author with an appointment in the category or categories indicated. *Legend*: TEM = technology, engineering, and math.

3.4.4 Primary care function

Diagnostic decision support was the most common primary care function addressed in studies (148 studies, 36.5%), followed by treatment decision support (56 studies, 13.8%), and then using AI for extracting information from data sources such as EHRs (49 studies, 12.1%). The most frequent combination of functions was information extraction and description (21 studies, 5.2%). Figure 3.3 summarizes primary care function counts; Appendix Figure C.2 presents more detail.

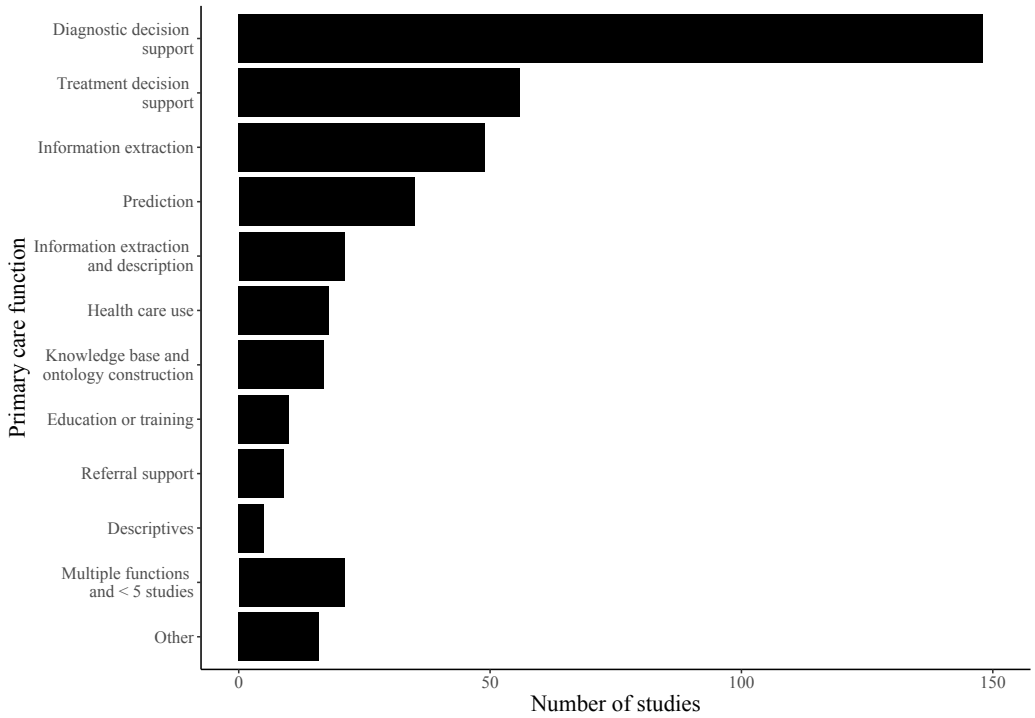


Figure 3.3: Primary care functions to be supported with artificial intelligence.

3.4.5 Reported target end user

The majority of studies reported physicians as a target end user, either alone or in combination with other target end users (243 studies, 60%). There appeared to be no positive association between having physicians as a target end user and having at least one author with a medical appointment: the percentage of studies with at least one author with any kind of medical appointment was similar between studies with physician and exclusively nonphysician target end users (51.9% and 46.3%, respectively). Twenty-six studies (6.4%) stated that their research was intended for patients, 25 (6.2%) for administrative use, and 9 (2.2%) for nurses or nurse practitioners, either alone or in combination with other end users. Appendix Figure C.3 shows the number of studies that included each of the target end user categories; Appendix Figure C.4 presents all combinations on a per-study basis.

3.4.6 Health condition

About one-quarter of studies (108 studies, 26.7%) focused on developing, using, or analyzing AI so that it would be relevant for most health conditions seen in primary care settings. Of studies that targeted a particular condition, chronic physical conditions were more frequent than acute or psychiatric conditions. We condensed target health conditions into 10 categories, with study distribution shown in Figure 3.4; Appendix Figure C.5 expands them into 27 categories.

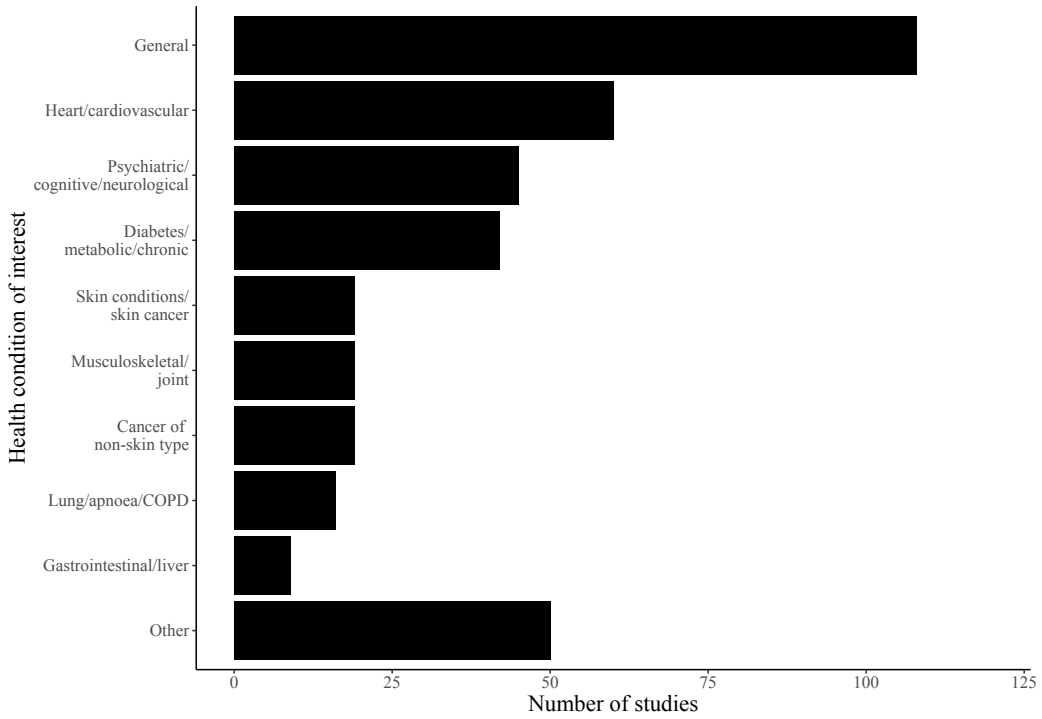


Figure 3.4: Health conditions studied.

Note: Includes only the 387 studies for which target condition(s) could be identified. *Legend*: COPD = Chronic obstructive pulmonary disease.

3.4.7 Geographic location

The location of most data source(s) used in a study or the intended location of AI implementation was higher-income countries belonging to the Organisation for Economic Co-operation and Development. Low- and middle-income countries were poorly represented. Most studies used data from a single country, with the United

States being the most common source (79 studies, 19.5%). Appendix Figure C.6 summarizes location counts and per capita rates; Appendix Table 3S C.5 contains a more detailed breakdown.

3.4.8 AI subfield

Most studies (363 studies, 89.6%) used methods within a single subfield of AI, and of these, supervised machine learning was the most common (162 studies, 40.0%), followed by expert systems (90 studies, 22.2%), and then natural language processing (35 studies, 8.6%). There were no articles on robotics. Expert systems had the earliest median year of publication (2007); data mining had the most recent (2015). Appendix Figure C.7 presents frequencies and median year of publication for 10 subfields of AI used by studies captured in our literature review; all AI subfield combinations are presented in Appendix Figure C.8.

3.5 Discussion

3.5.1 Key findings

We identified and summarized 405 research studies involving AI and primary care, and discerned three predominant trends. First, regarding authorship, the vast majority of studies did not have any primary care involvement. Second, in terms of methods, there was a shift over time from expert systems to supervised machine learning. And third, when it came to applications, studies most often developed AI to support diagnostic or treatment decisions, for chronic conditions, in higher-income countries. Overall, these findings show that AI for primary care is at an early stage of maturity for practice applications,[133,134] meaning more research is needed to assess its real-world impacts on primary care.

The dominance of TEM-appointed authors and AI methods development research is congruent with the early stage of this field. An AI-driven technology needs to be working well before real-world testing and implementation. Good performance is

achieved through methods development research, which is further reflected by most studies specifying researchers as an intended end user alongside clinicians—more work is required before implementing the AI in a practice setting. On the other hand, research focused on AI for analyzing health data is distinct and at a later stage of maturity. These AI applications are not intended for everyday clinical practice, so although their methodologic performance is important, longer-term health or workflow outcomes may not need to be assessed before real-world use.

The dominant subfields of AI identified by our review mirror trends in AI advances and align with other characteristics of the included studies. Expert systems comprise a substantial portion of the literature but are now less common (median publication year 2007 vs 2014 for supervised machine learning), reflecting a general shift in AI research from expert systems and rule-centric AI methods to machine learning and data-centric AI methods.[135] The latter are amenable to providing diagnostic and treatment recommendations as well as predicting future health, which supports primary care activities such as primary prevention and screening. This trend also aligns with the focus on physicians as target eventual end users.

Underlying drivers of AI research, and by extension maturation, are data availability and quality, particularly after the shift toward data-driven machine learning methods. The United States is the single dominant country in the field, which is unsurprising given its population, wealth, and research resources and output.[136–139] The high standing of the United Kingdom and Netherlands despite smaller populations may be attributable to primary care data availability,[140,141] facilitated by high adoption rates of EHRs,[142] and strong information technology academics and industries.[143,144] Investments in data generation, quality, and access will increase future possibilities for AI to be used to strengthen primary care in the corresponding region.

3.5.2 Strengths and limitations

Strengths of our review include a comprehensive search strategy, without date restriction, with use of inclusive eligibility criteria and conducted by an interdisciplinary

team. Limitations include multiple reviewers extracting data without double coding, English language restriction, and the lack of single widely accepted definitions for primary care or AI to guide screening. Proprietary research would not be captured by our review, nor would research completed after our search date.

3.5.3 Future research

Our next steps include further assessing the quality of the included studies and summarizing exemplary research projects. We additionally recommend a review on AI for the broader primary health care system that includes clinicians beyond physicians and nurses (e.g., social workers, physiotherapists).

For the field to mature, future research studies should have interdisciplinary teams with primary care end user engagement. Value must be placed both on developing rigorous methods and on identifying potential impacts of the developed AI on care delivery and longer-term health outcomes. Inclusion of nurses, patients, and administrators needs to increase—identifying relevant nonphysician end user activities that could be augmented by AI is an outstanding research endeavor on its own.

We expect future AI methods development to shift toward a middle ground between rule-centric and data-centric methods because interpretable models better support decisions and trust in the health care setting. For example, explainable AI is a paradigm whereby one can understand what a model is doing or why it arrives at a particular output.[145–147] Interpretability of models is additionally important from an equity lens to be able to identify and then avoid AI reproduction of biases in data, which is a present concern with data-driven methods.[148] It is also important to remember that AI is not always a superior solution: a literature review of studies published between January 2016 and August 2017 in Medline that compared prognostic prediction models for individualized prediction found comparable performance of machine learning compared with logistic regression based models.[149] This review further identified that studies generally had poor methodology and reporting, with a need for more calibration performance assessments.

3.5.4 Conclusions

Ours is the first comprehensive, interdisciplinary summary of research on AI and primary care. Two fundamental aims in the body of research emerged: providing support for clinician decisions and extracting meaningful information from primary care data. Overall, AI for primary care is an innovation that is in early stages of maturity, with few tools ready for widespread implementation. Interdisciplinary research teams including frontline clinicians and evaluation studies in primary care settings will be crucial for advancement and success of this field.

Chapter 4

Describing a Complex Primary Health Care Population in a Learning Health System to Support Future Decision Support and Artificial Intelligence Initiatives

Chapter 3 reviewed the field of of artificial intelligence (AI) and primary care research in general, finding notable gaps in research that is relevant to "real world" primary care settings. This Chapter focuses on a primary health care organization, the Alliance for Healthier Communities, to begin investigating possibilities for using AI and related techniques with their electronic health record (EHR) data to support care delivery. We generated an overview of their client population to help identify and support future initiatives, both in terms of substantive findings and in terms of methodological considerations that are relevant for work with similar populations

or primary health care organizations. More generally this Chapter demonstrates the value of descriptive epidemiology for informing learning health system (LHS) initiatives, and opportunities for unsupervised machine learning to play a role in descriptive studies of complex populations. The article of this Chapter is under revision for the *International Journal of Population Data Science.*[a]

4.1 Technical Background

This Chapter includes a large-scale descriptive epidemiology study that relies on standard techniques from epidemiology to characterize a population as well as unsupervised learning techniques to identify complex patterns. Background for both types of methods is provided herein.

4.1.1 Epidemiology

Integral to the design of an epidemiological study, and to the generalizability or impact of findings, is the cohort from which data are collected. A **cohort** refers to a group of people that are followed or observed over a period of time.[150] In a *closed cohort*, membership is defined based on eligibility criteria at the beginning of the time period and cannot change; new members cannot be added. In contrast, members of an *open cohort* can be added or removed depending on eligibility criteria that is assessed throughout the time period of interest. Eligibility criteria should include person, place, and time specifications.[151] The basis of the present study was an open cohort with membership defined based on being an adult that received primary care at the Alliance for Healthier Communities at any point in 2009-2019.

Two measures to assess the burden or risk of an outcome for a cohort include prevalence and incidence.

[a] A preprint is available: Kueper JK, Rayner J, Zwarenstein M, Lizotte DJ. Describing a complex primary health care population in a learning health system to support future decision support and artificial intelligence initiatives. medRxiv. Published online March 2, 2022. doi:10.1101/2022.03.01.22271714

- **Prevalence rate**: the prevalence of a particular outcome, e.g., disease diagnosis, is a count of the number of cases of the outcome in a defined population divided by the number of people at risk of the outcome in that population at the specified point in time.[151] *Period prevalence* is the number of "ever happened" (new, existing, recurrent) cases of an outcome across a specified time period, for a particular population, divided by the average (or midpoint estimate) size of the population at risk during that time period.[151] Note that for an open cohort this denominator is different from a count of the number of clients who ever had membership in the cohort.

- **Cumulative incidence rate**: the incidence of a particular outcome, e.g., disease diagnosis, is a count of the number of new (and/or repeat) cases within a specified time period for a particular population, divided by the number of at-risk members of that population for that same time period.[151]

Both prevalence and incidence rates depend on the denominator, data source(s), and outcome definition(s) used.

4.1.2 Unsupervised machine learning

Unsupervised machine learning algorithms are applied to unlabelled data to identify patterns or trends, which can then be interpreted by humans or used as inputs for another analysis. We used three well-established techniques in this study: Ising models to identify common co-occuring conditions, non-negative matrix factorization to identify patterns of care provider teams, and K-medoids time-series clustering to explore patterns in visit frequency.

- **Ising models:** A markov random field expresses a set of random variables (nodes) as an undirected graphical model.[152,153] An Ising model focuses on the pairwise connections (edges) between the nodes in a markov random field; learned edge weights between binary node variables represent the tendency for the two variables to be present as compared to one or both variables being absent, regardless of the state of the other variables in the graph.[154–156] These

graphs are related to Bayesian networks/directed acyclic graphs, but cyclical patterns are allowed and the goal is to understand co-occurence patterns rather than to develop a causal model or understand directions of effects. An Ising model can be developed by fitting an L1-penalized logistic regression for each variable and taking the mean of regression coefficients to arrive at "symmetised" edge weights.[156–158]

- **Non-negative matrix factorization (NMF):** NMF is a technique that factorizes a data matrix into two smaller matrices that approximate the original dataset, whereby all three matrices include only non-negative numbers.[159,160] As shown in Figure 4.1, the algorithm creates weighted collections ("topics") of codes, e.g., diagnostic codes in client EHRs. A distance metric to minimize the number of topics (k) allowed to explain the original matrix are set manually.[159–161] In our NMF analyses, the H matrix was our main focus, which has a row for each topic and a column for the amount or weight of each original code making up that topic: codes that show up together in topics tend to frequently co-occur in EHRs and vice versa. The W matrix can also be useful, such as for dimensionality reduction, as client EHRs are represented by a reduced vector of topic weights instead of all original codes.

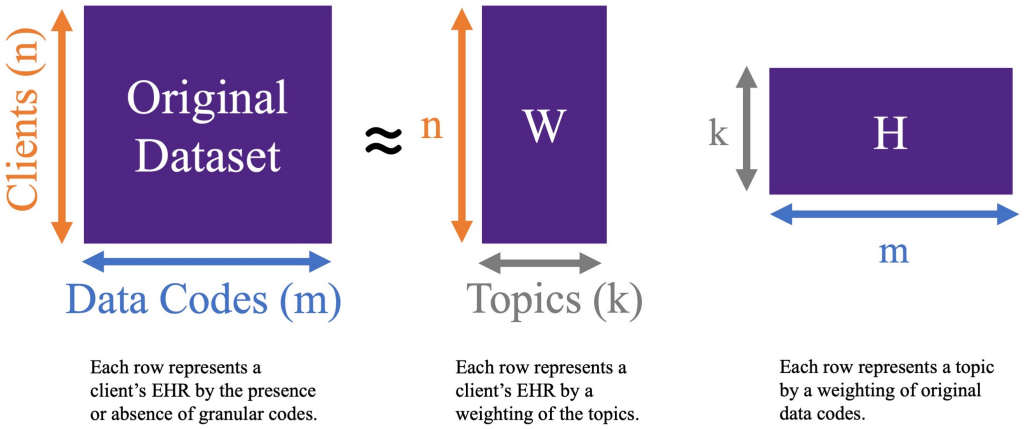


Figure 4.1: Non-negative matrix factorization example.

- **K-medoids time-series clustering:** Clustering is used to identify groups from a dataset such that e.g., clients, within the same group are more similar

to each other than to clients in other groups.[162] In contrast to topic modelling where a single client may pertain to more than one topic, in clustering the data are partitioned such that each client belongs to only one group.[162] Time-series clustering involves applying a clustering algorithm to time series data, which are ordered sequences of events.[163] We used K-medoids clustering, whereby each group is represented by the "prototype" client that has the smallest average distance to all other clients in their cluster (as opposed to the mean, which may be a fictitious value, as in K-means clustering).[163,164] We used dynamic time warping distance, which calculates the "best match" in terms of shape and magnitude between two time-series of possibly different lengths (smallest sum of absolute distances between indices, matched in a monotonically increasing fashion).[163,165–167]

4.2 Introduction

The recognized potential for analysis of EHR data to inform healthcare delivery led to the formalization of the concept of an LHS in 2007: a socio-technical system characterized by iterative cycles of data-to-knowledge-to-practice feedback.[1,3] LHS initiatives target quality improvement, research, or decision support; and usually rely on EHR data from the same population that the findings or end-product are intended to benefit.[1,4,74,75] These initiatives can support populations who have historically been excluded from medical research and clinical guideline development, such as those with complex health needs or barriers to participation.[21–23,25]

Primary care, first contact care provided in a community setting over the life course, is inherently complex.[5,6] The Alliance for Healthier Communities provides team-based primary health care through 72 Community Health Centres (CHCs) across Ontario to clients who face barriers to care and challenges, such as poverty and mental illness, that increase their risk for poor health.[9,168,169] Population health is a central element of their care model, and the Alliance officially adopted an LHS model in October 2020,[45,46] making them one of few documented primary care LHSs in North America.[4]

An LHS may pursue multiple initiatives to inform and improve care delivery. A first step towards any initiative is identifying needs of clients and providers, which is often driven by internal stakeholders.[75] Descriptive epidemiology is instrumental in outlining health states and needs of populations,[170] and may be beneficial to add into these early stages of LHS development both to identify new areas to explore and to support existing ideas. For example, describing how clients are represented in EHR data at a population level may complement clinical experience to identify potential bias or misrepresentation that analyses need to account for to obtain meaningful results.[171–173] In addition to proposed LHS benefits, descriptive studies can contribute towards closing the gap in understanding about the basic functions of primary care in general.[174]

To properly understand complex EHR data, we propose using both simple statistical techniques traditionally used in descriptive epidemiology and more complex techniques from AI, applied with an epidemiological lens. Simple techniques alone may provide an oversimplified or incorrect view of certain characteristics, which could lead to ineffective or harmful decisions later-on. So, in pursuing our primary purpose of better understanding care provided by the Alliance, we explored the suitability of a variety of techniques for epidemiology of a separate primary care system with its own EHR.

We performed the first large-scale descriptive and exploratory study of ongoing primary care clients served by the Alliance using statistical and machine learning methodology. Our *objective* was to summarize sociodemographic, clinical, and healthcare use characteristics of this population. We used unsupervised learning techniques to identify patterns of multimorbidity, care provider teams, and care access frequency. Findings provide a foundation for future Alliance LHS initiatives, including those related to their existing interest in using EHR data to segment populations and tailor care. In addition to substantive findings, this work more generally demonstrates the application of an epidemiological lens and use of a variety of methods from statistics and AI to effectively describe a complex population and

contribute to early stages of an LHS.

4.3 Methods

4.3.1 Study population and data source

We used a de-identified extract of the centralized, structured EHR database from all CHCs; unique identifiers allowed tracking of client care over time. Issues addressed during care were recorded using Electronic Nomenclature and Classification Of Disorders and Encounters for Family Medicine (ENCODE-FM)[175] and International Classification of Disease (ICD)-10 vocabularies.[176] Primary care EHRs represent an open cohort; Supplementary Figure D.1 shows the cohort size along calendar- and observation-based time definitions. Clients eligible for inclusion were over 18 years old in 2009, indicated a CHC as their primary care provider, and had at least one encounter at a CHC in 2009 to 2019. Any additional eligibility for specific analyses is described as needed below. We followed RECORD reporting guidelines (Appendix D).[177]

4.3.2 General analysis plan

Sociodemographic, clinical, and healthcare use characteristics are defined in Appendix Table D.1. Methods specific to each category are described below; we performed "table-based summaries" for all, whereby categorical variables were summarized by counts and percentages, and continuous variables by the range, median, mean, and standard deviation. Where specified, findings were stratified by client multimorbidity status (defined below) or CHC "urban at-risk" (UAR) status, referring to CHCs located in major urban geographical areas that serve priority populations defined by homelessness and/or mental health and substance use challenges.[178] CHCs without UAR designation still focus on clients with barriers to care but may be in rural or urban settings and do not solely serve clients with the aforementioned complexities.[178]

4.3.3 Sociodemographic characteristics

We conducted table-based summaries for select fields from the structured EHR client characteristic table and certain ENCODE-FM-derived variables. Missingness of the former occurred at the 1) CHC or provider level, whereby a client was not asked about the characteristic and 2) client level, whereby a client was asked and preferred to not respond. Results are presented overall and stratified by UAR and multimorbidity status.

4.3.4 Clinical characteristics

We investigated 20 chronic conditions that define multimorbidity in primary care research[179–181] and an additional four conditions of interest identified by Alliance stakeholders. For each condition, clients were assumed to receive related care upon the first record of a relevant code. We explored conditions in single, composite, and pairwise manners.

4.3.4.1 Prevalence and incidence

To provide different perspectives on clinical complexity, we calculated two measures of prevalence and one measure of incidence for each of the 24 conditions. We also calculated prevalence of multimorbidity. Our primary multimorbidity definition, including for stratification, was presence of at least three of the 20 chronic conditions.[179–181] We also looked at multimorbidity of at least two conditions, as this is another commonly used definition.[180]

1) *Eleven-year period prevalence*, based on calendar time, to assess the burden of conditions over the entire observation period (2009-2019). For each condition, we divided the number of clients who ever received a condition indication by an estimate of the average population size (technical details in Appendix D.4). Sensitivity analyses included the largest possible denominator: total number of eligible clients, and the smallest reasonable denominator: starting with the middle calendar year (2014), additional clients with at least one visit in adjacent

years were added until no prevalence estimate was over 100%. Results are shown overall and UAR-stratified.

2) *Observation-based period prevalence*, based on length of client observation, to assess the burden of conditions dependent on the number of years clients received care at a CHC. To calculate this, we separated clients into 11 sub-cohorts based on the number of years (consecutive 365.25 day intervals, rounded up) between their first and last recorded events. For each sub-cohort and condition, we divided the number of clients who ever received a condition indication by the number of clients in the sub-cohort. Results are presented as bar graphs.

3) *Cumulative incidence*, to assess the rate of condition indications by days of observation. We plotted cumulative incidence curves using the R package `survival`.[182] To prioritize capture of incident condition-related care, we excluded clients with conditions recorded in 2009 from this analysis.

4.3.4.2 Condition co-occurrence patterns

To assess co-occurrence for each pair of conditions while adjusting for all of the other conditions, we estimated an *Ising model* using R package `MRFcov`[157,158] for all conditions except Hepatitis C (Alliance-suggested condition that overlaps with one of the 20 chronic conditions). We converted coefficients, representing the strength of association between each condition pair adjusted for all other conditions, to odds ratios and interpreted size using Chen et al. (2010) guidelines.[183] We also viewed the top frequency-based co-occurrences.

4.3.5 Healthcare use characteristics

We performed table-based summaries of provider and care access characteristics overall and stratified by UAR CHC, Rural Geography CHC, and client multimorbidity status.

4.3.5.1 Providers involved

To identify common care provider teams that clients were exposed to across their care histories, we used *NMF*[184] to identify frequently-occurring: 1) *"Ever-seen" teams* whereby dummy variables were used to indicate whether each provider type was ever involved in care, and 2) *Relative "amount-seen" teams* based on volume of care whereby the number of events associated with each provider type was normalized within clients. For each version, we ran analyses allowing 2,3,5,10, and 15 topics (provider teams) with the Python package `sklearn.decomposition.NMF` and the Kullback-Leibler divergence distance metric.[185] We interpreted resulting topics by visual inspection. Provider types were maintained as recorded in the EHR except "Other", "Unknown", and "Undefined" were combined. We also summarized the top frequency-based provider types involved in care and referrals. Eligible clients required at least one provider type indication in their EHR.

4.3.5.2 Care access patterns

We measured *complexity of care* as the number of events (distinct issues addressed or types of care received) per visit (calendar day of access) to a CHC, and *care frequency* as the number of calendar days at least one event was recorded per year (365.25 day intervals) and per quarter-year (90.30 day intervals). To investigate frequency of care in terms of magnitude and shape (changes in magnitude across care histories), we performed *time series clustering* with the K-medoids algorithm and dynamic time warping distance metric[163] for 1) *short-term clients* with 2-3 observation years and 2) *long-term clients* with 8-10 observation years. For each time interval and cohort, we used R package `dtwclust`[166] to identify 2,3,4, and 5 clusters. Performance was assessed using the silhouette score and visual inspection.

4.4 Results

Of the 881,129 adult clients in the Alliance EHR database in 2009-2019, 232,529 (26.4%) had ongoing primary care client indications, and 221,047 (25.1%) had at

least one encounter in 2009-2019. Of these eligible clients, 64,504 (29.2%) received care at least once in 2009, 141,627 (64.1%) in 2019, and 40,704 (18.4%) received care in both years.

4.4.1 Sociodemographic characteristics

Sociodemographic characteristics are described in Table 4.1, with remaining substrata in Appendix Table D.2. The UAR CHCs tended to provide care to clients who were more commonly English-speaking, and had lower levels of education, household income, immigration, stable housing, and/or food security. Clients with multimorbidity tended to be older and more commonly female, reside in rural locations, and had lower levels of education, immigration, stable residence, and/or food security.

Table 4.1: Sociodemographic characteristics.

Variable	Values	All	UAR	MM
n		221047	35998	103172
Age in 2015	25-34	55505 (25.11)	7976 (22.16)	9346 (9.06)
	35-44	45646 (20.65)	7540 (20.95)	15542 (15.06)
	45-54	44653 (20.20)	8186 (22.74)	23982 (23.24)
	55-64	37848 (17.12)	6790 (18.86)	25578 (24.79)
	65-74	23162 (10.48)	3644 (10.12)	17780 (17.23)
	75+	14233 (6.44)	1862 (5.17)	10944 (10.61)
Rural Geography Residence	Rural	49275 (22.29)	6131 (17.03)	26818 (25.99)
	Urban	167728 (75.88)	28538 (79.28)	75011 (72.70)
	Missing	4044 (1.83)	1329 (3.69)	1343 (1.30)
Sex	Female	127070 (57.49)	18699 (51.94)	59946 (58.1)
	Male	93294 (42.21)	17151 (47.64)	43124 (41.80)
	Other	331 (0.15)	43 (0.12)	19 (0.02)
	Missing	352 (0.16)	105 (0.29)	83 (0.08)
Gender	Female	41352 (18.71)	5509 (15.30)	21831 (21.16)
	Gender Diverse	340 (0.15)	112 (0.31)	144 (0.14)
	Male	29366 (13.28)	4585 (12.74)	14733 (14.28)
	Prefer not to answer	1001 (0.45)	51 (0.14)	376 (0.36)
	Missing	148988 (67.40)	25741 (71.51)	66088 (64.06)

Sexual Orientation	Bisexual	1578 (0.71)	285 (0.79)	690 (0.67)
	Gay	708 (0.32)	192 (0.53)	306 (0.30)
	Heterosexual	57065 (25.82)	8447 (23.47)	29105 (28.21)
	Lesbian	485 (0.22)	70 (0.19)	244 (0.24)
	Queer	323 (0.15)	34 (0.09)	91 (0.09)
	Two-Spirit	128 (0.06)	80 (0.22)	61 (0.06)
	Other	246 (0.11)	34 (0.09)	143 (0.14)
	Do not know	924 (0.42)	201 (0.56)	485 (0.47)
	Prefer not to answer	7561 (3.42)	877 (2.44)	4078 (3.95)
	Missing	152029 (68.78)	25778 (71.61)	67969 (65.88)
Highest Level of Education	Post-secondary or equivalent	84888 (38.40)	12056 (33.49)	35763 (34.66)
	Secondary or equivalent	61831 (27.97)	11783 (32.73)	32617 (31.61)
	Less than high school	18941 (8.57)	3266 (9.07)	10618 (10.29)
	Other	8507 (3.85)	719 (2.00)	4078 (3.95)
	Do not know	4860 (2.20)	1318 (3.66)	2350 (2.28)
	Prefer not to answer	2950 (1.33)	422 (1.17)	1585 (1.54)
	Missing	39070 (17.67)	6434 (17.87)	16161 (15.66)
Primary Language	English	167163 (75.62)	31658 (87.94)	79599 (77.15)
	French	22547 (10.20)	944 (2.62)	11091 (10.75)
	Other	26847 (12.15)	2948 (8.19)	10710 (10.38)
	Missing	4490 (2.03)	448 (1.24)	1772 (1.72)
Race and Ethnicity	Black	8861 (4.01)	725 (2.01)	3757 (3.64)
	East/SouthEast Asian	3739 (1.69)	484 (1.34)	1545 (1.50)
	Indigenous	2944 (1.33)	1577 (4.38)	1641 (1.59)
	Latino	4350 (1.97)	206 (0.57)	1708 (1.66)
	Middle Eastern	2046 (0.93)	344 (0.96)	838 (0.81)
	Other	567 (0.26)	148 (0.41)	306 (0.3)
	South Asian	3597 (1.63)	323 (0.90)	1852 (1.80)
	White	38464 (17.40)	4531 (12.59)	21504 (20.84)
	Do not know	838 (0.38)	151 (0.42)	487 (0.47)
	Prefer not to answer	2649 (1.20)	261 (0.73)	1513 (1.47)
	Missing	152992 (69.21)	27248 (75.69)	68021 (65.93)
Years since Arrival in Canada	0-5 years	13654 (6.18)	1191 (3.31)	3047 (2.95)
	6+ years	51815 (23.44)	4940 (13.72)	22722 (22.02)

	None recorded	155578 (70.38)	29867 (82.97)	77403 (75.02)
Household Income	$0 to $14,999	40519 (18.33)	8729 (24.25)	17757 (17.21)
	$15,000 to $24,999	21102 (9.55)	3555 (9.88)	11081 (10.74)
	$25,000 to $39,999	20877 (9.44)	2988 (8.3)	10736 (10.41)
	$40,000 to $59,999	17245 (7.80)	2421 (6.73)	8671 (8.40)
	$60,000 or more	28494 (12.89)	3862 (10.73)	12868 (12.47)
	Do not know	15408 (6.97)	2658 (7.38)	6264 (6.07)
	Prefer not to answer	27621 (12.50)	4130 (11.47)	14890 (14.43)
	Missing	49781 (22.52)	7655 (21.27)	20905 (20.26)
Household Composition	Couple with children	53398 (24.16)	6759 (18.78)	20713 (20.08)
	Couple without child	39664 (17.94)	5945 (16.51)	22950 (22.24)
	Extended Family	7632 (3.45)	1123 (3.12)	3581 (3.47)
	Grandparents with Grand-child(ren)	1746 (0.79)	247 (0.69)	1183 (1.15)
	Siblings	1622 (0.73)	250 (0.69)	669 (0.65)
	Single Parent	14445 (6.53)	2527 (7.02)	6348 (6.15)
	Sole Member	32782 (14.83)	7445 (20.68)	18597 (18.03)
	Unrelated housemates	8622 (3.90)	1567 (4.35)	2849 (2.76)
	Other	8913 (4.03)	1476 (4.10)	4202 (4.07)
	Do not know	2475 (1.12)	643 (1.79)	1279 (1.24)
	Prefer not to answer	3727 (1.69)	491 (1.36)	1927 (1.87)
	Missing	46021 (20.82)	7525 (20.90)	18874 (18.29)
Stable Residence	True	199349 (90.18)	28227 (78.41)	90479 (87.70)
Food Insecurity	True	10985 (4.97)	2947 (8.19)	7323 (7.10)

Legend: MM = Multimorbidity; n = Number of clients; UAR = Urban At Risk

4.4.2 Clinical characteristics

4.4.2.1 Prevalence and incidence

Eleven-year period prevalence estimates ranged from 1.5% (Hepatitis C) to 81.0% (multimorbidity of two conditions) overall, with generally higher estimates in UAR strata (Table 4.2). The low sensitivity estimate for the denominator was based on 2012-2015 (n=148,595).

Table 4.2: Eleven-year period prevalence.

Variable	All	UAR
n	165125	27256
Hypertension	68177 (41.29)	12304 (45.14)
Depression or Anxiety	23828 (14.43)	5533 (20.30)
Chronic Musculoskleletal	104304 (63.17)	18842 (69.13)
Arthritis	37201 (22.53)	6906 (25.34)
Osteoporosis	11462 (6.94)	1950 (7.15)
Asthma or COPD or Chronic Bronchitis	43837 (26.55)	9190 (33.72)
CVD	23311 (14.12)	4673 (17.14)
Heart Failure	7994 (4.84)	1564 (5.74)
Stroke or TIA	2967 (1.8)	585 (2.15)
Stomach Problem	36175 (21.91)	7620 (27.96)
Colon Problem	24949 (15.11)	4974 (18.25)
Chronic Hepatitis	13288 (8.05)	2954 (10.84)
Diabetes	35704 (21.62)	6912 (25.36)
Thyroid Disorder	24793 (15.01)	4217 (15.47)
Any Cancer	14024 (8.49)	2636 (9.67)
Kidney Disease or Failure	8290 (5.02)	1555 (5.71)
Chronic Urinary Problem	59677 (36.14)	11131 (40.84)
Dementia or AD	4776 (2.89)	898 (3.29)
Hyperlipidemia	67175 (40.68)	11659 (42.78)
Obesity	38408 (23.26)	6455 (23.68)
Hepatitis C	2436 (1.48)	1173 (4.30)
Smoking or Tobacco Use	37355 (22.62)	9597 (35.21)
Substance Use	20853 (12.63)	7508 (27.55)
Lonely or Isolated	17947 (10.87)	5149 (18.89)
MM 2+	133704 (80.97)	24129 (88.53)
MM 3+	103172 (62.48)	19237 (70.58)

Note: Denominator was the approximated average population size across all years (2009-2019). *Legend*: AD = Alzheimer's Disease; COPD = Chronic Obstructive Pulmonary Disease, CVD = Cardiovascular Disease; MM = Multimorbidity; n = Number of clients; TIA = Transient Ischemic Attack.

Observation-based period prevalence estimates tended to increase with length of observation; however, cumulative incidence plots for the 156,543 (70.8%) clients without care recorded in 2009 showed the rate of condition indications notably decreased after the first year of observation. Sample plots are in Figure 4.2; all are in Appendix Figures D.2 and D.3.

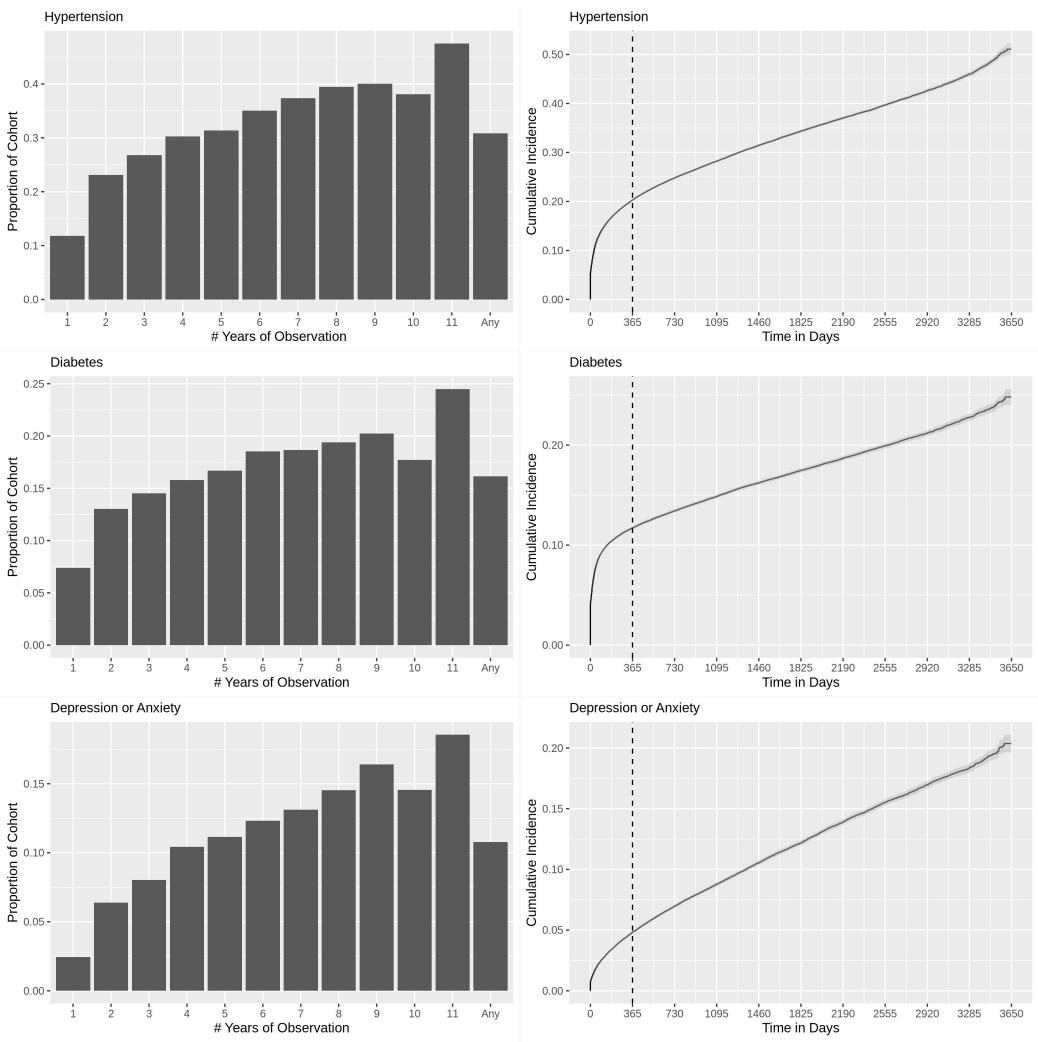


Figure 4.2: Example observation-based period prevalence and cumulative incidence plots.

Notes: Left column: Observation-based period prevalence. Right column: Cumulative incidence by days of observation.

4.4.2.2 Condition co-occurrence patterns

Among the 103,172 (46.7%) clients with multimobidity of at least three chronic conditions, there were 25,162 unique combinations ranging in frequency from 1 (<0.1%) to 845 (0.4%) clients.

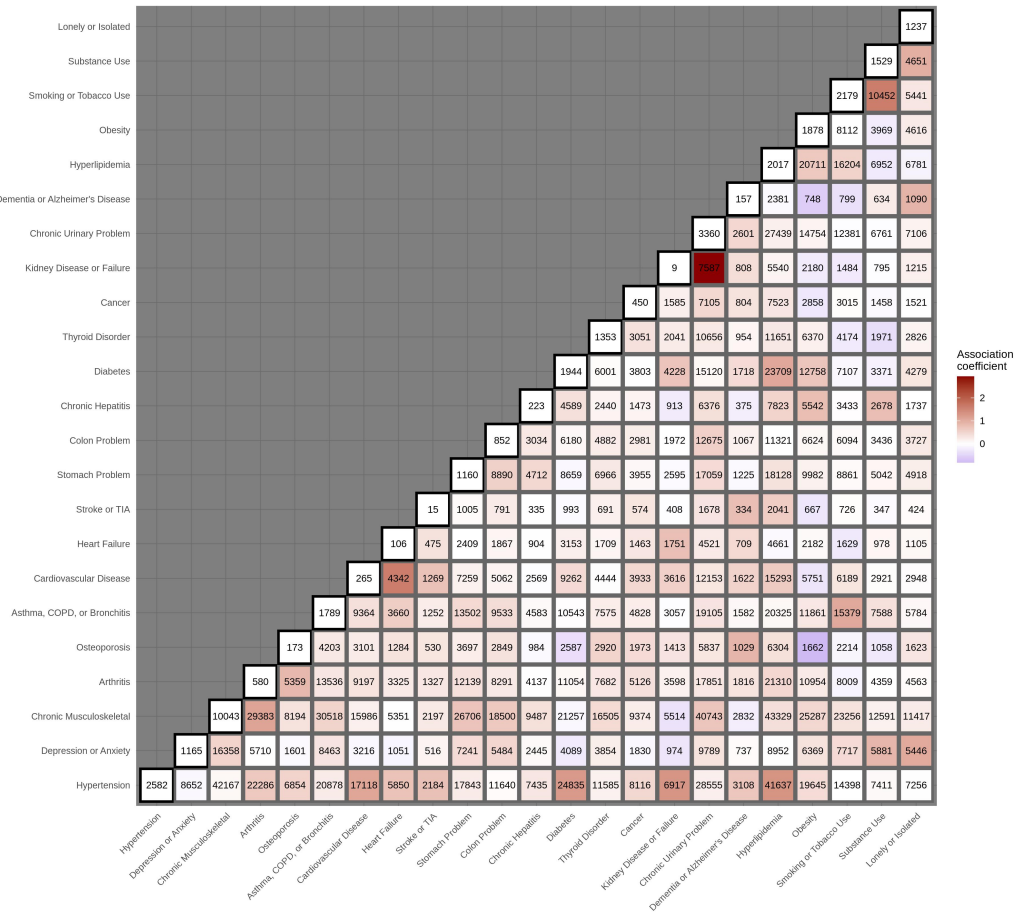


Figure 4.3: Condition co-occurrence patterns: heatmap representing the results of the Ising model.

Notes: Shading is relative to the edge weights or strength of condition co-occurrence. The numbers indicate raw counts in the data; diagonal counts represent clients who only had that single condition. *Legend*: TIA = Transient Ischemic Attack; COPD = Chronic Obstructive Pulmonary Disease.

Figure 4.3 presents the *Ising model* results. Pairwise associations between conditions on the log-odds scale ranged from -0.82 (Osteoporosis—Obesity) to 2.93 (Kidney disease or failure—Chronic urinary problem). There was one large, five medium, 40 small, and 207 very small associations based on odds ratio magnitude. The five largest positive associations were 1) Kidney Disease or Failure—Chronic Urinary Problem, 2) Smoking or Tobacco Use—Substance Use, 3) Cardiovascular Disease—Heart Failure,

4) Hypertension—Hyperlipidemia, and 5) Hypertension—Kidney Disease or Failure. In contrast, the top five co-occurring conditions based on raw frequency were 1) Hyperlipidemia—Chronic Musculoskeletal, 2) Hypertension—Chronic Musculoskeletal, 3) Hyperlipidemia—Hypertension, 4) Chronic Urinary Problem—Chronic Musculoskeletal, 5) Asthma or COPD or Chronic Bronchitis—Chronic Musculoskeletal. These directly correspond to the conditions that had the highest marginal frequencies.

4.4.3 Healthcare use characteristics

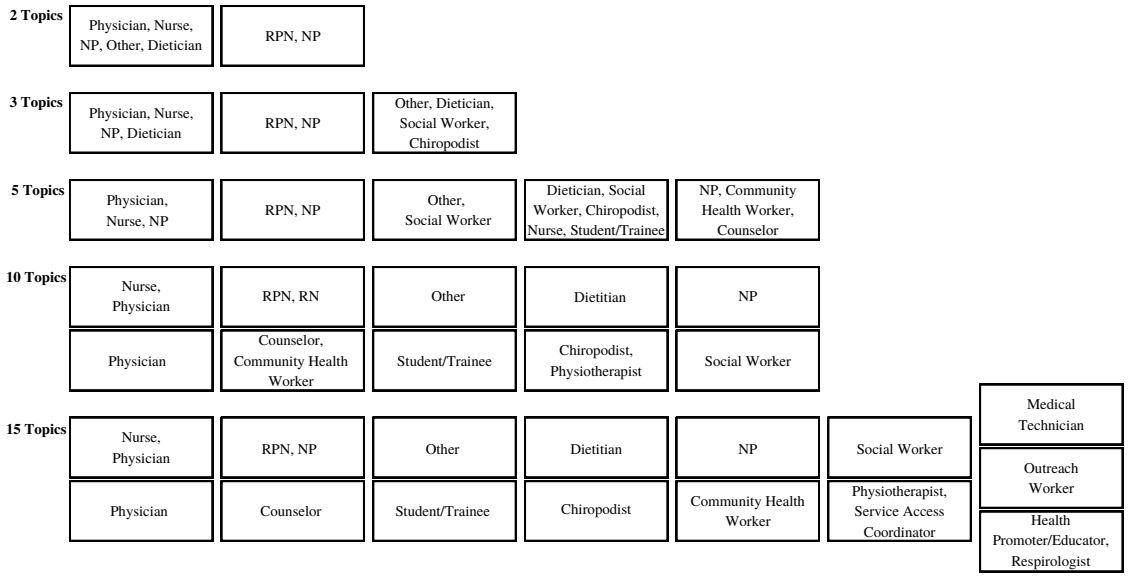
Table-based summaries of healthcare use characteristics are in Appendix Table D.3. In general, UAR CHCs had higher healthcare use while rural geography CHCs were closer to the overall population. Clients in multimorbidity strata exhibited higher healthcare use compared with the general base cohort.

4.4.3.1 Providers involved

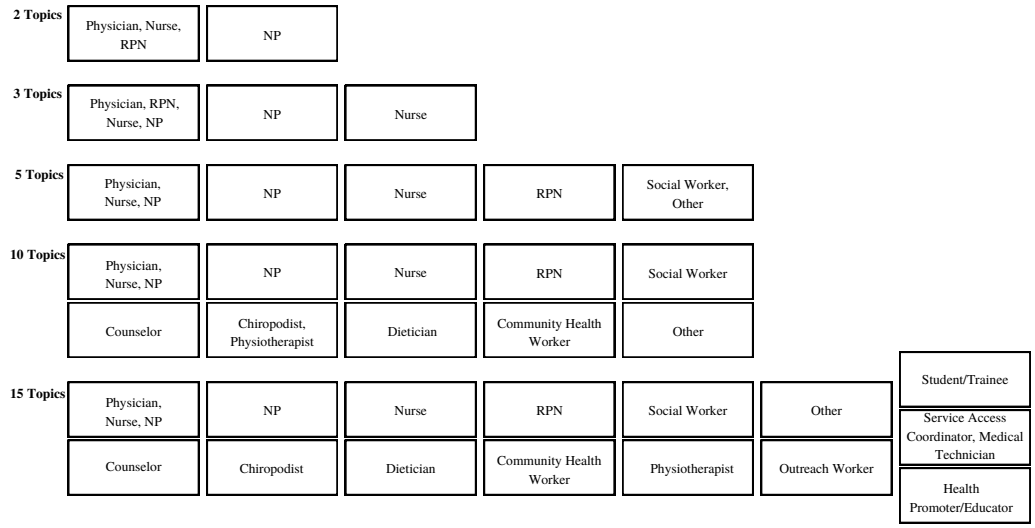
There were 19,394 unique combinations of the 68 distinct provider types seen across the 220,806 (99.9%) clients with at least one provider type recorded. In terms of referrals, 102,088 (46.2%) clients had at least one internal and 143,922 (65.1%) had at least one external referral recorded. Note internal referrals may not have captured "hallway referrals", whereby a nearby provider provides a quick consult that is not formally recorded.

Figure 4.4 shows results of the *NMF analysis*, listing the highest-weighted provider types in each topic down to a weight of three. For the *ever-seen* provider team analysis, physician and nursing provider types emerged most prominently overall. In general, as the number of topics increased, additional provider types emerged and then split apart to dominate separate topics. Exceptions were the high-weighted pairings of nurse and physician and of registered practical nurse and nurse practitioner. Overall, 18 of the 68 possible provider types emerged prominently in at least one topic; only one (respirologist) did not also appear in the amount-seen analysis.

The *amount-seen* provider team analysis had greater weight distributions between



(a) Ever-seen provider team analysis



(b) Relative amount seen provider team analysis

Figure 4.4: Common care provider teams.

Notes: Boxes represent the topics resulting from the non-negative matrix factorization analysis. Provider types are listed in order starting with the highest weighted provider; for any given topic, provider types with a weight less than three are not show. *Legend*: NP = Nurse Practitioner; RPN = Registered Practical Nurse.

provider types within topics. For example, the first of the three-topic analysis had an approximate 1:1:1:6 ratio of care provided by nurse practitioner:nurse:registered practical nurse:physician. In both versions, about half of clients had a non-zero weight for only one of the first two topics; in the amount-seen analysis more clients maintained a non-zero weight on only one topic as the number of topics increased, e.g., 16.6% versus 2.5% at five topics. In general, results suggest most clients received the majority of care from physician, nurse practitioner, or nurse provider types, usually in combination with other provider types at a lower volume of care and with heterogeneous co-occurrence. An example of patterns that emerged for other provider types include differences in timing and weight of dietician/nutritionist and social worker providers between the two analyses. Interpreted alongside the most common provider and referrals types (Appendix Table D.4), findings suggest referrals to dietitian/nutritionist were more common than to social worker, but frequent or longer-term care was more commonly provided by social workers.

4.4.3.2 Care access patterns

Complexity of care from a CHC-perspective was primarily low with 80.4% of client-visits associated with a single-issue and under 1.0% with over five issues addressed (higher intensity); however, from a client-perspective, 24,204 (11.0%) experienced at least one visit with over five issues while 38,533 (17.4%) experienced a maximum of one issue per visit across their care history. The mean *care access frequency* was 6 days per year (standard deviation=7.4). While 29,191 (13.2%) clients experienced at least one year with over 25 days, 7,455 (3.4%) averaged over 25 days per year across their entire care history. There were 8,700 (3.94%) clients with at least one frequent care period (year with over 25 days care accessed) and complex care episode (visit with over 5 issues addressed).

For the *time series clustering* analyses, the short-term cohort included 37,920 clients and 93,625 client-years of observation; the long-term cohort included 42,855 clients and 387,035 client-years of observation. The silhouette score was always highest for

two clusters (Appendix Table D.5). Visual inspection of plots (Figure 4.5) showed high variability within and between clients.

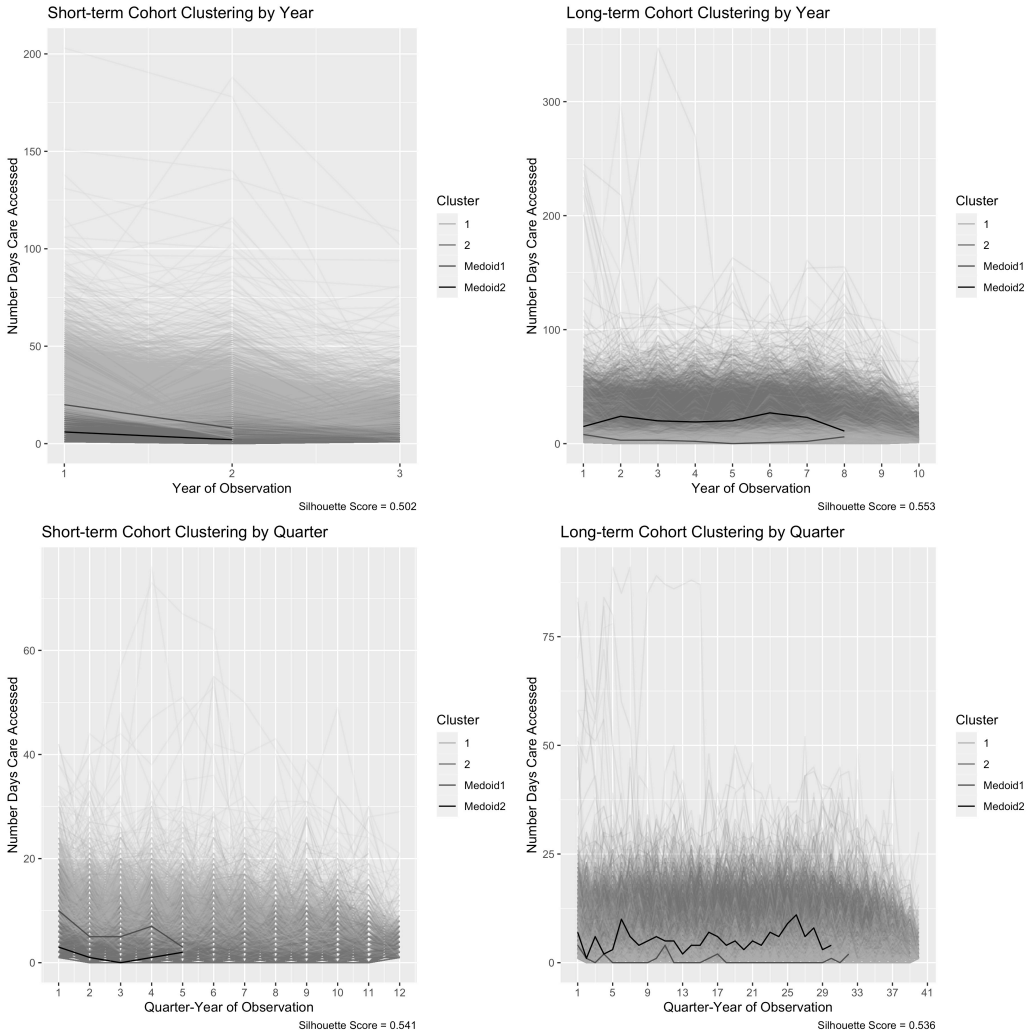


Figure 4.5: Care frequency clusters.

Notes: Results from the four time series clustering analyses for each cohort and data-representation combination. Medoids are shown with raw time series data, separated by cluster number, for the number of clusters that resulted in the highest silhouette score (SS).

4.5 Discussion

We used statistical and AI techniques to summarize sociodemographic, clinical, and healthcare use characteristics captured in the EHRs of ongoing primary care clients served by the Alliance. Substantive findings can motivate new topics for future LHS

initiatives, or help to refine existing ideas and selection of performance measures for long-term evaluation of implemented interventions. Methods-related findings may inform the approaches used in these endeavours. While our discussion focuses on LHS initiatives, as with any epidemiological study, substantive results may be immediately useful to the population of interest, e.g., to inform clinic-level case management and onboarding of new clients.

4.5.1 Sociodemographic characteristics

The CHC EHRs contain rich sociodemographic information, both the presence and absence of which is informative. Social determinants that may increase risk of poor health including lower household income and education, residence instability, and food insecurity were more prevalent in UAR CHC and multimorbidity strata. There appeared to be evidence for the healthy immigrant effect,[186] assessed by viewing the proportion of people in each category of the years since arrival in Canada variable across the multimorbidity strata: a lower proportion of people with 0-5 years in Canada had multimorbidity as compared to those with 6 or more years in Canada or no arrival information recorded (missing or born in Canada). Completeness rates varied by characteristic and may be due to client, provider, or CHC level decisions. For example, of the 72,059 (32.6%) clients asked about gender only 1001 (1.4%) preferred to not answer. In contrast, more clients, 171,266 (77.5%), were asked about household income but there was a higher tendency to not answer, 27,621 (16.1%). These findings align with a framework to assess selection bias in EHR data that suggested multiple mechanisms are usually responsible for missingness so the focus should be on "what data are observed [instead of missing] and why".[187] While provider-level decisions may be due to inferring certain characteristics or prioritizing information needed for them to direct care, completeness rates are important for decision support tool performance, which can improve with social determinants of health information.[188,189]

When assessing data quality and completeness, which is emphasized by LHS and machine learning for EHR guidelines,[1,30,75,172,190] the implications of pursuing LHS

initiatives at different levels should also be considered. For example, a subset of CHCs capture self-reported measures of health, which are valuable research outcomes.[191] While these measures are not suitable for analyses with data from all CHCs (CHC population-level initiatives), they should be considered for initiatives specific to the collecting CHCs.

4.5.2 Clinical characteristics

4.5.2.1 Prevalence and incidence

In operationalizing morbidity measures, the denominator must be defined with the intended end-goal in mind. The *eleven-year period prevalence* estimates relate to a CHC-based perspective and are useful for long-term system-level planning, while the *observation-based period prevalence* estimates are more aligned with a client-based perspective and absolute measure of risk. Another consideration is that just as ICD-10 or ENCODE-FM codes do not guarantee true condition presence, the absence of care does not verify absence of conditions.[192] For example, clients may not seek primary care when they are healthy, hospitalized, or experiencing barriers to care.

The *cumulative incidence* plots demonstrate that "risk" of condition codes is highest in the first year of observation. Clinically this makes sense, as new clients may have a build-up of unmet care needs. Nonetheless, there are important takeaways for LHS initiatives that require cohort construction. For example, predictive models developed for decision support need to account for the almost qualitative change in risk related to being a new client. Although this care pattern is somewhat unique to primary care settings, methods developed for related problems may be useful. For example, accounting for variable lengths of stay in intensive care unit EHRs,[193] or handling cold-starts and sparse data for recommender systems.[194]

4.5.2.2 Condition co-occurrence patterns

There was a high prevalence of multimorbidity, but with so many different multimorbidity "compositions" it is hard to see how to make use of the category of

multimorbidity. The *Ising model* demonstrated how to go beyond frequency-based comparisons and identify relationships between conditions irrespective of others, but again, this presents as a long tail problem, with very few combinations that are very prominent. Primary care decision support tools will face the challenge of making recommendations on many different and possibly co-occurring conditions. Most distinct multimorbidity compositions are rare events, making it unrealistic to generate standardized responses or specialized evidence bases for each, especially when medical condition combinations are considered alongside sociodemographics. The majority of existing decision support tools and clinical guidelines focus on a single condition at a time; new techniques for providing evidence-based guidelines or recommendations for these vast numbers of combinations are needed.[195–198]

4.5.3 Healthcare use characteristics

4.5.3.1 Providers involved

While care for ongoing primary care clients is typically led by physicians or nurse practitioners, CHCs include many provider types and LHS initiatives may choose to focus on particular provider type(s). The *NMF analyses* more easily identify reliable patterns of commonly seen provider types and teams than manually sifting through extensive count-based tables. Another use for NMF is dimensionality reduction or data pre-processing, whereby data are summarized to reduce the number of variables that need to be included in an analysis.[184] For example, NMF-derived topics could be used as inputs to a predictive model instead of separate variables to represent each provider type or specific, manually selected combinations.

4.5.3.2 Care access patterns

Complexity of care from a CHC system-level perspective was primarily low intensity (few problems addressed per visit), although this may be partly due to data quality such as if only one issue was recorded in the EHR when multiple were actually addressed in the appointment. The subset of clients who experienced higher care

complexity did not tend to also have high frequency of care. Sporadic visit patterns may be due to unstable living arrangements or demanding life responsibilities; when there is uncertainty about when a client will return, providers may pack together multiple types of care. The marginal distribution of *care frequency* was right-skewed without a distinct break; most clients experienced lower care frequency, but higher frequencies were also observed. In contrast to expectations, we did not identify consistent, distinct client groupings through the time-series clustering, e.g., to indicate a subpopulation of "frequent visitors." This may be due to restrictions in the types of similarity that dynamic time warping captures. Future analyses could try a different similarity metric or including covariates to account for baseline variability.

4.5.4 Strengths and limitations

Strengths included the strong interdisciplinary approach used to assess complex, longitudinal EHR data. We used chronic condition definitions recommended for primary care research,[179–181] although the algorithms have not been validated for CHCs specifically. Our broad cohort definition supported a high-level overview of the population, but may not be appropriate for specific research questions.

4.5.5 Conclusions

We demonstrated the use of simple statistics and AI techniques, applied with an epidemiological lens, to describe EHR data from a budding LHS. Substantive findings lay a foundation for future Alliance initiatives and may be informative for other organizations serving complex primary care populations.

Key suggestions for future LHS initiatives include the need to carefully deliberate the level of analysis, or who a given initiative should be targeted at (e.g., population or specific CHCs, one or many clinical presentations, all or subset of providers), and the associated implications for how clients will be represented in the data. Representation will depend on analytical-, system-, provider-, and client-level factors. Decision support initiatives need to consider heterogeneity in conditions and care access patterns,

including non-uniform risk of condition indications across observation history.

Chapter 5

Hybrid Feature- and Similarity-Based Models for Prediction and Interpretation on Large-Scale Observational Data

Chapter 3 concluded with a call for more artificial intelligence (AI) research that is relevant to primary care settings and Chapter 4 highlighted complexity that can be present within these settings. Chapter 5 presents machine learning methodology that is designed to take advantage of rich but heterogeneous observational data sources like primary care electronic health records (EHRs). Prediction and interpretation using the proposed hybrid feature- and similarity-based model is demonstrated with synthetic data and in a case study with the Alliance for Healthier Communities, informed by Chapter 4. The work in this Chapter will be submitted as a full research paper to a *machine learning conference*, whereby it undergoes review and if accepted, will be published in the associated proceedings.[a]

[a] A preprint is available: Kueper JK, Rayner J, Lizotte DJ. Hybrid Feature- and Similarity-based models for prediction and interpretation using large-scale observational data. arXiv:2204.06076v1 [cs.AI]. Published online April 12, 2022.

5.1 Technical Background

This Chapter addresses supervised learning tasks where the outcome of interest $y(o)$ for a particular observation o can be explained partly by constructed features $\phi(o)$ (i.e. scalar or fixed-length vectors that represent a characteristic or property of o) and partly by more complex information $\Psi(o)$ (e.g., high-dimensional, time-varying, variable-length data). For example, in a primary health care setting each observation could be a client and the outcome of interest a condition or situation, such as diabetes or food insecurity, that the client is at risk for and early intervention may help to prevent. In this setting, $\phi(o)$-type information may include sociodemographic characteristics and current diagnoses while $\Psi(o)$-type information may include years of encounter data representing the subset of thousands of possible tests, diagnoses, and procedures that the client has received in their lifetime.

5.1.1 Feature-based learning

Feature-based supervised models, such as trees, learn a mathematical function that takes features as input and provides an estimate about the outcome as output.[55,153] To use a feature-based approach for the above scenario, the $\Psi(o)$-type information must be converted into $\phi(o)$-type information. This can be done in a data-driven way and/or based on medical or social theories of health. For example, a data-driven dimensionality reduction approach such as topic modelling could be applied to client histories of diagnostic codes and the resulting topic weights for each client used as features. A theoretical approach may include identifying known risk-factors for the outcome based on research literature or clinical expertise, and then collapsing specific subsets of codes to generate binary indicators for whether or not the client has ever experienced each risk factor. Oftentimes feature construction loses or misses information and in general the extent to which this is a disadvantage will depend on the complexity of the data and predictive task.

In this Chapter we use logistic regression (LR) to relate $\phi(o)$-type information to the outcome. A prediction for a client of interest o based on a set of j features can be

written $\hat{y}(o) = \varsigma\left(\sum_j \phi_j(o)\beta_j\right)$ where ς is the sigmoid function.[55,153]

5.1.2 Similarity-based learning

Similarity-based approaches, such nearest-neighbour and kernel methods, can handle both $\phi(o)$-type and $\Psi(o)$-type information as inputs.[55,153] Instead of learning explicit relationships between individual inputs and the outcome, these methods use similarities or distances between observations by assuming that similar o are likely to experience the same outcome, and if the same o is entered into the model twice the same prediction will result. There are potentially two challenging aspects of $\Psi(o)$ that make similarity-based approaches attractive: the dimensionality and the proper form. In some situations, the best form for $\Psi(o)$ is known but its dimension is too large or challenging to construct with traditional feature-based approaches; other times, even if $\Psi(o)$ is a manageable size, the most useful way to incorporate it into a model is unknown.

Kernels can handle both of these challenges and are the similarity-based approach used in the remainder of this Chapter. A kernel function $k : \mathbb{R}^m \times \mathbb{R}^m \mapsto \mathbb{R}$ expresses the inner product between two inputs that have been mapped to some high-dimensional feature space.[55,153] The feature mapping defines the notion of similarity captured by the scalar output, and can be non-linear with infinite dimensions; the mapping does not need to be made explicit to use a kernel function. A valid *kernel function* must be symmetric and result in a positive semi-definite kernel matrix. The notion(s) of similarity to capture, and whether data pre-processing is warranted, will depend on the specific scenario.

5.1.2.1 Select kernel functions

The simplest kernel function is the **linear kernel**, $k(o_i, o_j) = o_i^T o_j$, which can be used to create a dual formulation of linear regression that uses o directly as features.

An example of a more complex and commonly used kernel function is the **Gaussian or Radial Basis Function (RBF)**, $k(o_i, o_j) = e^{-\frac{||o_i - o_j||_2^2}{2\sigma^2}}$, which has a feature space

with infinite dimensions and one hyperparameter σ to be tuned.[153] The RBF kernel is used in our simulation study experiments.

The **Jaccard kernel function**, described here, and proposed extension, described in the article, is used in our clinical case study. The Jaccard similarity between two sets of codes, A and B, is $J(A,B) = \frac{(A \cap B)}{(A \cup B)}$. The highest similarity is when all codes in A are in B and vice versa, and the lowest similarity is when no codes that are in A are also in B, regardless of the number of codes. The relative prevalence of codes does not matter, which leads to potentially limiting characteristics in situations with a large number of possible codes. For example, when trying to capture similarity with the Electronic Nomenclature and Classification Of Disorders and Encounters for Family Medicine (ENCODE-FM)[175] vocabulary that includes over 4,000 unique codes to record care activities. Strengths and limitations of the Jaccard are introduced below through simple examples:

Imagine a scenario where a "model" client comes in for a blood test, receives a diagnosis of hypertension, and then comes in for a follow-up appointment that includes a prescription renewal. As shown in Table 5.1a, there are three other clients that have the same diagnosis, but different surrounding treatment. The Jaccard similarity (Table 5.1b) appears to work well in that:

- Client 1 similarity with others is proportional to the number of codes the other clients have.
- Clients 2 and 3 have comparable similarity profiles.

Table 5.1: Jaccard similarity score example.

(a) Client care profiles

Client	Blood	Diagnosis	Prescription
C1	TRUE	TRUE	TRUE
C2	TRUE	TRUE	FALSE
C3	FALSE	TRUE	TRUE
C4	FALSE	TRUE	FALSE

(b) Jaccard similarity

	C1	C2	C3	C4
C1	1.00	0.67	0.67	0.33
C2	0.67	1.00	0.33	0.50
C3	0.67	0.33	1.00	0.50
C4	0.33	0.50	0.50	1.00

Now imagine there is another code, "Com", that is so common everyone has it:

similarity scores all increase, but the ratio in terms of who is most similar to who else stays the same. While this type of universally present code is not overly concerning for predictive purposes, a limitation emerges when we imagine what happens to the similarity score for two clients that *do not have a very common code* (Table 5.2a and 5.2b):

- Their similarity with everyone else decreases. (This is desirable.)
- Their similarity with each other does not increase even though they share an "abnormality" from the population expectation. (We hypothesize that this is not desirable.)

Table 5.2: Jaccard similarity score example with common code.

(a) Client care profiles

Client	Blood	Diagnosis	Prescription	Com
C1	TRUE	TRUE	TRUE	TRUE
C2	TRUE	TRUE	FALSE	TRUE
C3	FALSE	TRUE	TRUE	TRUE
C4	FALSE	TRUE	FALSE	TRUE
C3b	FALSE	TRUE	TRUE	FALSE
C4b	FALSE	TRUE	FALSE	FALSE

(b) Jaccard similarity

	C1	C2	C3	C4	C3b	C4b
C1	1.00	0.75	0.75	0.50	0.50	0.25
C2	0.75	1.00	0.50	0.67	0.25	0.33
C3	0.75	0.50	1.00	0.67	0.67	0.33
C4	0.50	0.67	0.67	1.00	0.33	0.50
C3b	0.50	0.25	0.67	0.33	1.00	0.50
C4b	0.25	0.33	0.33	0.50	0.50	1.00

Sharing in the absence of common codes is not explicitly worked into the Jaccard similarity score. There are situations where the absence of common codes may matter more than the presence, for example missing check-ups after the diagnosis of a new condition or not receiving screening tests (when eligible). The same is not true for rare codes, where the presence of codes is generally expected to be more important than the absence. So, simply reverse coding everything or removing codes from the

universe is not a suitable solution. A modification to the Jaccard function to account for limitations is presented in this Chapter's article.

5.1.3 Model selection and performance

Supervised machine learning model development for predictive tasks typically includes training and then comparing several different candidate models to select the best one for use or perhaps further development and testing. For the purposes of this section, "best" refers to predictive performance.

Model selection refers to the process of training and comparing several different models, such as a feature-based model with a similarity-based model, and/or comparing models that have the same form but different hyperparameters, such penalization strength.[162] A key component of model selection is that models are trained on different data than used to assess predictive performance, so that potential issues like overfitting can be identified.[162,199] Two main ways to achieve this are with data splitting and cross validation.

- *Data Splitting* is when a dataset is segmented into separate training and validation sets.[162,199] The training set is used to learn parameters (e.g., feature coefficients), and then the trained model is applied to the validation set to get predictions with new observations.[162,199] Performance is compared using the validation set predictions.

- *Cross validation (CV)* uses resampling to generate multiple (k) training/validation splits of a dataset, often referred to as "folds".[162,199] The entire dataset is divided into k equal sized subsets; each subset is given a turn at being the validation set, whereby predictions are made from models trained on the remaining $k-1$ subsets.[162,199] Model comparison is typically based on the average performance across the k validation sets. Nested CV is when the CV process is performed for each of the training folds; the model selected based on the "inner loop" is re-trained on all "outer loop" training fold data before making predictions on the outer loop validation data. An example use of nested CV is

to select hyperparameters on an inner loop, so that the "best form" of several different model types can be compared on the outer loop validation folds.[162,199]

Model performance: Three common performance metrics to assess a model that predicts the probability (risk) of a binary outcome:

- *Discrimination* refers to the ability to assign a higher probability of the outcome to an observation with the outcome present as compared to one without the outcome.[200,201] Discrimination performance can be summarized by the area under the receiver operating characteristic curve ($AUROC$) or c-statistic.[200,201]

- *Calibration* assesses how well predicted probabilities match observed proportions. For a well-calibrated model, of observations assigned a given probability of the outcome, a similar proportion will truly be assigned with the outcome.[200,201] Calibration can be assessed with a calibration plot. The intercept indicates whether the predicted probabilities are generally overestimates (< 0) or underestimates (> 0) and the slope represents whether the estimates are generally too high (< 1) or low (> 1).[201]

- *Precision* is the proportion of correct positive predictions made from all positive predictions (true and false) and *recall* is the proportion of correct positive predictions made from all possible positive cases (true positive and false negatives).[202,203] A *precision-recall curve* plots these two metrics across different probability thresholds, with the area under the curve ($AUPRC$) serving as a summary performance metric.[202,203]

Of note, if a single cut-off value is selected, such that probabilities above the threshold are considered positive outcome predictions, a confusion matrix can be constructed and additional metrics like accuracy and positive predictive value calculated.

Above describes the model selection process and performance metrics that may inform the selection as well as more general evaluation. Once a final model is selected, an estimate of its generalization error should be obtained using data that were not part of the model selection process. In internal validation, this would be a held out subset

of data from the same data source used for model training and selection.[162,204,205] In external validation this would be data from a different data source, e.g., different healthcare system population.[162,204,205] Note that for models intended for use in healthcare, clinical validation is also needed.[204,205]

5.1.4 Standardization

Standardization is a technique used in epidemiology to investigate outcome rate(s), such as to estimate how many cases of an outcome are in a subpopulation where data are not available or to assess whether the number of cases differs from what would be expected based on rates from a reference population. Rates may be adjusted for characteristics, such as sex or age, by which the outcome prevalence is expected to differ. In the following study we use indirect standardization to calculate standardized morbidity ratios, which assess the sex-adjusted rate of the outcome in a subset of the eligible cohort as compared to the rest of the eligible cohort. $Standardized\ Morbidity\ Ratio = \frac{Observed\ Outcome\ Cases}{Expected\ Outcome\ Cases} = \frac{\sum_i x_i}{\sum_i (M_i * n_i)}$ where x_i represents the number of outcome cases in strata i of the population under investigation, M represents the outcome rate in strata i of the reference population, and n represents the number of clients in strata i in the population under investigation.[206] If $SMR = 1$ then the sex-adjusted rates are the same in the two populations.

5.2 Introduction

Health care settings generate large amounts of data and yet it can be challenging to fully harness these data for machine learning applications. For machine learning tasks with large-scale observational data, there are often known, informative features as well as additional data that may be useful for the task but are challenging to summarize into meaningful features due to size or complexity. For example, EHRs capture client characteristics (e.g., year of birth) in structured fields and record information arising from each encounter (e.g., date-stamped diagnosis and procedure codes) in dynamic tables. The former may be well suited for features while the latter high-dimensional,

variable length data may be better represented in terms of similarity to other clients in the database. Explainable, reproducible methods that take full advantage of these rich data are needed to support further advancements in the field of machine learning for healthcare.[32,38–41,43,207]

Feature- and similarity-based models have complimentary characteristics. Feature-based approaches, such as logistic regression (*LR*), tend to be more familiar to end-users, less susceptible to overfitting, and easier to interpret (e.g., viewing regression coefficients or the structure of a decision tree); however, not all valuable information can be captured with features and model performance may suffer from underfitting, especially for heterogeneous populations. In contrast, similarity-based approaches such as multiple kernel learning have a higher computation cost but can incorporate more complex or time-varying data that may account for additional variability in the outcome.[208–210] Interpretation of similarity-based approaches is not as straightforward as for feature-based methods, but can include strategies such as summarizing characteristics about the most similar training examples used to train a model to the one for whom a prediction is being made.[40,211]. Similarity-based approaches are not interpreted on their own for the purpose of causal inference, while feature based approaches may be, either explicitly in estimating a treatment effect or implicitly by interpreting feature coefficients to identify risk factors to intervene on.

We present two variations on an intrinsically interpretable *hybrid feature- and similarity-based model* (*HFSM*) and demonstrate their use with synthetic data and with EHR data from a complex primary health care population. The model form is able to support traditional causal interpretations of feature coefficients while reaping additional benefits from similarity based approaches, such as improved absolute risk prediction while maintaining traditional feature interpretations, or adjustment for complex confounders. Our experiments found the *HFSM* approach can outperform solely feature- or similarity-based methods while retaining or enhancing interpretability.

5.2.1 Generalizable insights about machine learning in the context of healthcare

Our primary contributions are through the hybrid model, both in being the first to present the model structure, and in describing the types of supervised learning scenarios where combining feature- and similarity-based approaches within an inherently interpretable model may be beneficial. Our hybrid model enables incorporation of key prior knowledge in ways not possible with existing methods. Applications include prediction-oriented tasks, e.g., to underlie a clinical decision support system, and exploratory or causal analyses, e.g., to learn about a population. We describe how our model can be used for "traditional" clinical epidemiology modelling and use simple examples to demonstrate situations where feature coefficients may become more or less biased depending on other characteristics of the model—these concepts apply to *HFSM* as well as to any other multivariable/multicomponent model that may be interpreted for decision making.

Additional contributions are made through our clinical case study, wherein we applied *HFSM* using a new strategy for building kernels that assesses similarity in terms of both the presence of rare care characteristics and the absence of common care characteristics. Assessing similarity in terms of what expected characteristics are missing may be useful for other settings (e.g., public health, emergency room triage) where two people that deviate from population-level expectations are more similar than if they fit the expected profile. We discuss additional challenges encountered in our applied setting that are relevant to other health care contexts as well, such as the open cohort nature of primary health care and decisions related to features that are informative but rare.

5.3 Related Work

Our hybrid model approach contributes to two general areas of research: 1) methods designed to incorporate multiple sources or types of data and 2) combining simple and

complex models to make a single prediction. Our approach to measuring similarity in the clinical case study additionally contributes to research on kernel functions for clinical data.

5.3.1 Integrating multiple types of data

Recommender systems, e.g., for movies or products, are often designed for settings with two distinct types of data: 1) user attributes, such as demographic information and 2) time-varying, high-dimensional information arising from user interactions with a system, such as histories of movie viewings or ratings. Fan et al. (2017) developed *RIT-UA*, which makes predictions based on a weighted linear combination of two similarity scores: one based on a weighted count of common attributes and one based on sigmoid functions applied to historical data about user preferences and ratings.[212] Our *HFSM* approach is designed to handle data with a similar structure; however, the *RIT-UA* generates scalar similarity scores to combine information from the two data sources whereas *HFSM* maintains separate model structure to use information from features directly.

Multiview learning combines multiple data types to improve predictive performance. Lian et al. (2015) proposed a framework that assumes all feature and/or similarity matrices contribute a different "view" of the data.[213] A shared latent factor matrix is learned to serve as a global representation of the data. Multiple kernel learning is a special case within this framework where each view is treated as a kernel matrix.[208–210,213] While there is overlap in the similarity-based part of this approach to *HFSM*, including the possibility to incorporate multiple kernel learning techniques, *HFSM* maintains separation of the feature matrix in a way that also prioritizes interpretation of individual feature coefficients.

5.3.2 Combining model types

Boosting approaches may also handle diverse data types by combining complimentary model forms to improve predictions. Hothorn et al. (2010) developed *mboost*, a

component-wise boosting algorithm that combines penalized least square estimates and/or regression tree base learners in an additive model structure.[214] Each component may be applied to all or a subset of data, is weighted in the fitted model, and can be interpreted separately.[214,215] Our sequentially-optimized *HFSM* approach is similar to *mboost*, but uses different components and does not employ an overall weight for each model component. Building on *mboost*, Sigrist et al. (2021) developed *KTBoost*, which learns both a regression tree and a reproducing kernel Hilbert space regression function on all available data in each iteration, and then adds the one that is expected to result in better performance to the ensemble of base learners.[216] This approach does not segregate data and does not allow for feature and kernel coefficients to be jointly optimized as in our simultaneous *HFSM* approach. A popular gradient boosting technique is *XGBoost*, which continues to fit new decision tree models to account for residual errors from previous models until performance stops improving.[217] *XGBoost* has demonstrated excellent predictive performance in several settings, but as with the other boosting techniques, the focus is on predictions. Our *HFSM* is parametric with a fully convex objective function; this supports reproducibility, which is particularly important when interpretation of the model may be used to learn about a population or to support clinical decision making.

5.3.3 Kernel functions for clinical data

Kernel functions are commonly used to capture similarity or distance, and several functions exist that could be applied to indicator data in healthcare, such as diagnostic and procedure codes from client care histories. Shawe-Taylor and Cristianini (2004) review standard kernels for sets or strings that could be applied (e.g., intersection kernel, union complement kernel, agreement kernel); however, these are solely based on present elements, all equally weighted.[208] For many clinical scenarios, tailoring similarity measures based on the frequency or type of input data elements is expected to be advantageous.

Klenk et al. (2010) proposed using regression techniques to weight the importance of

each input variable for assessing overall similarity between two clients.[218] Belanche et al. (2013) defined a similarity score for input variables that depends on their probability of occurring in the training data; "rarer" variables are given higher weights.[219] Similar to these two studies, we are interested in applying weights derived based on training data to differentiate the treatment of input variables; however, while these methods focus solely on presence of variables we additionally explore similarity due to shared absence of common variables. Our kernel approach is also related to work that applies different functions to different types of input variables, e.g., ordinal versus nominal,[220,221] and to work on learning composite kernels as a structure discovery problem.[222] We have not found previous work considering both presence and absence of input variables for similarity assessments within an indicator dataset.

5.4 Methods

We address supervised learning tasks where the outcome of interest $y(o)$ for a particular observation o can be explained partly by constructed features $\boldsymbol{\phi}(o)$ (i.e. scalar or fixed-length vectors that represent a characteristic or property of o) and partly by more complex information $\Psi(o)$ (e.g., high-dimensional, time-varying, variable-length data). For example, in a primary health care setting each observation could be a client and the outcome of interest a condition or situation, such as diabetes or food insecurity, that the client is at risk for and early intervention may help to prevent. In this setting, $\phi(o)$-type information may include sociodemographic characteristics and core diagnoses while $\Psi(o)$-type information may include years of encounter data representing the subset of thousands of possible tests, diagnoses, and procedures that the client has received in their lifetime.

Additional technical background on feature- and similarity-based approaches was provided at the beginning of this Chapter. These two approaches can also be contrasted from a clinical standpoint, whereby there are often documented risk factors for a given outcome that can be assessed for each individual client; however, clinicians may also assess clients by thinking about similar clients they have previously cared for. When

developing a model, similarity-based approaches may be advantageous over feature-based approaches when there are a large number of data elements or characteristics to consider and/or when the proper way to enter data into a model is unknown. Example applications include using kernel methods to capture genomic similarity[223,224] or to improve-upon threshold-based alerts for intracranial hypertension in the Intensive Care Unit.[225]

5.4.1 The hybrid feature- and similarity-based model

The proposed *HFSM* combines a feature-based component and a similarity-based component with an additive model structure. Prediction $\hat{y}(o)$ for observation o is given by

$$\hat{y}(o) = h\left(\sum_j \phi_j(o)\beta_j + \sum_i \alpha_i \sum_l k_l(\Psi_l(o), \Psi_l(o_i))\right) = h\left(\boldsymbol{\phi}(o)^T\boldsymbol{\beta} + \boldsymbol{k}(o)^T\boldsymbol{\alpha}\right)$$

where j indexes the features; i indexes the observations or clients in the training data; l indexes the kernel domains, if there are multiple; and h is a monotonic function, e.g., sigmoid or identity. $\boldsymbol{k}(o)$ is the vector of kernel values between o and each training data point. All analyses in this paper use the sigmoid function to estimate the probability of a binary outcome occurring. Thus, the estimated probability of an outcome occurring for o is based on 1) their feature values and the corresponding coefficients (β_j) and 2) similarity to clients from the training data and the overall influence (α_i) of each client.

To train the model, we optimize a penalized log likelihood training criterion given by

$$LL(\boldsymbol{\beta}, \boldsymbol{\alpha}; \lambda) = \left(\sum_i y_i(\boldsymbol{\phi}(o_i)^\mathsf{T}\boldsymbol{\beta} + \boldsymbol{k}(o_i)^\mathsf{T}\boldsymbol{\alpha}) - \log(1 + e^{\boldsymbol{\phi}(o_i)^\mathsf{T}\boldsymbol{\beta} + \boldsymbol{k}(o_i)^\mathsf{T}\boldsymbol{\alpha}})\right)/n - \lambda||\boldsymbol{\alpha}||_1.$$

The L1-penalty on $\boldsymbol{\alpha}$ controls overfitting and produces a sparse model whose kernel component only depends on a subset of the training data; this is different from the original kernel logistic regression formulation which penalizes the norm of the regression function in its Hilbert space but does not induce sparsity.[226] Training o_i

that maintain non-zero α can be thought of as "representatives" for groups of similar clients. We solve this problem using the convex programming language `cvxpy` in Python.[227,228] An illustrative example relating the hybrid model to a special case of kernel logistic regression and Python code for *HFSM* are provided in Appendices E.1 and E.2, respectively. Required memory for model fitting, assuming n clients and m features, is $\mathcal{O}[mn+n^2]$ for *HFSM* as compared to $\mathcal{O}[mn]$ for *LR* and $\mathcal{O}[n^2]$ for kernel logistic regression (*KLR*). Solve time will be compared in the experiments.

5.4.1.1 Fitting and interpretation

We consider two variations on fitting *HFSM* that have different interpretations: 1) *HFSM-Sequential* (*HFSM-Seq*), which learns the feature coefficients fixing $\boldsymbol{\alpha} = 0$ and then fixes the learned feature coefficients while learning the kernel coefficients, and 2) *HFSM-Simultaneous* (*HFSM-Sim*), which optimizes the feature and kernel coefficients jointly. The simultaneous model fit is expected to result in better predictive performance since there is more flexibility to maximize the objective function, but the resulting model has a more complex causal interpretation. In *HFSM-Seq*, the feature coefficients represent their impact on the outcome adjusted for all of the other features in the model but *averaged* over the information in the kernel, whereas *HFSM-Sim* feature coefficients are additionally *adjusted* for the information in the kernel. We discuss the implications for interpretation below. If the feature and kernel matrices are orthogonal, the models produced by the two procedures will be identical.

A series of illustrative examples contrast the performance and interpretation of *HFSM-Seq* and *HFSM-Sim* in terms of causal inference. For each example there was a binary outcome y, one continuous feature $X_1 \sim N(0,1)$ that maintained a direct relationship with y, and a binary feature X_2 whose relationship with y and K was manipulated. For simplicity, K was unpenalized and constructed from a linear kernel function applied to a single binary variable. We designed four examples, represented in Figure 5.1:

1. **Independent contributions**. $P(Y) = \varsigma(0.25 - 1X_1 + 2X_2 + 3K)$ where $X_2 \sim$

$B(p = 0.5)$, and $K \sim B(p = 0.5)$.

2. The kernel operated as a **confounder** between the second feature and the outcome. $P(Y) = \varsigma(0.25 - 1X_1 + 3K)$ where $K \sim B(p = 0.5)$ and $P(X_2) = \varsigma(2K)$.
3. The kernel operated as a **collider** between the outcome and on the second feature. $P(Y) = \varsigma(0.25 - 1X_1)$ where $P(K) = \varsigma(3Y + 2X_2)$ and $X_2 \sim B(p = 0.5)$.
4. The kernel operated as a **mediator** between the second feature and the outcome. $P(Y) = \varsigma(0.25 - 1X_1 + 3K)$ where $P(K) = \varsigma(2X_2)$ and $X_2 \sim B(p = 0.5)$.

For each example, feature coefficients were compared for *HFSM-Seq*, *HFSM-Sim*, and *LR* fit on 3,000 training examples and predictive performance was compared based on AUROC for 1,000 new test examples.

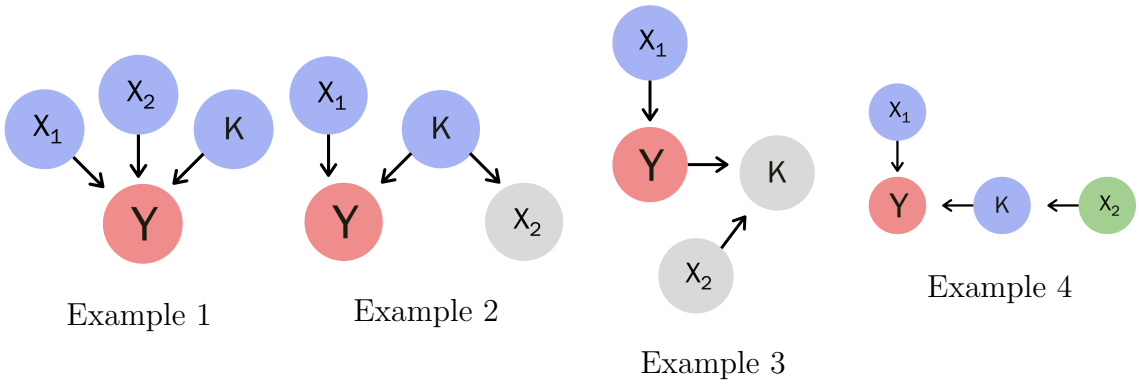


Figure 5.1: Data generating mechanisms used to contrast sequential and simultaneous hybrid model optimization.

As seen in Table 5.3a, *HFSM-Sim* had the best predictive performance for all examples; however, as seen in Table 5.3b the corresponding *HFSM-Sim* feature coefficient estimates could be closer to the truth, further from the truth, or similar to the feature coefficients learned in *HFSM-Seq*. While the impacts of adjusting the feature coefficients by the kernel were predictable for these simple experiments, in practice the direction of bias, if any, may be hard to determine. This uncertainty is analogous to situations with solely feature-based approaches where the relationships between the features and the outcome are unknown, or when automatic feature selection methods

are used.[229–232]

Table 5.3: AUROCs and coefficients for interpretation example.

(a) Test AUROCs

	LR	HFSM-Seq	HFSM-Sim
Ex. 1	0.758	0.860	0.863
Ex. 2	0.721	0.834	0.854
Ex. 3	0.777	0.800	0.845
Ex. 4	0.739	0.852	0.888

(b) Coefficients

	HFSM-Seq	HFSM-Sim	True
Ex. 1: Independent Contributions			
β_0	1.345	0.395	0.25
β_1	-0.796	-0.995	-1.00
β_2	1.515	1.896	2.00
Ex. 2: K was a Confounder			
β_0	0.595	0.264	0.25
β_1	-0.723	-0.960	-1.00
β_2	1.044	0.092	0.00
Ex. 3: K was a Collider			
β_0	0.382	-2.325	0.25
β_1	-0.950	-0.979	-1.00
β_2	-0.020	-0.578	0.00
Ex. 4: K was a Mediator			
β_0	1.213	0.295	0.25
β_1	-0.793	-0.984	-1.00
β_2	1.246	0.158	2.00

Note: HFSM-Seq and logistic regression coefficients are equal. *Legend*: HFSM-Seq = Hybrid Model - Sequential Fit, HFSM-Sim = Hybrid Model - Simultaneous Fit, True = Coefficients used to generate the data.

An interpretation advantage of the *HFSM-Sim* approach over solely feature-based approaches is the opportunity to adjust for more complex types of confounding information than can be adequately captured through features. When the information captured by the kernel is uncertain, *HFSM-Seq* can be used to maintain straightforward feature coefficient interpretation while still improving the absolute risk prediction through the addition of the kernel. Thus, hybrid models may be used much in the

same way that logistic regression can be used for prediction or inference depending on whether one focuses on the predicted outcome or model coefficients, respectively, with the advantage of using a kernel to account for additional variability. Closeness to the truth needs to be assessed on a model-by-model basis.

The kernel coefficients $\boldsymbol{\alpha}$ may also be informative in and of themselves. Whereas applying an L1-penalty to features is a form of feature selection, applying an L1-penalty to $\boldsymbol{\alpha}$ selects "representative observations" to include while adjusting for the features. The higher the penalty, the fewer observations are allowed. The most influential clients in the training data (highest magnitude α) can be investigated to explore kernel behaviour. A prediction for an individual client is based on their feature values and corresponding β, and then will be further increased or decreased depending on similarity in terms of the kernel to clients in the training data that have non-zero α. Similarity to clients with positive α will increase the predicted probability while similarity to clients with negative α will decrease the predicted probability of the outcome.

5.5 Evaluation

We compared the performance of the *HFSM* approach to solely feature- or similarity-based approaches with 1) a simulation study of three synthetic data scenarios where the relative importance of the feature- and kernel-based data was varied and 2) a clinical case study with EHRs from a primary health care organization in Ontario.

5.5.1 Simulation study

This study compared the *HFSM* approaches to the two most direct sub-component models as in an ablation study. We followed the ADEMP framework for planning and reporting on this study.[233] The *data generating mechanism* was based on a parametric model that most closely corresponds to Example 1 (independent contributions) in the illustrative examples above, with four binary features and additional complex

information that cannot be well represented by features in a linear model. This latter information was based on the classic Monk-1 data problem, which includes 6 categorical variables (two-level variables: $a3$, $a6$; three-level variables: $a1$, $a2$, $a4$; four-level variable: $a5$) and an outcome M that is the result of the Boolean statement $(a1 = a2) \vee (a5 = 1)$.[234] The outcome was used in the data generating mechanism. For models with a similarity-based component, the RBF kernel function was applied to the six categorical variables. We selected the RBF kernel due to its popularity in machine learning applications and its use in previous work with Monk's data problems.[219,235]

The data generating mechanism was $P(Y) = \varsigma(\beta_0 + 0.3X_1 + 0.4X_2 + 0.6X_3 + 0.7X_4 + \delta M)$. Coefficients were decided such that if $\beta_0 = 0$, $\delta = 0$, and $\sum_{m=1}^{4} \beta_m \tilde{=} 2$ then $P(Y)$ ranges from 0 to 0.88 and δ can be used to further increase the maximum probability of the outcome. Across the three experiments the β were fixed and δ was changed to vary the relative importance of the feature- and kernel-based data. The intercept β_0 was used to bring the prevalence of the outcome below 50% to be more similar to most clinical outcomes. Three scenarios were set up with 10,000 observations generated from each, which is similar to the number of clients expected across a few small primary health care clinics:

1. Kernel had a similar effect to a single feature: $\delta = \text{mean}(\boldsymbol{\beta})$ and $\beta_0 = -1.5$
2. Kernel had a similar effect to the set of features: $\delta = \text{sum}(\boldsymbol{\beta})$ and $\beta_0 = -2.1$
3. Kernel had a larger effect than the set of features: $\delta = 2 \cdot \text{sum}(\boldsymbol{\beta})$ and $\beta_0 = -3.2$

For each scenario, we implemented a nested CV procedure whereby for each of five outer folds, the outer fold training data were split 75/25 into inner fold training and validation data. To reduce random variation between the models, the same outer and inner CV folds were used for each model; seeds were re-set between scenarios. Hyperparameters, if any, were selected through a grid search for the best AUROC on the inner validation data. Models were then re-trained with the selected hyperparameter(s) on all outer fold training data, and predictions of the target binary outcome were made on the outer fold test data. Folds were trained in parallel using the python package `multiprocessing`.[236,237]

Four models were compared: 1. Feature only model: *LR*, 2. Similarity only model: *KLR*, 3. *HFSM-Seq*, and 4. *HFSM-Sim*. For kernel-containing models, three candidate hyperparameters σ for the RBF kernel were considered (0.01, 0.1, 1). These values provided a range of similarity patterns on the Monk's data, based on a measure of matrix diagonal dominance and visual exploration of the RBF kernel calculated for a random sample of 1000 observations (Appendix E.3.1). For each σ, five candidate values for the L1-regularization strength λ on the kernel coefficients, ranging from 0.001 to 1, were considered.

Model performance was compared using measures averaged across the five outer test folds. The primary metric of interest was discrimination, assessed through AUROC. Secondary metrics of interest included AUPRC, calibration plot slopes and intercepts, and time to re-train the model with selected hyperparameters. Predictive performance metrics were also calculated for a "best possible model" that made predictions based on applying the known coefficients to all data. Hyperparameters and parameters were viewed and compared between models.

The expected trends emerged across the three scenarios: the hybrid models always performed similar to or better than the single component models, with a notable advantage for the second scenario (Table 5.4). Selected hyperparameters and learned parameters are in Appendix E.3.2. In the two extreme scenarios, the hybrid models performed similarly to whichever single-component model captured the more important portion of the data. In the "feature heavy" scenario 1, *HFSM* performance came with an increase in computation time (seconds vs. hours) as compared to *LR*. In the "kernel heavy" scenario 3, the discrimination performance of *LR* approached a dummy classifier while *KLR* showed similar predictive performance to the hybrid models and increased fitting time as compared to *HFSM-Sim*. In the "middle ground" scenario 2, *HFSM* demonstrated the best discrimination and precision-recall performance, but neither *HFSM* version outperformed *LR* in terms of calibration. These findings show the advantage of using *HFSM* when both feature- and kernel-based data are important and when there is uncertainty about their relative importance. For all models with

a feature component, as the relative importance of the kernel-based data increased, the feature coefficient estimates got further from the truth (Appendix E.3.2).

Table 5.4: Synthetic data study results.

	LR	KLR	HFSM-Seq	HFSM-Sim	Best
Scenario 1: Kernel data had similar effect to a single feature					
AUROC	0.647	0.504	0.648	0.647	0.655
AUPRC	0.571	0.436	0.572	0.573	0.581
Calibration Slope	-0.025	-0.362	-0.032	-0.013	-0.008
Calibration Intercept	1.035	-0.517	1.038	1.031	0.989
Time (hours)	< 1	6.553	7.815	7.177	
Scenario 2: Kernel data had similar effect to the set of features					
AUROC	0.614	0.712	0.725	0.726	0.781
AUPRC	0.587	0.672	0.708	0.710	0.759
Calibration Slope	-0.001	0.043	0.044	0.008	0.027
Calibration Intercept	0.993	1.415	1.305	1.257	1.017
Time (hours)	0.001	9.901	10.574	9.316	
Scenario 3: Kernel data had a larger effect than the set of features					
AUROC	0.558	0.872	0.877	0.877	0.903
AUPRC	0.538	0.825	0.840	0.846	0.879
Calibration Slope	-0.001	0.010	0.012	-0.058	0.018
Calibration Intercept	0.980	1.557	1.575	1.534	1.018
Time (hours)	< 1	7.148	7.419	5.111	

Note: AUPRC = Area Under Precision Recall Curve, AUROC = Area Under Receiver Operating Characteristic Curve, Best = Hardcoded true coefficients applied to all data, HFSM-Seq = Hybrid Model - Sequential Fit, HFSM-Sim = Hybrid Model - Simultaneous Fit, KLR = Kernel logistic regression, LR = Logistic Regression.

5.5.2 Methods for clinical case study

We present a case study with EHR data from the Alliance for Healthier Communities, which provides inter-professional, team-based primary health care through Community Health Centres (CHCs) across Ontario, Canada.[9,238] All CHCs record standardized sociodemographical information (e.g., birth date, education, household income) and appointment details (e.g., care provider type, diagnosis codes) in a centralized, structured EHR database. We used de-identified data from January 1, 2009 to December 31, 2019 to predict two-year risk of first incidence loneliness or social isolation for middle-aged clients being served by the "urban-at-risk" (UAR) peer group

of CHCs. This subgroup of CHCs provides care to clients with pre-existing substance use, homelessness, or mental health challenges. This study was approved by Western University ethics board (project ID 111353).

Although this case study was primarily intended to test the proposed methods, we selected this outcome because it aligns with the Alliance and it is increasingly recognized as a serious health challenge in the literature. Research has largely focused on sequelae and comorbidities in older adults, finding associations with several other poor health outcomes.[239–246] There are a range of services and programs offered through CHCs that may help mitigate the risk of social isolation and loneliness, such as their social prescribing initiatives.[247–249] Identifying people at risk of social isolation or loneliness may provide an opportunity for early intervention; this case study did not result in a model ready for deployment, but substantive findings may inform future work in the area.

5.5.2.1 Cohort

The cohort of interest included ongoing primary care clients at UAR CHCs without the outcome at baseline. To restrict the sample to new or newly returning, mid- to long-term clients, only those whose first event was recorded in 2010 or later and who had at least one event three years from the first recorded event were eligible for inclusion. Primary health care is provided at all stages of life and health and social isolation or loneliness may occur at any point, so we randomly selected two-year periods from each client's observation history to serve as the prediction interval. Feature and kernel input data used to make predictions were from the first recorded event to the beginning of the randomly selected prediction interval. The start of the prediction interval had to be at least one year from the first recorded event as the first year of care provision in this population is associated with a distinct risk profile, likely due to "catch-up" on unresolved care and diagnoses.[250] We restricted our cohort to those 45-64 years old at the end of their baseline period as age is associated with the outcome and may influence the risk factors and potential interventions to help

someone at high-risk.

5.5.2.2 Feature choices

We identified 19 candidate features based on evidence in the literature,[239–242,244,245,251] perceived importance with input from Alliance stakeholders, and feasibility to construct with available data. Features from the client characteristic table were handled with complete case analysis if under 1% missingness and with a missingness indicator approach otherwise for 1) client was asked the question and preferred not to respond and 2) client was never asked. Features constructed using International Classification of Disease (ICD-10)[176] and Electronic Nomenclature and Classification Of Disorders and Encounters for Family Medicine (ENCODE-FM)[175] vocabularies were assumed absent if no appropriate codes were present during baseline. Three of these features had under 1% prevalence in baseline data and were excluded from the model (people with the features present were not excluded from the cohort); we performed indirect standardization to assess sex-adjusted risk in associated subpopulations.

We also constructed a feature to represent general clinical complexity as the count of the number of chronic conditions identified as important for multimorbidity research in primary care present during baseline,[179] scaled to 0,1 range. This type of non-specific, complex information was what we designed the kernels to capture; this composite feature represented what we may try to include instead for a solely feature-based model.

5.5.2.3 Kernel choices

In addition to the specific conditions identified for features, there is a sense that general health complexity may be positively associated with the outcome. We used three types of *kernel input data* based on appointment-associated care characteristics to capture this additional complex information: 1) the provider type(s) involved in care (e.g., nurse practitioner, social worker), 2) the service type(s) provided during an appointment, which represents the general type of care functions provided (e.g.,

assessment, treatment management) without specifying conditions, and 3) both 1 and 2. There are many ways data could be pre-processed and combined for kernel inputs; we worked with sets and added together distinct codes experienced at least once during baseline care.

A valid *kernel function* must be symmetric and result in a positive semi-definite kernel matrix. Additional properties that we wanted our kernel function to have include:

1. Holding all else constant, two clients who both have or do not have a specific code should be more similar than when only one of them has the code present.
2. Two clients who do not have a code that is common in the population of interest should be more similar than two people who both have the common code present.
3. Two clients who have a rare code present should be more similar than if they did not, but sharing in the absence of rare codes should not have a large impact on similarity.

We developed kernel functions based on Gower's (1971) work on the coefficient of similarity.[252] The similarity between two individuals i and j based on character c can be assigned the similarity score $S_{i,j,c}$ ranging from 0 (no similarity) to 1 (the same). An indicator $\delta_{i,j,c}$ is used to represent whether or not a comparison can be made. In the case of a binary variable, if one or both people have the variable present the indicator is 1; if neither person has the variable present the indicator is 0 and $S_{i,j,c}$ is set to 0.[252] Gower (1971) further demonstrated that a weight can be introduced for each code w_c; if there are no missing values and all $w_c \geq 0$ the following allows for a positive semidefinite similarity matrix with entries $S_{i,j} = \sum_{c=1}^{v} S_{i,j,c} w_c / \sum_{c=1}^{v} \delta_{i,j,c} w_c$.

The commonly-used J similarity is equivalent to setting $w_c = 1$ for all c. In the case where both people's sets are empty, we set $S_{i,j}$ to 1. It meets the first of our desired properties and is the first candidate kernel function we used. The second property can be addressed by reverse coding common data elements and assigning weights such that only common codes are considered; the third property is addressed by maintaining traditional coding based on presence and setting weights such that

only rare codes are considered. We used a cut-off based on prevalence in the training data to define common (prevalence ≥ 0.70) and rare (prevalence < 0.30) codes such that codes above/below the threshold are assigned $w_c = 1$ and remaining codes are assigned $w_c = 0$. Our second candidate kernel function (SCR) adds together the "common absence" and "rare presence" similarity scores.

5.5.2.4 Application 1: Prediction

We assessed predictive performance using a similar nested CV procedure as was used for the simulation studies, with 80/20 splits to define inner training/validation data for each of five outer folds. The same four models of interest (LR, KLR, HFSM-Seq, HFSM-Sim) were compared alongside two additional models: *LR-E*, which was *LR* with the extra count of chronic conditions feature; and a more complex model (*XGBoost*) that included all features and all kernel data input represented as dummy variables. Hyperparameters for kernel-containing models selected based on a grid search for the highest AUROC on the inner fold validation data included L1 penalty strength (0.0001, 0.001, or 0.01), kernel data inputs (providers involved, service types, or both), and kernel function (J or SCR).

5.5.2.5 Application 2: Interpretation

To demonstrate model interpretability, we re-trained *HFSM-Seq* and *HFSM-Sim* on all data using the Jaccard kernel function on both types of data, using the mode of the selected L1 penalty in Application 1 divided by five to scale for the increase in amount of data. We examined whether feature coefficients changed between the two models similar to in the illustrative examples. We then moved our focus to *HFSM-Seq* to examine the type of information captured by the kernel after accounting for the features. We split the cohort into clients with positive, negative, and zero-valued kernel coefficients. Feature-based characteristics were compared with descriptive table-based summaries across the three strata. Kernel-based characteristics were explored by applying non-negative matrix factorization with five topics, using Python package `sklearn.decomposition.NMF` and the Kullback-Leibler divergence distance

metric, to each of the three strata.[185]

5.5.3 Results for clinical case study

There were 5,070 eligible clients with a 5.4% cumulative incidence (n=276) of the outcome across all client-specific two-year prediction intervals. See Appendix Figure E.2 for a cohort flow diagram and Appendix Table E.8 for select baseline characteristics.

5.5.3.1 Application 1: Prediction

Performance metrics are in Table 5.5. The SCR kernel function was selected three times for HFSM-Seq and KLR, and two times for HFSM-Sim; for all models the combined provider and service type data were selected the majority of the time (Appendix Table E.9).

Table 5.5: Clinical case sudy predictive performance results.

	LR	LR-E	KLR	HFSM-Seq	HFSM-Sim	XGBoost
AUROC	0.753	0.754	0.734	0.774	0.778	0.727
AUPRC	0.146	0.148	0.139	0.185	0.184	0.137
Calibration Slope	0.852	0.848	0.698	0.788	0.875	0.868
Calibration Intercept	-0.367	-0.378	-0.788	-0.521	-0.294	-0.621
Time (minutes)	< 1	< 1	42	115	89	< 1

Note: Results were averaged across the five outer folds. *Legend*: LR = Logistic Regression; LR-E = Logistic Regression-Extra Clinical; KLR = Kernel Logistic Regression; HFSM-Seq = Hybrid Feature- and Similarity-based Model-Sequential; HFSM-Sim = HFSM-Simultaneous.

General trends across predictive performance metrics from worst to best were *KLR* and *XGBoost*, *LR*, *HFSM-Seq*, and then *HFSM-Sim*. For discrimination and precision-recall performance, *HFSM-Seq* and *HFSM-Sim* were best. Calibration was best for *HFSM-Sim*; all models tended to overestimate risk. While there were some instances of a model having notably worse performance in terms of calibration, there were no instances of one model that had very large performance gains over all other models; discrimination performance was comparable across all models. *LR* and *LR-E* performed similarly on all metrics. Kernel containing models were the least efficient, even with pre-computed kernel matrices, but still ran within a feasible amount of

time.

5.5.3.2 Application 2: Interpretation

In general, *HFSM-Seq* feature coefficients (Appendix Table E.10) were larger in magnitude and consistent in direction to those of *HFSM-Sim*, suggesting that the kernel information adjusted for some of the feature relationships. We suspect that for some of the coefficients with the largest change in magnitude the kernel (capturing health care use information) was acting as a mediator (e.g., stable housing), and for others it was acting as a confounder (e.g., depression or anxiety). This or colliding bias could also explain features where associated coefficients tended to increase in magnitude when adjusted for the kernel (e.g., primary language) and features where there was a qualitative change (e.g., food insecurity; coefficient switched from positive to negative after adjusting for the kernel). Importantly, we did not set up this model to be causal and do not know which feature coefficients are closer to the truth; we would not want to deploy it in its current form for clinical decision making.

Examining *HFSM-Seq*, α coefficients ranged from -1.70 to 1.30 and when rounded to five significant digits there were 5,038 zero, 13 positive, and 19 negative. Feature and outcome values stratified across these three groups are in Appendix Table E.11. Distinct trends for clients with negative α, and thus decreased the predicted probability for similar clients, were that none had the outcome, all lived in an urban geography, and they tended to have higher levels of obesity than the other strata. Clients with positive α, and thus increased the predicted probability for similar clients, all had English as their primary language and tended to have lower household income and higher levels of stable housing, substance use, smoking or tobacco use, and food insecurity. The top ten weighted codes (from provider type and service type data, coded for presence) from NMF for the three subgroups are in Appendix E.4.3. The group with negative α's had a unique topic characterized by diagnosis and treatment with physician and nurse providers; a topic related to counselling and foot care with counsellor and chiropodists; and one related to counselling with nurse practitioners. The

group with positive α had a unique topic strongly characterized by external referral and consult; and a topic strongly characterized by social worker, nurse practitioner, and individual counselling. Codes related to diagnosis, treatment, and management were not as prominent in topics as for the other groups. The zero α topics included one strongly characterized by community resources and community health workers, which only weakly entered topics for the other subgroups.

5.6 Discussion

The *HFSM* approach captures relationships within large-scale observational data in an interpretable form when some but not all data and desired information to capture are suitable for simple feature representation. Simulation studies confirmed that *HFSM* is best suited for situations where the feature- and kernel-based data are both important for the outcome, and our clinical case study demonstrated how it can be used to build a predictive model and develop understanding of risk drivers within a complex primary health care population. Of note, while our case study was situated in primary health care, the methods are applicable to other sectors as well.

5.6.1 Hybrid model methodology

The predictive performance of *HFSM-Sim* is always expected to be as good or better than *HFSM-Seq*, assuming appropriate set up and tuning, while *HFSM-Seq* provides more certainty in feature coefficient interpretation when the role of the kernel in terms of causal structures is uncertain. If the goal is to prioritize absolute risk predictions, *HFSM-Sim* is recommended; however, if the model is intended to support decision making with interpretation of feature coefficients, then greater care is needed. Feature coefficients learned under *HFSM-Seq* as shown in this paper are adjusted for each other and averaged over the kernel, so can be interpreted analogously as for *LR*. An additional option is to fix some or all of the feature coefficients based on previous research studies or epidemiological analyses, and learn the rest from the training data. The feature coefficients for *HFSM-Sim* are adjusted for each other and adjusted

for the kernel matrix, which may be favourable for interpretation in situations such as when the kernel is constructed to adjust for complex confounders. Frameworks commonly used in clinical epidemiology studies, such as directed acyclic graphs, may help guide model development.[253] The degree to which the kernel and feature matrices are independent will determine the difference between *HFSM-Seq* and *HFSM-Sim*.

The Illustrative Examples demonstrating feature coefficient changes depending on the "causal role" of the kernel reinforce the importance of careful modelling not only for explicit causal inference but also when a risk prediction model might be interpreted as an "upstream" treatment effect model. For example, if feature coefficients are interpreted as identifying modifiable risk factors (e.g., hypertension, smoking) important for a client's estimated outcome, and then inform risk prevention strategy selection. Formal techniques for multiple causal inference, such as the deconfounder approach informed by a directed acyclic graph based on clinical and epidemiological input, may be useful here.[65,254,255] Future work is needed to determine what "pragmatic" level of causality is sufficient to support decisions in these settings. Avenues for future work on the hybrid model structure include interactions between features and the kernel, other outcomes types (e.g., time to event), multilevel modelling, and adding an L2 penalty to the kernel coefficients. The most closely related work to the latter applies an elastic net penalty to the dual form of the problem only.[256]

5.6.2 Clinical case study

Predicting the rare outcome of social isolation or loneliness in middle-aged clients served by UAR CHCs is a challenging supervised learning problem. The *HFSM* models performed as well or better than solely feature- and similaritly-based models, including `XGboost`, while providing superior interpretability. Our proposed kernel function that calculates similarity based on the absence of common codes and the presence of rare codes was selected more often than Jaccard, which considers presence of all codes. We used a basic cut-off to define rare and common codes, but future work could expand this to obtain more sophisticated weightings, such as with a probability

mass function on the training data,[219] and to explore methods to identify the ideal switching point between presence and absence coding.

We explored model behaviour and learned about the cohort by viewing feature values and NMF-derived topics on kernel data stratified by positive, negative, or zero α. Of note, the kernel data included provider types; some, e.g., social work, may have a higher index of suspicion for the outcome and care for correlated conditions, but these data are only from baseline and all providers could code the outcome. We could extend this work into the causal setting and intentionally set up the model to use the features and kernel to capture relationships between some or all of the features and the outcome. Future work could compare coefficients from *HFSM* that e.g., uses to kernel to adjust for complex potential confounders, against other approaches or known effect estimates.

In addition to insights about *HFSM* and the kernel, our clinical case study demonstrates insights relevant to future primary health care machine learning applications. In contrast to settings where care is initiated due to a problem (e.g., cancer diagnosis, emergency room visit), primary health care is sought out during all stages of health, there is variability in visit patterns, and risk patterns change across across the care history due to cumulative and acute factors.[250,257] Outcomes such as ours are relevant across the entire care trajectory, which induces challenges for determining a prediction interval as "lifetime risk" is unhelpful. Future research is needed on the best way to define prediction intervals for these types of outcomes in primary health care; we selected two-year prediction interval periods to support generalizability of the resulting model across the care continuum within a time frame that allows preventative intervention, but the most appropriate choice will depend on context.

A significant challenge we encountered, which is relevant to other health sectors, is rare features. Three features (Sensory Disability, Social Phobia, Dementia or Alzheimer's Disease) representing characteristics that literature suggests are risk factors for social isolation and loneliness had less than 1% prevalence in the baseline cohort data. The standardized morbidity ratios (SMR), representing the ratio of

observed to expected number of outcome cases based on sex-specific rates in the remaining eligible population, showed higher than expected risk in each of the rare feature sub-populations (Sensory Disability SMR = 2.44; Social Phobia SMR = 3.27, Dementia or Alzheimer's Disease SMR = 2.11). The number of clients with these characteristics was too small to meaningfully do statistics with, and we could not find an explainable AI framework that addressed this type of scenario. If the model were implemented, it would be important to communicate to care providers that these risk factors are not considered by the model, such as with a flag when making predictions for a client with one of the characteristics present. Qualitative research could investigate risks and needs of these subpopulations specific to the health care setting of implementation.

5.6.3 Limitations

We have not provided confidence intervals or hypothesis tests particularly in the case of *HFSM-Sim* because although the objective function is convex, the non-smoothness of the L1 penalty is expected to require use of techniques like the m-out-of-n bootstrap[258] or potentially a selective inference framework[259] to account for non-regularity in the estimators. Developing these is beyond the scope of this work.

In our clinical case study, we restricted based on age and CHCs within the UAR peer group; however, remaining variability within these strata was not taken into account. Some feature construction was based on the client characteristic table, which included rich sociodemographic information but was not time stamped in our data extraction. Outcome recording was not blinded and can only be considered a proxy for "true" social isolation and loneliness. The majority of clients excluded for having less than three years of observation had their first event in 2017 or later so there was not enough calendar time for sufficient observation. There were 1,430 clients with a first event early enough and who met other eligibility criteria. If we were to proceed with this model we would perform sensitivity analyses to assess whether there is bias due to

their "true" loss to follow up.

5.7 Conclusion

We presented a hybrid feature- and similarity-based model that combines well-established approaches (LR and KLR) into a single machine learning model. The hybrid approach provides a way to take advantage of large-scale datasets information about features where the relationship with the outcome can be specified in a linear model (e.g., known informative risk factors or structured one-time question fields), as well as more complex data that may be better captured in terms of similarity to other training examples (e.g., historical data on care and diagnoses received). Maintaining separation of feature and similarity based components supports interpretability of the final model, and the option to fix or learn feature coefficients in advance of the kernel coefficients provides additional flexibility over feature coefficient interpretation. The inherent model interpretability and the reproducibility due to a fully convex objective function supports the extension of model use from prediction to causal inference tasks both within health care and in other domains with complex data and causal structures.

Chapter 6

Discussion

This thesis integrated epidemiology and computer science through three research studies, each of which built upon the other: first, a summary of the state of artificial intelligence (AI) and primary care research at the outset of this body of work; second, the first large-scale descriptive and exploratory study of adult primary care clients served by the Alliance for Healthier Communities; and third, proposed supervised machine learning methodology that combines feature and kernel learning for prediction and interpretation. The major contributions of this work can be organized by those primarily to computer science and to epidemiology, noting that altogether the work constitutes an interdisciplinary contribution to primary care:

To **computer science** and more specifically the subfield of machine learning, we contributed 1) the two forms of the hybrid feature- and similarity-based model (HFSM), which can be extended to any setting with large-scale observational data; 2) a new framework for thinking about kernel based similarity in terms of rare and common characteristics; and 3) demonstration of how techniques frequently used in epidemiology (scoping review and population level descriptions) can help select and inform problems or projects that involve AI, including development of decision support tools with electronic health record (EHR) data.

To **epidemiology** and more specifically the subfield of clinical epidemiology, we con-

tributed 4) the first comprehensive overview of AI and primary care research; 5) an extensive description of sociodemographic, clinical, and healthcare use characteristics of adult primary care clients served by Community Health Centres (CHCs) in Ontario from 2009 through 2019; 6) demonstration of how unsupervised machine learning techniques can be used in this type of population-level exploration of a complex population, and demonstration of how the hybrid model can be used to investigate causal relationships or further explore population characteristics within a supervised learning task.

Throughout the body of work three themes emerged. First, *AI for primary care is at an early stage of maturity, but progressing*, both in terms of the field in general and in terms of individual primary care systems being able to harness value from their data. Second, *primary care is complex in ways that are unique within the healthcare system and requires innovative, careful approaches to study design and methods.* Third, *epidemiology and computer science are complementary fields*, as seen in our choice of methods and approaches for various problems and sub-problems. Each theme is elaborated on in turn before discussing future directions and final conclusions.

6.1 Summary of Major Themes

6.1.1 Artificial intelligence for primary care is at an early stage of maturity, but progressing

Primary care is understudied in comparison to other sectors of health care and AI is no exception.[174,260–262] The scoping review in Chapter 3 summarized all research until 2018 that included AI and primary care, finding that the field was at an early stage of maturity in terms of widespread adoption of AI in practice, with few examples of studies in "real-world" primary care settings, and not enough interdisciplinary collaboration. Since then progress has been made. There are more and more examples of research with genuine primary care involvement, such as a study by Wingrove et al. (2020) to develop and test a machine learning model to predict family medicine

specialty from Medicare data;[263] or the use of machine learning to predict frailty from the Canadian Primary Care Sentinel Surveillance Network database by Aponte-Hao et al. (2021).[264] Another type of progress has been increased knowledge about organizations, such as the Alliance for Healthier Communities, that are motivated to use their data to improve care; we demonstrate steps taken to inform and support success of future AI-related initiatives in these types of settings.

In addition to an increase in activity of primary care communities with respect to AI research in general, work has been conducted that provides additional direction to the field. A scoping review by Rahimi et al. (2021) took a deeper dive into AI applications that have been tested or deployed in community practice settings up to 2020, highlighting gaps to address in future evaluation research;[265] Ronquillo et al. (2021) established the Nursing and AI Leadership (NAIL) Collaborative and held an international think-tank to identify implications, opportunities, and needs in AI for nursing;[266] Yang et al. (2022) held a two day virtual meeting that resulted in the "IDEAS" framework outlining major domains for AI integration in primary care in the United States;[267] and Kueper et al. (2022) conducted a one-day multi stakeholder event to identify priorities for AI and primary care in Ontario, Canada.[268] These more recent research studies built upon the scoping review to fill some of the knowledge gaps around AI for primary care, although in general, there is still a lot of work to be done.

Chapters 4 and 5 of this thesis make additional contributions, especially in terms of identifying methodological considerations for AI in primary care settings. We have not found other research that explicitly outlines what primary care complexity means for machine learning methods.

6.1.2 Methodology needs to account for primary care complexity

The large-scale descriptive and exploratory study in Chapter 4 showcased the rich data collected through care encounters at CHCs across Ontario, and found complex-

ity in all domains that were explored: sociodemographic characteristic explorations identified high prevalence of social determinants of health; clinical characteristic explorations highlighted the heterogeneity in co-occurrence of conditions; and care characteristic explorations found variability in care provider teams and in care access frequency. While findings were primarily intended to support work within the Alliance, they also contributed to primary care and AI research literature more broadly. First in terms of substantive findings, the population-level overview based on over a decade of care encounters at CHCs provides a unique look into client and care profiles for team-based primary care, which may be similar or different than other care models in Ontario and CHCs in other geographical regions.[9,178,262,269] Second, the associated methodological insights can inform machine learning work both within the Alliance, as exemplified through decisions made for the case study in Chapter 5, and in other primary care settings.

The methodological considerations of primary interest for this discussion are those that may be unique or exacerbated in primary care as compared to other sectors of healthcare, and especially for clients experiencing clinical and social complexity. Table 6.1 presents considerations for machine learning or decision support tool problem selection that arise due to the wide breadth or scope of team-based primary care, with examples from the thesis work. Each of the contributing primary care challenges are studied in the primary care literature outside of the context of AI, and how to best account for or select areas of focus amid heterogeneity and scope will require interdisciplinary collaborations and a clear understanding of the intended impacts of a potential machine learning project. Table 6.2 presents additional methodological challenges for prediction of future events that arise due to the provision of primary care across the life course. Each of these challenges highlights the need for more research to understand how to best develop machine learning for primary care; this research can build off of existing work and solutions from other sectors. For example, phenotyping to identify "true" client characteristics from EHRs is an active machine learning research field,[270] covariate shift is a well-established problem of study related to differences between the data used to train a machine learning model and testing

or "real life use" data,[271] and there are existing theories and techniques for outcome measure development, evaluation, and selection in clinical epidemiology that can be extended or applied to AI for healthcare settings.[272]

Table 6.1: Machine learning or decision support tool problem selection challenges due to the breadth of team-based primary care practice.

Primary care characteristic	**Machine learning considerations**	**Example from thesis work**
Heterogeneity in client characteristics and care needs, e.g., screening healthy people vs. caring for palliative or complex clients. Furthermore, within an organization certain providers or clinics may specialize towards certain populations.	Generalizability of models across populations within primary care in terms of performance and utility.	The case study in Chapter 5 to predict social isolation and loneliness was restricted to adults aged 45-64 at UAR CHCs; unsure if the trained models would be applicable in non-UAR clinics, which serve similar clients to UAR CHCs as well as other, lower-risk clients.
	Differential data collection based on person and care context.	Completeness of sociodemographic characteristic data varies due to client, provider, and clinic level factors.
Care team compositions (type(s) of provider(s), family members, etc. involved in any single client's care) may differ by client needs and/or preferences.	The type of end user for a given tool may differ by client.	NMF on provider types in Chapter 4 showed that client care was often lead by a medical doctor or nurse practioner, and may have included seeing several other types of care providers in a variety of amounts.
Similar to other sectors of healthcare, **alert fatigue and care provider burn out** has been exacerbated by EHRs.	There are hundreds of conditions cared for in primary care that machine learning could be developed for; developing separate models for each condition is not practical. Potential benefits of AI may be highest for complex clients or rare conditions, but these are more challenging settings for AI.	Chapter 4 found thousands of distinct multimorbidity compositions among Alliance clients, and the most "prominent" grouping of conditions differed depending on prioritization of frequency (count-based table) or tendency to co-occur after adjusting for other conditions (Ising model).

Table 6.2: Methodological challenges for prediction of future events related to providing care across the life course.

Primary care characteristic	**Machine learning or related methodological considerations**	**Example from thesis work**
Variable visit frequency between clients (e.g., some clients come multiple times a month; others have years between visits) and within clients across their lifetime (e.g., a single client may experience periods with frequent care).	Timescale of something truly starting to impact a client may differ from when it is recorded in their EHR.	Cumulative incidence plots in Chapter 4 showed a higher "risk" profile associated with being a new or newly returning client. It is not that being a new client increases risk for conditions; it is that being a new primary care client includes assessments and catch-up on unmet care needs and diagnoses. This may be exacerbated for clients with barriers to care access. The case study in Chapter 5 excluded "year 1" of observation history from potential prediction intervals.
Many different types of **data are collected across a potentially long period**.	Different data elements/domains may need to be treated differently, and the proper treatment/operationalization may depend on the problem. For example, childhood trauma may be operationalized as an "ever-happened" variable vs. the "number of" chronic conditions present vs. the "most recent" x-ray being more important than older ones vs. lab results being most informative when analyzed as "trends over time".	The hybrid model in Chapter 5 addresses the need for a single model to be able to take into account different types of data. The rare-common kernel function introduces the idea that the presence of some codes may be most informative whereas the opposite may be true for others.
Outcomes of interest happen on **different intervals or timescales**.	Cohort and prediction interval construction requires deciding 1) when someone starts being "at-risk", e.g., always vs. once they are X years old vs. if diagnosed with Y; and 2) what a meaningful "risk prediction" interval is, e.g., is it on the scale of days, years, number of visits.	The case study in Chapter 5 predicted first-episode social isolation and loneliness. Someone is at risk of this across their entire life course, but predicting "lifetime risk" is not meaningful. Instead, we selected two-year risk intervals across the observation history (excluding first year).

Some of these challenges additionally highlight the importance of explainable AI (XAI) for primary care. Not only is XAI important for the same reasons that it is important to understand how a model is behaving in other contexts, but it is also important because end-users may need to do an additional assessment on the relevancy or recency of the data that an EHR-embedded machine learning model is using to arrive at a prediction for a given client. Primary care clients may have long periods between appointments during which their social circumstances may have changed since the last time data were recorded in their EHR. An additional challenge not explicitly explored in our work is related to the presentation of conditions to primary care often being early-on, before symptoms and signs are clearly developed; it may take a while to arrive at a diagnosis, or problems may be resolved before a final diagnosis is reached. Distinguishing signal from noise at these early stages and arriving at performance between over- and under-diagnosis will be challenging for diagnostic tools.

In addition to methodological considerations, AI for primary care populations has sociocultural and ethical implications. The Alliance for Healthier Communities developed an evidence-informed Model of Health and Wellbeing that is used to guide care delivery and their research initiatives.[51] Example components of the model include equity and social justice, population needs-based, and care that is based on the determinants of health.[44] Tools that are developed for pilot testing and implementation in clinical settings need to maintain the Model from project outset through to long-term evaluation of whether the developed tool supports their care model and values. A learning health system (LHS) as a socio-technical system; even the best technology will be useless if it does not match the needs and values of the end-users and beneficiaries.[1,3,273] Nash et al. (2022) performed semi-structured qualitative interviews with Alliance stakeholders, finding organizational goals and culture as one of three foundational elements (in addition to data quality and resources) for the success of them becoming an LHS, with the potential to improve care as a key motivational factor.[274]

6.1.3 Epidemiology and computer science are complementary fields

This thesis relies on and contributes towards epidemiology and computer science. At the core of this intersection is the rapidly increasing amount and types of "everyday" health data being generated, the strength of epidemiology in investigating health-related questions, and advancements in computer science to support the processing, storage, and analysis of these data. The number of ways these disciplines have already and are expected to work together extends far beyond the scope of this thesis.[275–277] Nonetheless, a few examples of how both fields were integral to this body of work are highlighted below.

First, the use of descriptive epidemiology to understand a population and inform future machine learning and related data-driven projects, combined with the use of unsupervised learning techniques from computer science to do these types of descriptive studies more effectively for a complex population. While our work was situated in the context of an LHS, the value of early-stage descriptive studies is expected to apply in other settings where data-driven initiatives are being developed for known populations. Chapter 4 and the discussion above outline some of the more generalizable insights derived from this intersection, such as the need to carefully consider heterogeneity and visit patterns in primary care. Machine learning for healthcare guidelines emphasize co-development and early stage end-user engagement, as well as post hoc analyses to assess machine learning model performance across different subgroups.[30,172,190,278–280] These practices are partially motivated by the desire to produce relevant, meaningful tools that do not exacerbate inequities.[34,279] Early stage descriptive study findings can be brought into conversations with end-users and community advocates, to inform specific projects and preemptively highlight potential equity issues, as well as to track progress over time for a system-wide transition into an LHS framework. For example, we used findings from Chapter 4 to inform the clinical study in Chapter 5, such as outcome definition, predictors, and cohort construction and eligibility criteria. Of note, while our findings from Chapter 4 constitute novel

contributions, they are based on well-established epidemiology and machine learning techniques.

Descriptive epidemiology has a rich history of being integral to understanding health related states and needs of a population, such as by identifying the burden of disease to aid in resource planning for a particular geographical region, or identifying subsets of a population that have higher than expected rates of disease and may require tailored intervention. It is also a mainstay in public health practices, for example, in efforts related to the current COVID-19 pandemic, including the identification of heath inequities.[281–283] In its simplest presentation, although not necessarily simple to conduct, descriptive epidemiology includes summary counts and proportions related to a health outcome of interest for a carefully defined population.[206,281] Despite widespread utility and demonstrated value, research to advance methods for descriptive epidemiology is far less prominent than that to advance analytic epidemiology objectives, such as causal modelling. In comparison is unsupervised machine learning, which sees a larger methods research focus,[284] but there is also a need to better understand how existing, well-established techniques may advance other disciplines that are now using larger and more heterogeneous datasets. We found some of these unsupervised learning techniques to be useful in understanding the Alliance adult primary care population, e.g., non-negative matrix factorization of provider types involved in care, while others left us with outstanding questions to follow-up on, e.g., K-medoids time-series clustering with dynamic time warping distance on visit frequency. There is an opportunity for future research to further refine techniques for these types of population-level descriptions.

Moving from a descriptive to an analytic focus, the hybrid models directly addressed to the need for explainable machine learning methods with observational data such as that from healthcare. The models combined the well-established techniques of logistic regression, which is heavily used by both computer science and epidemiology, and kernel learning, which is more common in computer science. Our HFSM represents a contribution to the subfield of machine learning as well as a new tool for analytic

epidemiology, including to investigate individual causal relationships or prediction. Prediction is a core goal of supervised learning methods, whereby performance is impacted both by the specific methods involved and by the many decisions made throughout the process to get from the form of a function or algorithm to a trained model that is ready for use. In addition, to those from computer science and end-users, inputs based on epidemiology can be used to arrive at the best possible model for a given scenario. For example, the body of literature pertaining to a particular outcome could be critically analyzed to arrive at suggestions for feature and kernel based parts of the model, as well as to assess the potential causal role of a given kernel and the implications for sequential as compared to simultaneous model fitting. Causal inference from observational data is an active topic of study in both machine learning and epidemiology communities.[285–287] Another more general example of this intersection is our use of thinking from epidemiology about population-level characteristics to begin exploring the potential of extending a common computer science technique, kernel functions, into capturing similarity that is based on deviations from broader population health expectations.

6.2 Avenues for Further Study

Each integrated article included discussion about future directions specific to that study; herein we focus on additional or more general future work related to 1) AI for primary care, 2) the Alliance for Healthier Communities, and 3) methods research at the intersection of epidemiology and computer science.

6.2.1 Artificial intelligence for primary care

Despite growing recognition that there are potential benefits of AI for primary care, there are fewer examples of AI research or implementations in primary care settings as compared to other sectors. Progress may be supported by understanding what is similar and what is different about AI for primary care as compared to other sectors, both within and outside of healthcare, and by describing these challenges in terms

of specific research problems. The above discussion demonstrates how some of the unique characteristics of primary care can be broken down and described in terms of challenges for machine learning; on a fundamental level many of these challenges exist in other settings and solutions may be applied or adapted for primary care purposes. In addition to technical questions, there are other areas in need of study, such as to understand ethical, workflow, and legal implications of AI-based tools intended for long-term implementation in primary care settings.

The scoping review in Chapter 3 found a need for more interdisciplinary collaboration in AI for primary care research; future work could focus on two facilitators of this: incentives and education. There needs to be infrastructure, such as research funding, and recognition of discipline-specific needs, such as promotion requirements, to support and sustain deep interdisciplinary work. Education may include basic training to understand what different fields, including one's own, have to offer towards solving a particular problem. In addition to supporting multidisciplinary and interdisciplinary teams, where each member contributes their disciplinary expertise to the formulation or solving of a problem, there is a need for training of individuals in multiple disciplines, to be bridges for the aforementioned, and to strive for transdisciplinary thought, where problems can be approached from a new way of thinking. This thesis provides one example of infrastructure and training that supported the combination of multiple disciplines to tackle challenges related to primary care for complex populations.

6.2.2 The Alliance for Healthier Communities

Future work with the Alliance will include further consultations with stakeholders regarding the substantive findings from Chapter 4, to better understand how identified characteristics do or do not align with current perceptions of their primary care population, and to identify priority areas for future engagement. This knowledge sharing can inform current and future projects at the CHC and/or Alliance-wide level. For example, certain CHCs may have greater capacity to investigate potential

next steps on a project that is specific to their client population or priority care areas. An area of general interest across the Alliance is how to best care for clients who seek care more frequently than expected or necessary. CHCs provide various individual and group programs that may benefit these "frequent visitors", such as social prescribing initiatives to address upstream drivers of care access.[46,247,249]

Two analyses were done to investigating factors associated with high frequency of care visits to primary care providers (nurse practitioner or medical doctor). First, table-based comparisons of sociodemographic, clinical, and healthcare use characteristics of all-time frequent visitors with those not meeting that definition. For each CHC, frequent visitors included clients in the top 10% of average days of primary care provider visits per observation year. Second, a risk factor analysis for a period of frequent visits. This analysis treated each quarter-year as a possible episode of interest: being in the top 10% of the maximum quarter-year care access frequencies (not including the first year) across all clients from the same CHC as the client of interest. To estimate the risk of frequent visitor status in the next quarter-year, multilevel modelling was performed using logistic regression and blocked variable entry. Detailed method and results are in Appendix F. A key finding was an association between social isolation or loneliness with frequent visits, which partially motivated social isolation or loneliness as the outcome in the Chapter 5 case study.

Of note, this thesis frames the population served by the Alliance in terms of complexity. Other areas where complexity is a focus and advanced methods may be able to contribute to is syndemics, where co-occurrence of clinical conditions is viewed in conjunction with social conditions, and research investigates how multiple intersecting factors can contribute towards exacerbated health disparities.[288] The methods used in the present thesis may be useful in syndemics and techniques and theory from syndemics may be able to contribute towards understanding of complex primary care populations. For example, CHC clients with mental health and/or substance use disorders have been found to have higher healthcare use than people receiving care elsewhere.[169] Social determinants of health are expected to play a large role in these

types of findings, and likely have an influence on the types of interventions that will or will not work well.

6.2.3 Epidemiology and computer science methods research

A first avenue for future research is the predictive and interpretation impacts of penalization on features as compared to kernel coefficients. In our HFSM work we placed an L1 penalty on the kernel coefficients to obtain a sparse model that is more practical to deploy; feature coefficients were left unpenalized as we developed these models under a framework where feature selection was done based on theory rather than purely data-driven methods. Nonetheless, penalization of features, such as with lasso logistic regression, is a well-known practice.[289] Feature data entered into a linear kernel as compared to logistic regression are equivalent if left unpenalized (primal vs. dual forms); however, an L1-penalty operates slightly differently on the two forms. An L1-penalty applied to features roughly represents selection of client characteristics to maintain in the model as compared to the selection of entire client records from training data, as is the case for an L1 penalty on kernel coefficients.

A second area for future work includes how to best capture similarity in high-dimensional indicator data for complex populations. In Chapter 5 we introduced the idea of using population-level prevalence to inform whether it is the presence or absence of any given indicator that may be most informative. In addition to extending this work, research could explore how to best incorporate recency of codes and temporality. The latter is related to work on similarity based on sequences, where the exact code sequence matters[223,224,290,291]; in primary care temporality may matter for certain outcomes but the ordering is likely only informative up to a certain point, after which it becomes noise. For example, lab values entering abnormal ranges may matter on a short-term scale while the order that chronic conditions are diagnosed in may only matter on the scale of years; the ordering of these types of diagnoses within a short time period is likely more due to care characteristics than preventable or long-term changes in underlying physiology. This

work is also expected to be related to that on missing data, both missingness in terms of individual data elements and in terms of the timescale between when something becomes true or relevant for a client and when it appears in their EHR.

Finally, most information derived from clinical epidemiology and machine learning is population-specific, and how far outside of the population represented in a particular research study or training dataset the information pertains to is a challenging question. The two fields approach these knowledge generalizability questions from slightly different angles, and there may be value to investigating and expanding their overlap. For example, can techniques intended to prevent or assess selection bias in epidemiology be used to proactively inform or improve machine learning model generalizability planning or assessment; can post hoc machine learning fairness assessments help understand situations where epidemiological findings may apply outside of the original target population; and can new statistical techniques be developed that apply to both settings?

6.3 Conclusions

Over the past few years, the state of AI for primary care has been transitioning; this thesis is positioned within that transition. We began with the first scoping review on AI and primary care research, which called for an increase in high-quality, interdisciplinary research with more primary care leadership. The Alliance for Healthier Communities is an example of an organization that is committed to learning from their data to support and improve care for the clients they serve; we provided the first large-scale description of this population, demonstrating the use of both simple statistical and unsupervised learning techniques to properly capture complexity, and deriving insights to inform future LHS initiatives. We then proposed hybrid feature- and similarity-based supervised learning methodology for prediction and interpretation, demonstrating and testing their use with synthetic data and in a case study to predict social isolation and loneliness within the Alliance. These interpretable methods are well-suited for co-design studies around predictive models, and can serve as

a tool for investigation of relationships in epidemiological studies. Finally, we began thinking about how best to capture similarity among primary care clients by extending the coefficient of similarity to consider rare and common codes differently.

In summary, motivated by the potential to improve primary care for complex clients, this thesis integrated epidemiology and computer science to understand the state of AI for primary care, to identify opportunities and challenges to AI and other data-driven initiatives within a complex primary care population, and to develop novel machine learning methodology for prediction and interpretation with large scale observational data.

6.4 References

1. Foley T, Horwitz L, Zahran R. *Realising the Potential of Learning Health Systems*. Newcastle University; 2021:101. https://learninghealthcareproject.org/wp-content/uploads/2021/05/LHS2021report.pdf.
2. CIFAR. *Building a Learning Health System for Canadians. Report of the Artificial Intelligence for Health Task Force*. Canada; 2020:32. https://cifar.ca/wp-content/uploads/2020/11/AI4Health-report-ENG-10-F.pdf.
3. Friedman CP, Allee NJ, Delaney BC, et al. The science of Learning Health Systems: Foundations for a new journal. *Learn Health Syst*. 2017;1(1):e10020. doi:10.1002/lrh2.10020
4. Nash DM, Bhimani Z, Rayner J, Zwarenstein M. Learning health systems in primary care: A systematic scoping review. *BMC Fam Pract*. 2021;22(1):126. doi:10.1186/s12875-021-01483-z
5. Starfield B. *Primary Care. Balancing Health Needs, Services, and Technology*. New York, NY: Oxford University Press, Inc.; 1998.
6. *CIHR Primary Healthcare Summit 2010 Final Report Summary*. Toronto, Ontario: Canadian Institutes of Health Research; 2010.
7. Hogg W, Rowan M, Russell G, Geneau R, Muldoon L. Framework for primary care organizations: The importance of a structural domain. *Int J Qual Health Care*. 2008;20(5):308-313. doi:10.1093/intqhc/mzm054
8. Mercer S, Salisbury C, Fortin M. *ABC of Multimorbidity*. Wiley Blackwell; 2014.
9. Albrecht D. Community health centres in Canada. *Leadersh Health Serv*. 1998;11(1):5-10. doi:10.1108/13660759810202596

10. Karam M, Chouinard M-C, Poitras M-E, et al. Nursing care coordination for patients with complex needs in primary healthcare: A scoping review. *Int J Integr Care*. 2021;21(1):16. doi:10.5334/ijic.5518

11. Smith SM, Soubhi H, Fortin M, Hudon C, O'Dowd T. Managing patients with multimorbidity: Systematic review of interventions in primary care and community settings. *BMJ*. 2012;345(sep03 1):e5205-e5205. doi:10.1136/bmj.e5205

12. Kastner M, Cardoso R, Lai Y, et al. Effectiveness of interventions for managing multiple high-burden chronic diseases in older adults: A systematic review and meta-analysis. *CMAJ*. 2018;190(34):E1004-E1012. doi:10.1503/cmaj.171391

13. Boyd C, Smith CD, Masoudi FA, et al. Decision making for older adults with multiple chronic conditions: Executive summary for the American Geriatrics Society guiding principles on the care of older adults with multimorbidity. *J Am Geriatr Soc*. 2019;67(4):665-673. doi:10.1111/jgs.15809

14. McGilton KS, Vellani S, Yeung L, et al. Identifying and understanding the health and social care needs of older adults with multiple chronic conditions and their caregivers: A scoping review. *BMC geriatr*. 2018;18(1):231. doi:10.1186/s12877-018-0925-x

15. Fortin M, Contant E, Savard C, Hudon C, Poitras M-E, Almirall J. Canadian guidelines for clinical practice: An analysis of their quality and relevance to the care of adults with comorbidity. *BMC Fam Prac*. 2011;12(1). doi:10.1186/1471-2296-12-74

16. Galopin A, Bouaud J, Pereira S, Séroussi B. Clinical practice guidelines consistency for patients with multimorbidity: A case-study in the management of type 2 diabetes and hypertension. *Stud Health Technol Inform*. 2015;210:344-348.

17. Abidi S. A knowledge-modeling approach to integrate multiple clinical practice guidelines to provide evidence-based clinical decision support for managing comorbid conditions. *J Med Syst*. 2017;41(12). doi:10.1007/s10916-017-0841-1

18. Pathirana TI, Jackson CA. Socioeconomic status and multimorbidity: A systematic review and meta-analysis. *Aust N Z J Public Health*. 2018;42(2):186-194. doi:10.1111/1753-6405.12762

19. Muth C, Blom JW, Smith SM, et al. Evidence supporting the best clinical management of patients with multimorbidity and polypharmacy: A systematic guideline review and expert consensus. *J Intern Med*. 2019;285(3):272-288. doi:10.1111/joim.12842

20. Rothwell PM. External validity of randomised controlled trials: "To whom do the results of this trial apply?" *Lancet*. 2005;365(9453):82-93. doi:10.1016/S0140-6736(04)17670-8

21. Robinson JM, Trochim WMK. An examination of community members', researchers' and health professionals' perceptions of barriers to minority participation in medical research: An application of concept mapping. *Ethn Health*. 2007;12(5):521-539. doi:10.1080/13557850701616987

22. George S, Duran N, Norris K. A systematic review of barriers and facilitators to minority research participation among African Americans, Latinos, Asian Americans, and Pacific Islanders. *Am J Public Health*. 2014;104(2):e16-e31. doi:10.2105/AJPH.2013.301706

23. Odierna DH, Schmidt LA. The effects of failing to include hard-to-reach respondents in longitudinal surveys. *Am J Public Health*. 2009;99(8):1515-1521. doi:10.2105/AJPH.2007.111138

24. Heiat A, Gross CP, Krumholz HM. Representation of the elderly, women, and minorities in heart failure clinical trials. *Arch Intern Med*. 2002;162(15). doi:10.1001/archinte.162.15.1682

25. Bonevski B, Randell M, Paul C, et al. Reaching the hard-to-reach: A systematic review of strategies for improving health and medical research with socially disadvantaged groups. *BMC Med Res Methodol*. 2014;14:42. doi:10.1186/1471-2288-14-42

26. Kwiatkowski K, Coe K, Bailar JC, Swanson GM. Inclusion of minorities and women in cancer clinical trials, a decade later: Have we improved? *Cancer*. 2013;119(16):2956-2963. doi:10.1002/cncr.28168

27. Lin SY, Mahoney MR, Sinsky CA. Ten ways artificial intelligence will transform primary care. *J Gen Intern Med*. 2019;34(8):1626-1630. doi:10.1007/s11606-019-05035-1

28. Hinton G. Deep Learning - A technology with the potential to transform health care. *JAMA*. 2018;320(11):1101-1102.

29. Jiang F, Jiang Y, Zhi H, et al. Artificial intelligence in healthcare: Past, present and future. *Stroke Vasc Neurol*. 2017;2(4):230-243. doi:10.1136/svn-2017-000101

30. Wiens J, Saria S, Sendak M, et al. Do no harm: A roadmap for responsible machine learning for health care. *Nat Med*. 2019;25:1337-1340. doi:10.1038/s41591-019-0548-6

31. Wiens J, Shenoy ES. Machine learning for healthcare: On the verge of a major shift in healthcare epidemiology. *Clin Infect Dis*. 2018;66(1):149-153. doi:10.1093/cid/cix731

32. Ghassemi M, Naumann T, Schulam P, Beam AL, Chen IY, Ranganath R. A review of challenges and opportunities in machine learning for health. *AMIA Summits on Translational Science Proceedings*. 2020;2020:191-200. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7233077/.

33. Navarro CLA, Damen JAA, Takada T, et al. Risk of bias in studies on prediction models developed using supervised machine learning techniques: Systematic review. *BMJ*. 2021;375:n2281. doi:10.1136/bmj.n2281

34. Chen IY, Pierson E, Rose S, Joshi S, Ferryman K, Ghassemi M. Ethical machine learning in healthcare. *Annu Rev Biomed Data Sci*. 2021;4(1):123-144. doi:10.1146/annurev-biodatasci-092820-114757

35. Weissler EH, Naumann T, Andersson T, et al. The role of machine learning in clinical research: Transforming the future of evidence generation. *Trials*. 2021;22(1):537. doi:10.1186/s13063-021-05489-x

36. Gianfrancesco MA, Tamang S, Yazdany J, Schmajuk G. Potential biases in machine learning algorithms using electronic health record data. *JAMA Intern Med*. August 2018. doi:10.1001/jamainternmed.2018.3763

37. Liaw W, Kakadiaris I. Artificial intelligence and family medicine: Better together. *Fam Med*. 2020;52(1):8-10. doi:10.22454/FamMed.2020.881454

38. Guo Y, Hao Z, Zhao S, Gong J, Yang F. Artificial intelligence in health care: Bibliometric analysis. *J Med Internet Res*. 2020;22(7):e18228. doi:10.2196/18228

39. Doshi-Velez F, Kim B. Considerations for Evaluation and Generalization in Interpretable Machine Learning. In: Escalante HJ, Escalera S, Guyon I, et al., eds. *Explainable and Interpretable Models in Computer Vision and Machine Learning*. Cham: Springer International Publishing; 2018:3-17. doi:10.1007/978-3-319-98131-4_1

40. Tonekaboni S, Joshi S, McCradden MD, Goldenberg A. What clinicians want: Contextualizing explainable machine learning for clinical end use. In: *Proceedings of Machine Learning Research*.; 2019:21.

41. Rudin C. Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead. *Nat Mach Intell*. 2019;1(5):206-215. doi:10.1038/s42256-019-0048-x

42. Ghassemi M, Oakden-Rayner L, Beam AL. The false hope of current approaches to explainable artificial intelligence in health care. *Lancet Digit Health*. 2021;3(11):e745-e750. doi:10.1016/S2589-7500(21)00208-9

43. Payrovnaziri SN, Chen Z, Rengifo-Moreno P, et al. Explainable artificial intelligence models using real-world electronic health record data: A systematic scoping review. *J Am Med Infrom Assoc*. 2020;27(7):1173-1185. doi:10.1093/jamia/ocaa053

44. Alliance for Healthier Communities. 2020. https://www.allianceon.org/. Accessed January 13, 2020.

45. Alliance for Healthier Communities. Moving Forward as a Learning Health System. *Alliance for Healthier Communities*. November 2020. https://myemail.constantcontact.com/EPIC-News--Issue-1.html?soid=1108953382524&aid=uzy8bphr91U. Accessed November 23, 2020.

46. Alliance for Healthier Communities. *Towards a Learning Health System: Better Care Tomorrow When We Learn from Today*. Alliance for Healthier Communities; 2020:15. https://www.allianceon.org/sites/default/files/documents/Learning%20Health%20System%20report%202020-10-20%20-%20FINAL_JR.pdf. Accessed November 23, 2020.

47. Greenland S, Rothman K. Chapter 3. Measures of Occurence. In: *Modern Epidemiology*. 3rd ed. Philadelphia, USA: Lippincott Williams & Wilkins; 2008:32-33.

48. Baron J. Chapter 19. Clinical epidemiology. In: *Teaching Epidemiology a Guide for Teachers in Epidemiology, Publis Health and Clinical Medicine*. 2nd ed. United States: Oxford University Press; 2001:237-251.

49. Kaufman J. Chapter 26. Social Epidemiology. In: *Modern Epidemiology*. 3rd ed. Philadelphia, USA: Lippincott Williams & Wilkins; 2008:532-533.

50. What are the social determinants of health? *Canadian Public Health Association*. 2020. https://www.cpha.ca/what-are-social-determinants-health. Accessed May 4, 2020.

51. Model of Health and Wellbeing. *Alliance for Healthier Communities*. https://www.allianceon.org/model-health-and-wellbeing. Accessed May 17, 2022.

52. Government of Canada. Health services research. Canadian institutes for health research. https://cihr-irsc.gc.ca/e/48809.html. Published May 15, 2019. Accessed June 7, 2022.

53. Denning PJ, Comer DE, Gries D, et al. Computing as a discipline. *Computer*. 1989;22(2):63-70. doi:10.1109/2.19833

54. Kueper JK, Terry AL, Zwarenstein M, Lizotte DJ. Research on artificial intelligence and primary care: A scoping review. *medRxiv*. August 2019. doi:10.1101/19003913

55. Russell S, Norvig P. *Artificial Intelligence A Modern Approach*. 3rd ed. Pearson Education Inc.; 2010.

56. Turing A. Computing machinery and intelligence. *Mind*. 1950;59(236):433-460. http://www.jstor.org/stable/2251299.

57. Kueper JK. Primer for artificial intelligence in primary care. *Can Fam Physician*. 2021;67(12):889-893. doi:10.46747/cfp.6712889

58. Rabe-Hesketh S, Skrondal A. Classical latent variable models for medical research. *Stat Methods Med Res*. 2008;17(1):5-32. doi:10.1177/0962280207081236

59. Formann AK, Kohlmann T. Latent class analysis in medical research. *Stat Methods Med Res*. 1996;5(2):179-211. doi:10.1177/096228029600500205

60. Arrieta A, Diaz-Rodriguez N, Ser J, et al. Explainable Artificial Intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI. *Inf Fushion*. 2020;58:82-115. doi:10.1016/j.inffus.2019.12.012

61. Clinciu M-A, Hastie H. A survey of explainable AI terminology. In: *Proceedings of the 1st Workshop on Interactive Natural Language Technology for Explainable Artificial Intelligence (Nl4xai 2019)*. Tokyo, Japan: Association for Computational Linguistics; 2019:8-13. doi:10.18653/v1/W19-8403

62. Adadi A, Berrada M. Peeking inside the black-box: A survey on Explainable Artificial Intelligence (XAI). *IEEE Access*. 2018;6:52138-52160. doi:10.1109/ACCESS.2018.2870052

63. Black J, Terry A, Lizotte D. FRAMR-EMR: Framework for Prognostic Predictive Model Development Using Electronic Medical Record Data with a case study in osteoarthritis risk. *arXiv:170509563 [stat]*. September 2017. http://arxiv.org/abs/1705.09563. Accessed May 5, 2020.

64. Pearl J. *Causality. Models, Reasoning, and Inference*. United States of America: Cambridge University Press; 2000.

65. Greenland S, Pearl J, Robins JM. Causal diagrams for epidemiologic research. *Epidemiology*. 1999;10(1):37-48. doi:10.1097/00001648-199901000-00008

66. Muldoon LK, Hogg WE, Levitt M. Primary Care (PC) and Primary Health Care (PHC): What is the difference? *Can J Public Health*. 2006;97(5):409. doi:10.1007/BF03405354

67. Pimlott N. For family physicians, by family physicians? *Can Fam Physician*. 2016;62:699.

68. Glouberman S, Zimmerman B. *Complicated and Complex Systems: What Would Successful Reform of Medicare Look Like?* Ottawa, ON: Commission on the Future of Health Care in Canada; 2002. doi:10.3138/9781442672833-004

69. Horwitz L. *A Clinician's Perspective on Machine Learning in Healthcare*. Machine Learning for Healthcare; 2020. https://www.youtube.com/watch?v=7skU_7MNzRw.

70. Liaw W, Kakadiaris IA. Primary care artificial intelligence: A branch hiding in plain sight. *Ann Fam Med*. 2020;18(3):194-195. doi:10.1370/afm.2533

71. Electronic Health Records. *Canada Health Infoway*. 2020. https://www.infoway-inforoute.ca/en/solutions/digital-health-foundation/electronic-health-records. Accessed May 4, 2020.

72. What is a learning health system? The learning healthcare project. May 2021. https://learninghealthcareproject.org/introduction-and-rationale/what-is-a-learning-health-system/.

73. Smith MD, Institute of Medicine (U.S.), eds. *Best Care at Lower Cost: The Path to Continuously Learning Health Care in America*. Washington, D.C: National Academies Press; 2012.

74. Delaney BC, Peterson KA, Speedie S, Taweel A, Arvanitis TN, Hobbs FDR. Envisioning a Learning Health Care System: The Electronic Primary Care Research Network, a case study. *Ann Fam Med*. 2012;10(1):54-59. doi:10.1370/afm.1313

75. Lindsell CJ, Gatto CL, Dear ML, et al. Learning from what we do, and doing what we learn: A learning health care system in action. *Acad Med*. 2021;96(9):1291-1299. doi:10.1097/ACM.0000000000004021

76. CFPC. Imagining the future of family medicine. May 2019.

77. Girouard A. The new CFPC-AMS TechForward fellow will explore artificial intelligence opportunities in family medicine. May 2021. https://www.cfpc.ca/en/news-and-events/news-events/news-events/news-releases/2021/the-new-cfpc-ams-techforward-fellow-will-explore-a. Accessed January 5, 2022.

78. American Board of Family Medicine SU The Center for Professionalism and Value in Health Care. Setting a research agenda for the use of artificial intelligence & machine learning in primary care. The center for professionalism and value in health care. https://professionalismandvalue.org/setting-a-research-agenda-for-the-use-of-artificial-intelligence-machine-learning-in-primary-care/. Published March 18, 2021. Accessed June 16, 2021.

79. Tricco AC, Lillie E, Zarin W, et al. PRISMA extension for scoping reviews (PRISMA-ScR): Checklist and explanation. *Ann Intern Med*. 2018;169(7):467. doi:10.7326/M18-0850

80. Levac D, Colquhoun H, O'Brien KK. Scoping studies: Advancing the methodology. *Implement Sci*. 2010;5(69). doi:10.1186/1748-5908-5-69

81. Arksey H, O'Malley L. Scoping studies: Towards a methodological framework. *Int J Soc Res Methodol*. 2005;8(1):19-32. doi:10.1080/1364557032000119616

82. Thomas J, Kneale D, McKenzie JE, Brennan SE, Bhaumik S. Chapter 2: Determining the scope of the review and the questions it will address. In: *Cochrane Handbook for Systematic Reviews of Interventions*. Vol 6.3.; 2022. https://training.cochrane.org/handbook/current/chapter-02. Accessed May 18, 2022.

83. Beam AL, Kohane IS. Big data and machine learning in health care. *JAMA*. 2018;319(13):1317. doi:10.1001/jama.2017.18391

84. Gray R. The a-z of how artificial intelligence is changing the world. November 2018. http://www.bbc.com/future/gallery/20181115-a-guide-to-how-artificial-intelligence-is-changing-the-world. Accessed March 18, 2019.

85. Davenport TH, Ronanki R. *Artificial Intelligence for the Real World*. Harvard Business Review; 2018. https://hbr.org/2018/01/artificial-intelligence-for-the-real-world.

86. Rajkomar A, Dean J, Kohane I. Machine learning in medicine. *NEJM*. 2019;380(14):1347-1358. doi:10.1056/NEJMra1814259

87. Topol E. *The Topol Review. Preparing the Healthcare Workforce to Deliver the Digital Future.* NHS; 2019.

88. U.S. Agency for International Development. *Artificial Intelligence in Global Health: Defining a Collective Path Moving Forward*. USAID; The Rockefeller Foundation; Bill & Melinda Gates Foundation; 2019. https://www.usaid.gov/cii/ai-in-global-health. Accessed June 17, 2019.

89. Hitching R. Primary care 2.0. M. https://becominghuman.ai/the-changing-face-of-healthcare-artificial-intelligence-primary-care-an-empowered-patients-8ba638ffacd6. Published February 20, 2019. Accessed July 8, 2019.

90. Pratt M. Artificial intelligence in primary care. Medical economics. https://www.medicaleconomics.com/business/artificial-intelligence-primary-care. Published 2018. Accessed July 8, 2019.

91. Moore SF, Hamilton W, Llewellyn DJ. Harnessing the power of intelligent machines to enhance primary care. *Br J Gen Pract*. 2018;68(666):6-7. doi:10.3399/bjgp17X693965

92. Huang M, Gibson C, Terry A. Measuring electronic health record use in primary care: A scoping review. *Appl Clin Inform*. 2018;09(1):15-33. doi:10.1055/s-0037-1615807

93. Terry AL, Stewart M. How does canada stack up? A bibliometric analysis of the primary healthcare electronic medical record literature. *Inform Prim Care*. 2012;20:233-240.

94. Kringos DS, Boerma WG, Hutchinson A, Zee J van der, Groenewegen PP. The breadth of primary care: A systematic literature review of its core dimensions. *BMC Health Serv Res*. 2010;10(65). doi:10.1186/1472-6963-10-65

95. Friedman C, Elhadad N. Natural language processing in health care and biomedicine. In: Shortliffe EH, Cimino JJ, eds. *Biomedical Informatics*. London: Springer London; 2014:255-284. doi:10.1007/978-1-4471-4474-8_8

96. Yu K-H, Beam AL, Kohane IS. Artificial intelligence in healthcare. *Nat Biomed Eng*. 2018;2(10):719-731. doi:10.1038/s41551-018-0305-z

97. Talukder AK. Big data analytics advances in health intelligence, public health, and evidence-based precision medicine. In: Reddy PK, Sureka A, Chakravarthy S, Bhalla S, eds. *Big Data Analytics*. Vol 10721. Cham: Springer International Publishing; 2017:243-253. doi:10.1007/978-3-319-72413-3_17

98. Greenhalgh T, Potts HWW, Wong G, Bark P, Swinglehurst D. Tensions and paradoxes in electronic patient record research: A systematic literature review using the meta-narrative method: Electronic patient record research. *Milbank Q*. 2009;87(4):729-788. doi:10.1111/j.1468-0009.2009.00578.x

99. Greenhalgh T, Hinder S, Stramer K, Bratan T, Russell J. Adoption, non-adoption, and abandonment of a personal electronic health record: Case study of HealthSpace. *BMJ*. 2010;341:c5814-c5814. doi:10.1136/bmj.c5814

100. Terry AL, Ryan BL, McKay S, et al. Towards optimal electronic medical record use: Perspectives of advanced users. *Fam Pract*. 2018;35(5):607-611. doi:10.1093/fampra/cmy002

101. Shortliffe EH, Axline SG, Buchanan BG, Green CC, Cohen SN. Computer-based consultations in clinical therapeutics: Explanation and rule acquisition capabilities of the MYCIN system. *Comput Biomed Res.* 1975;8:303-320.

102. Alanazi HO, Abdullah AH, Qureshi KN. A critical review for developing accurate and dynamic predictive models using machine learning methods in medicine and health care. *J Med Syst.* 2017;41(4). doi:10.1007/s10916-017-0715-6

103. Miotto R, Wang F, Wang S, Jiang X, Dudley JT. Deep learning for healthcare: Review, opportunities and challenges. *Brief Bioinform.* 2018;19(6):1236-1246. doi:10.1093/bib/bbx044

104. Miotto R, Li L, Kidd BA, Dudley JT. Deep patient: An unsupervised representation to predict the future of patients from the electronic health records. *Sci Rep.* 2016;6. doi:10.1038/srep26094

105. Esteva A, Kuprel B, Novoa RA, et al. Dermatologist-level classification of skin cancer with deep neural networks. *Nature.* 2017;542(7639):115-118. doi:10.1038/nature21056

106. Stanganelli I, Brucale A, Calori L, et al. Computer-aided diagnosis of melanocytic lesions. *Anticancer Res.* 2005;25:4577-4582.

107. Collier R. Electronic health records contributing to physician burnout. *CMAJ.* 2017;189(45):E1405-E1406. doi:10.1503/cmaj.109-5522

108. Arndt BG, Beasley JW, Watkinson MD, et al. Tethered to the EHR: Primary care physician workload assessment using EHR event log data and time-motion observations. *Ann Fam Med.* 2017;15(5):419-426. doi:10.1370/afm.2121

109. Tricco AC, Lillie E, Zarin W, et al. A scoping review on the conduct and reporting of scoping reviews. *BMC Med Res Methodol.* 2016;16(1). doi:10.1186/s12874-016-0116-4

110. *Covidence Systematic Review Software.* Melbourne, Australia: Veritas Health Innovation www.covidence.org.

111. Craske MG, Rose RD, Lang A, et al. Computer-assisted delivery of cognitive behavioral therapy for anxiety disorders in primary-care settings. *Depress Anxiety.* 2009;26(3):235-242. doi:10.1002/da.20542

112. Cucciare MA, Curran GM, Craske MG, et al. Assessing fidelity of cognitive behavioral therapy in rural VA clinics: Design of a randomized implementation effectiveness (hybrid type III) trial. *Implement Sci.* 2015;11(1). doi:10.1186/s13012-016-0432-4

113. Graaf LE de, Gerhards SAH, Arntz A, et al. Clinical effectiveness of online computerised cognitive–behavioural therapy without support for depression in primary care: Randomised trial. *Br J Psychiatry.* 2009;195(1):73-80. doi:10.1192/bjp.bp.108.054429

114. Graaf LE de, Hollon SD, Huibers MJH. Predicting outcome in computerized cognitive behavioral therapy for depression in primary care: A randomized trial. *J Consult Clin Psychol*. 2010;78(2):184-189. doi:10.1037/a0018324

115. Graaf LE de, Gerhards SAH, Arntz A, et al. One-year follow-up results of unsupported online computerized cognitive behavioural therapy for depression in primary care: A randomized trial. *J Behav Ther Exp Psychiatry*. 2011;42(1):89-95. doi:10.1016/j.jbtep.2010.07.003

116. Chang EK, Yu CY, Clarke R, et al. Defining a patient population with cirrhosis: An automated algorithm with natural language processing. *J Clin Gastroenterol*. 2016;50(10):889-894. doi:10.1097/MCG.0000000000000583

117. Widmer G, Horn W, Nagele B. Automatic knowledge base refinement: Learning from examples and deep knowledge in rheumatology. *Artif Intell Med*. 1993;5(3):225-243. doi:10.1016/0933-3657(93)90026-Y

118. Michalowski M, Michalowski W, Wilk S, O'Sullivan D, Marc Carrier M. AFGuide system to support personalized management of atrial fibrillation. In: *Joint Workshop on Health Intelligence.*; 2017. https://aaai.org/ocs/index.php/WS/AAAIW17/paper/view/15215.

119. Glasspool DW, Fox J, Coulson AS, Emery J. Risk assessment in genetics: A semi-quantitative approach. *Stud Health Technol Inform*. 2001;2001(84):459-463.

120. Flores CD, Marcelo FJ, Rosecler BM, Ana R, Helder C. Method for building a medical training simulator with bayesian networks: SimDeCS. In: *2nd KES International Conference on Innovation in Medicine and Healthcare*. Vol 207. San Sebastian, Spain: Stud Health Technol Inform; 2014:102-114. doi:10.3233/978-1-61499-474-9-102

121. Astilean A, Avram C, Folea S, Silvasan I, Petreus D. Fuzzy petri nets based decision support system for ambulatory treatment of non-severe acute diseases. In: *2010 IEEE International Conference on Automation, Quality and Testing, Robotics (AQTR)*. Cluj-Napoca, Romania: IEEE; 2010:1-6. doi:10.1109/AQTR.2010.5520706

122. Dominguez-Morales JP, Jimenez-Fernandez AF, Dominguez-Morales MJ, Jimenez-Moreno G. Deep neural networks for the recognition and classification of heart murmurs using neuromorphic auditory sensors. *IEEE Trans Biomed Circuits Syst*. 2018;12(1):24-34. doi:10.1109/TBCAS.2017.2751545

123. Zamora M, Baradad M, Amado E, et al. Characterizing chronic disease and polymedication prescription patterns from electronic health records. In: *2015 IEEE International Conference on Data Science and Advanced Analytics (DSAA)*. Campus des Cordeliers, Paris, France: IEEE; 2015:1-9. doi:10.1109/DSAA.2015.7344870

124. Orient JM. Evaluation of abdominal pain: Clinicians' performance compared with three protocols. *South Med J*. 1986;79(7):793-799.

125. Sumner 2nd. W, Truszczynski M, Marek VW. Simulating patients with parallel health state networks. *Proceedings / AMIA Annual Symposium AMIA Symposium*. 1998:438-442. https://www.scopus.com/inward/record.uri?eid=2-s2.0-0032239216&partnerID=40&md5=23e9e4bc917433bed8b821cc2311e64a.

126. Sumner 2nd. W, Xu JZ, Roussel G, Hagen MD. Modeling relief. In: *AMIA Annual Symposium Proceedings*.; 2007:706-710.

127. Sumner 2nd. W, Xu JZ. Modeling fatigue. In: *AMIA Annual Symposium Proceedings*.; 2002:747-751.

128. Sumner 2nd W, Hagen MD, Rovinelli R. The item generation methodology of an empiric simulation project. *Adv Health Sci Educ*. 1999;4:49-66.

129. Zhuang ZY, Churilov L, Sikaris K. Uncovering the patterns in pathology ordering by australian general practitioners: A data mining perspective. In: *Proceedings of the 39th Annual Hawaii International Conference on System Sciences (HICSS'06)*. Kauia, HI, USA: IEEE; 2006:92c-92c. doi:10.1109/HICSS.2006.513

130. Zhuang ZY, Amarasiri R, Churilov L, Alahakoon D, Sikaris K. Exploring the clinical notes of pathology ordering by australian general practitioners: A text mining perspective. In: *2007 40th Annual Hawaii International Conference on System Sciences (HICSS'07)*. Waikoloa, HI: IEEE; 2007:136-136. doi:10.1109/HICSS.2007.220

131. Zhuang ZY, Churilov L, Burstein F, Sikaris K. Combining data mining and case-based reasoning for intelligent decision support for pathology ordering by general practitioners. *Eur J Oper Res*. 2009;195(3):662-675. doi:10.1016/j.ejor.2007.11.003

132. Zhuang ZY, Wilkin CL, Ceglowski A. A framework for an intelligent decision support system: A case in pathology test ordering. *Decis Support Sys*. 2013;55(2):476-487. doi:10.1016/j.dss.2012.10.006

133. Voracek D. NASA innovation framework and center innovation fund. 2018.

134. Gupta A, Thorpe C, Bhattacharyya O, Zwarenstein M. Promoting development and uptake of health innovations: The nose to tail tool. *F1000Res*. 2016;5:361. doi:10.12688/f1000research.8145.1

135. Hao K. We analyzed 16,625 papers to figure out where AI is headed next. *MIT Technology Review*. January 2019. https://www.technologyreview.com/s/612768/we-analyzed-16625-papers-to-figure-out-where-ai-is-headed-next/. Accessed March 15, 2019.

136. Scimago journal and country rank. SJR. https://www.scimagojr.com/countryrank.php. Accessed May 31, 2019.

137. Conte ML, Liu J, Schnell S, Omary MB. Globalization and changing trends of biomedical research output. *JCI Insight*. 2017;2(12):e95206. doi:10.1172/jci.insight.95206

138. Hajjar F, Saint-Lary O, Cadwallader J-S, et al. Development of primary care research in north america, europe, and australia from 1974 to 2017. *Ann Fam Med*. 2019;17(1):49-51. doi:10.1370/afm.2328

139. 15 countries with highest technology in the world. *41Studio*. 2019. https://www.41studio.com/blog/2018/15-countries-with-higest-technology-in-the-world/. Accessed May 31, 2019.

140. Sturgiss E, Boven K van. Datasets collected in general practice: An international comparison using the example of obesity. *Aust Health Rev*. 2018;42(5):563. doi:10.1071/AH17157

141. Smeets HM, Kortekaas MF, Rutten FH, et al. Routine primary care data for scientific research, quality of care programs and educational purposes: The julius general practitioners' network (JGPN). *BMC Health Serv Res*. 2018;18(1). doi:10.1186/s12913-018-3528-5

142. eHealth WGO for, Organization WH, eHealth WGO for, eds. *Atlas of eHealth Country Profiles: The Use of eHealth in Support of Universal Health Coverage: Based on the Findings of the Third Global Survey on eHealth, 2015*. Geneva, Switzerland: World Health Organization; 2016.

143. Boyrikova A. Why the netherlands is the new silicon valley. Your sneak preview of the future. https://innovationorigins.com/why-netherlands-is-the-new-silicon-valley/. Published 2017. Accessed June 14, 2019.

144. Europe's hub for r & d innovation. Invest in holland. https://investinholland.com/business-operations/research-and-development/. Published 2019. Accessed June 14, 2019.

145. Doshi-Velez F, Kim B. Towards a rigorous science of interpretable machine learning. *arXiv:170208608 [cs, stat]*. February 2017. http://arxiv.org/abs/1702.08608. Accessed May 31, 2019.

146. Holzinger A, Biemann C, Pattichis CS, Kell DB. What do we need to build explainable AI systems for the medical domain? 2017. http://arxiv.org/abs/1712.09923. Accessed May 31, 2019.

147. Miller T. Explanation in artificial intelligence: Insights from the social sciences. June 2017. http://arxiv.org/abs/1706.07269. Accessed May 31, 2019.

148. Verghese A, Shah NH, Harrington RA. What this computer needs is a physician: Humanism and artificial intelligence. *JAMA*. 2018;319(1):19. doi:10.1001/jama.2017.19198

149. Christodoulou E, Ma J, Collins GS, Steyerberg EW, Verbakel JY, Van Calster B. A systematic review shows no performance benefit of machine learning over logistic regression for clinical prediction models. *J Clin Epidemiol*. 2019;110:12-22. doi:10.1016/j.jclinepi.2019.02.004

150. Rothman KJ. *Epidemiology: An Introduction*. 2nd ed. New York, NY: Oxford University Press; 2012.

151. Bhopal R. *Concepts of Epidemiology. Integrating the Ideas, Theories, Principles and Methods of Epidemiology.* 2nd ed. United Kingdom: Oxford University Press; 2008.

152. Kindermann R, Snell JL. *Markov Random Fields and Their Applications.* American Mathematics Society; 1980. https://web.archive.org/web/20170810092327/http://www.cmap.polytechnique.fr/~rama/ehess/mrfbook.pdf. Accessed June 10, 2022.

153. Bishop CM. *Pattern Recognition and Machine Learning.* New York: Springer; 2006.

154. Ising E. Beitrag zur Theorie des Ferromagnetismus. *Z Physik.* 1925;31(1):253-258. doi:10.1007/BF02980577

155. Haslbeck JMB, Epskamp S, Marsman M, Waldorp LJ. Interpreting the ising model: The input matters. *Multivariate Behavioral Research.* 2021;56(2):303-313. doi:10.1080/00273171.2020.1730150

156. Cheng J, Levina E, Wang P, Zhu J. A sparse ising model with covariates. *Biometrics.* 2014;70(4):943-953. doi:10.1111/biom.12202

157. Clark NJ, Wells K, Lindberg O. Unravelling changing interspecific interactions across environmental gradients using markov random fields. *Ecology.* 2018;99(6):1277-1283. doi:10.1002/ecy.2221

158. Borkulo CD van, Borsboom D, Epskamp S, et al. A new method for constructing networks from binary data. *Sci Rep.* 2014;4(1):5918. doi:10.1038/srep05918

159. Lee DD, Seung HS. Learning the parts of objects by non-negative matrix factorization. *Nature.* 1999;401(6755):788-791. doi:10.1038/44565

160. Lee D, Seung HS. Algorithms for non-negative matrix factorization. In: *Advances in Neural Information Processing Systems.* Vol 13. MIT Press; 2000. https://proceedings.neurips.cc/paper/2000/hash/f9d1152547c0bde01830b7e8bd60024c-Abstract.html. Accessed June 10, 2022.

161. Yang Z, Zhang H, Yuan Z, Oja E. Kullback-leibler divergence for nonnegative matrix factorization. In: Honkela T, Duch W, Girolami M, Kaski S, eds. *Artificial Neural Networks and Machine Learning – ICANN 2011.* Lecture notes in computer science. Berlin, Heidelberg: Springer; 2011:250-257. doi:10.1007/978-3-642-21735-7_31

162. Hastie T, Tibshirani R, Friedman J. *The Elements of Statistical Learning.* 2nd ed. California: Springer; 2008.

163. Aghabozorgi S, Seyed Shirkhorshidi A, Ying Wah T. Time-series clustering – a decade review. *Inf Syst.* 2015;53:16-38. doi:10.1016/j.is.2015.04.007

164. Schubert E, Rousseeuw PJ. Fast and eager k-medoids clustering: O(k) runtime improvement of the PAM, CLARA, and CLARANS algorithms. *Information Systems.* 2021;101:101804. doi:10.1016/j.is.2021.101804

165. Berndt DJ, Clifford J. Using dynamic time warping to find patterns in time series. In: *Proceedings of the 3rd International Conference on Knowledge Discovery and Data Mining*. AAAIWS'94. Seattle, WA: AAAI Press; 1994:359-370.

166. Montero P, Vilar JA. TSclust: An r package for time series clustering. *J Stat Softw*. 2015;62:1-43. doi:10.18637/jss.v062.i01

167. Oates T, Firoiu L, Cohen PR. Clustering time series with hidden markov models and dynamic time warping. In: *In Proceedings of the Ijcai-99 Workshop on Neural, Symbolic and Reinforcement Learning Methods for Sequence Learning.*; 1999:17-21.

168. Glazier RH, Zagorski B, Rayner J. *Comparison of Primary Care Models in Ontario by Demographics, Case Mix and Emergency Department Use, 2008/09 to 2009/10*. Toronto, Ont.: Institute for Clinical Evaluative Sciences; 2012. https://www.deslibris.ca/ID/232144. Accessed May 6, 2020.

169. Booth RG, Richard L, Li L, Shariff SZ, Rayner J. Characteristics of health care related to mental health and substance use disorders among Community Health Centre clients in Ontario: A population-based cohort study. *CMAJ Open*. 2020;8(2):E391-E399. doi:10.9778/cmajo.20190089

170. Cameron D, Jones IG. John snow, the broad street pump and modern epidemiology. *Int J Epidemol*. 1983;12(4):393-396. doi:10.1093/ije/12.4.393

171. Thuraisingam S, Chondros P, Dowsey MM, et al. Assessing the suitability of general practice electronic health records for clinical prediction model development: A data quality assessment. *BMC Med Inform Decis Mak*. 2021;21(1):297. doi:10.1186/s12911-021-01669-6

172. Verma AA, Murray J, Greiner R, et al. Implementing machine learning in medicine. *CMAJ*. 2021;193(34):E1351-E1357. doi:10.1503/cmaj.202434

173. Lee S, Xu Y, D'Souza AG, et al. Unlocking the potential of electronic health records for health research. *Int J Popul Data Sci*. 2020;5(1):02. doi:10.23889/ijpds.v5i1.1123

174. Westfall JM, Wittenberg HR, Liaw W. Time to invest in primary care research—commentary on findings from an independent congressionally mandated study. *J Gen Intern Med*. 2021;36(7):2117-2120. doi:10.1007/s11606-020-06560-0

175. ENCODE-FM. Electronic Nomenclature and Classification Of Disorders and Encounters for Family Medicine. *ENCODE-FM*. 2020. http://aix1.uottawa.ca/~fammed/fmcencod.htm. Accessed April 6, 2020.

176. ICD-10 version:2019. World health organization. https://icd.who.int/browse10/2019/en. Published 2020. Accessed April 6, 2020.

177. Benchimol EI, Smeeth L, Guttmann A, et al. The REporting of studies conducted using observational routinely-collected health data (RECORD) statement. *PLoS Med*. 2015;12(10):e1001885. doi:10.1371/journal.pmed.1001885

178. Glazier RH, Rayner J, Kopp A. *Examining Community Health Centres According to Geography and Priority Populations Served, 2011/12 to 2012/13: An ICES Chartbook*. Toronto, Ontario: Institute for Clinical Evaluative Sciences in Ontario; 2015. http://www.deslibris.ca/ID/248807. Accessed January 20, 2020.

179. Fortin M, Almirall J, Nicholson K. Development of a research tool to document self-reported chronic conditions in primary care. *JOC*. 2017;7(1):117-123. doi:10.15256/joc.2017.7.122

180. Lee ES, Lee PSS, Xie Y, Ryan BL, Fortin M, Stewart M. The prevalence of multimorbidity in primary care: A comparison of two definitions of multimorbidity with two different lists of chronic conditions in singapore. *BMC Public Health*. 2021;21(1):1409. doi:10.1186/s12889-021-11464-7

181. Lee YAJ, Xie Y, Lee PSS, Lee ES. Comparing the prevalence of multimorbidity using different operational definitions in primary care in singapore based on a cross-sectional study using retrospective, large administrative data. *BMJ Open*. 2020;10(12):e039440. doi:10.1136/bmjopen-2020-039440

182. Therneau T. *A Package for Survival Analysis in r.*; 2021. https://CRAN.R-project.org/package=survival.

183. Chen H, Cohen P, Chen S. How big is a big odds ratio? Interpreting the magnitudes of odds ratios in epidemiological studies. *Commun Stat Simul Comput*. 2010;39(4):860-864. doi:10.1080/03610911003650383

184. Wang Y-X, Zhang Y-J. Nonnegative matrix factorization: A comprehensive review. *IEEE Trans Knowl Data Eng*. 2013;25(6):1336-1353. doi:10.1109/TKDE.2012.51

185. Pedregosa F, Varoquaux G, Gramfort A, et al. Scikit-learn: Machine learning in python. *J Mach Learn Res*. 2011;12(85):2825-2830. http://jmlr.org/papers/v12/pedregosa11a.html. Accessed January 1, 2022.

186. McDonald JT, Kennedy S. Insights into the "healthy immigrant effect": Health status and health service use of immigrants to canada. *Soc Sci Med*. 2004;59(8):1613-1627. doi:10.1016/j.socscimed.2004.02.004

187. Haneuse S, Daniels M. A general framework for considering selection bias in EHR-based studies: What data are observed and why? *eGEMs*. 2016;4(1). doi:10.13063/2327-9214.1203

188. Chen M, Tan X, Padman R. Social determinants of health in electronic health records and their impact on analysis and risk prediction: A systematic review. *J Am Med Inform Assoc*. 2020;27(11):1764-1773. doi:10.1093/jamia/ocaa143

189. Zhao Y, Wood EP, Mirin N, Cook SH, Chunara R. Social determinants in machine learning cardiovascular disease prediction models: A systematic review. *Am J Prev Med*. 2021;0(0). doi:10.1016/j.amepre.2021.04.016

190. Arbet J, Brokamp C, Meinzen-Derr J, Trinkley KE, Spratt HM. Lessons and tips for designing a machine learning study using EHR data. *J Clin Transl Res*. 2020;5(e21):1-10. doi:10.1017/cts.2020.513

191. CIHI. Patient-reported outcome measures (PROMs). Canadian institute for health information. https://www.cihi.ca/en/patient-reported-outcome-measures-proms. Published 2022. Accessed February 7, 2022.

192. Bagley SC, Altman RB. Computing disease incidence, prevalence and comorbidity from electronic medical records. *J Biomed Inform.* 2016;63:108-111. doi:10.1016/j.jbi.2016.08.005

193. Zhang L, Chen X, Chen T, Wang Z, Mortazavi BJ. DynEHR: Dynamic adaptation of models with data heterogeneity in electronic health records. In: *2021 IEEE EMBS International Conference on Biomedical and Health Informatics (BHI).*; 2021:1-4. doi:10.1109/BHI50953.2021.9508558

194. Alyari F, Jafari Navimipour N. Recommender systems: A systematic review of the state of the art literature and suggestions for future research. *Kybernetes.* 2018;47(5):985-1017. doi:10.1108/K-06-2017-0196

195. Moons KGM, Altman DG, Vergouwe Y, Royston P. Prognosis and prognostic research: Application and impact of prognostic models in clinical practice. *BMJ.* 2009;338. doi:10.1136/bmj.b606

196. O'Caoimh R, Cornally N, Weathers E, et al. Risk prediction in the community: A systematic review of case-finding instruments that predict adverse healthcare outcomes in community-dwelling older adults. *Maturitas.* 2015;82(1):3-21. doi:10.1016/j.maturitas.2015.03.009

197. Goldstein BA, Navar AM, Pencina MJ, Ioannidis JPA. Opportunities and challenges in developing risk prediction models with electronic health records data: A systematic review. *J Am Med Inform Assoc.* 2017;24(1):198-208. doi:10.1093/jamia/ocw042

198. Guthrie B, Boyd CM. Clinical guidelines in the context of aging and multimorbidity. *Public Policy Aging Rep.* 2018;28(4):143-149. doi:10.1093/ppar/pry038

199. Harrell Jr. FE, Lee KL, Mark DB. Multivariable prognostic models: Issues in developing models, evaluating assumptions and adequacy, and measuring and reducing errors. *Stat Med.* 1996;15(4):361-387. doi:10.1002/(SICI)1097-0258(19960229)15:4<361::AID-SIM168>3.0.CO;2-4

200. Alba AC, Agoritsas T, Walsh M, et al. Discrimination and calibration of clinical prediction models: Users' guides to the medical literature. *JAMA.* 2017;318(14):1377-1384. doi:10.1001/jama.2017.12126

201. Van Calster B, McLernon DJ, Smeden M van, et al. Calibration: The achilles heel of predictive analytics. *BMC Medicine.* 2019;17(1):230. doi:10.1186/s12916-019-1466-7

202. Buckland M, Gey F. The relationship between recall and precision. *J Am Soc Inf Sci.* 1994;45(1):12-19. doi:10.1002/(SICI)1097-4571(199401)45:1<12::AID-ASI2>3.0.CO;2-L

203. Sofaer HR, Hoeting JA, Jarnevich CS. The area under the precision-recall curve as a performance metric for rare binary events. *Methods Ecol Evol*. 2019;10(4):565-577. doi:10.1111/2041-210X.13140

204. Altman DG, Royston P. What do we mean by validating a prognostic model? *Statistics in Medicine*. 2000;19(4):453-473. doi:10.1002/(SICI)1097-0258(20000229)19:4<453::AID-SIM350>3.0.CO;2-5

205. Altman DG, Vergouwe Y, Royston P, Moons KGM. Prognosis and prognostic research: Validating a prognostic model. *BMJ*. 2009;338(7708):1432-1435. https://www.jstor.org/stable/25671796. Accessed June 10, 2022.

206. Szklo M, Nieto FJ. *Epidemiology: Beyond the Basics*. Fourth edition. Burlington, Massachusetts: Jones& Bartlett Learning; 2019.

207. Morgenstern JD, Buajitti E, O'Neill M, et al. Predicting population health with machine learning: A scoping review. *BMJ Open*. 2020;10(10):e037860. doi:10.1136/bmjopen-2020-037860

208. Shawe-Taylor J, Cristianini N. *Kernel Methods for Pattern Analysis*. Cambridge University Press; 2004.

209. Conroy B, Xu-Wilson M, Rahman A. Patient similarity using population statistics and multiple kernel learning. In: *Proceedings of Machine Learning for Healthcare 2017*. Vol 68. J Mach Learn Res; 2017.

210. Gonen M, Alpaydın E, Tr BE, Tr BE. Multiple kernel learning algorithms. *J Mach Learn Res*. 2011;12:2211-2268.

211. Fang HSA, Tan NC, Tan WY, Oei RW, Lee ML, Hsu W. Patient similarity analytics for explainable clinical risk prediction. *BMC Medical Informatics and Decision Making*. 2021;21(1):207. doi:10.1186/s12911-021-01566-y

212. Fan X, Chen Z, Zhu L, Liao Z, Fu B. A novel hybrid similarity calculation model. *Sci Program*. 2017;2017:e4379141. doi:10.1155/2017/4379141

213. Lian W, Rai P, Salazar E, Carin L. Integrating features and similarities: Flexible models for heterogeneous multiview data. In: *Proceedings of the Twenty-Ninth AAAI Conference on Artificial Intelligence*.; 2015:7.

214. Hothorn T, Buhlmann P, Kneib T, Schmid M, Hofner B. Model-based boosting 2.0. *J Mach Learn Res*. 2010;11:2109-2113.

215. Hothorn T, Buehlmann P, Kneib T, Schmid M, Hofner B. *Model-Based Boosting*.; 2021. https://github.com/boost-R/mboost.

216. Sigrist F. KTBoost: Combined kernel and tree boosting. *Neural Process Lett*. 2021;53(2):1147-1160. doi:10.1007/s11063-021-10434-9

217. Chen T, Guestrin C. XGBoost: A scalable tree boosting system. In: *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. San Francisco California USA: ACM; 2016:785-794. doi:10.1145/2939672.2939785

218. Klenk S, Dippon J, Fritz P, Heidemann G. Determining patient similarity in medical social networks. In: *MEDEX 2010 Proceedings*.; 2010:6-13.

219. Belanche LA, Villegas MA. Kernel functions for categorical variables with application to problems in the life sciences. In: *Proceedings of the 16th International Conference of the Catalan Association for Artificial Intelligence*. Vol 256. Catalonia, Spain: IOS Press; 2013:171-180. doi:https://doi.org/10.3233/978-1-61499-320-9-171

220. Daemen A, Timmerman D, Van den Bosch T, et al. Improved modeling of clinical data with kernel methods. *Artif Intell Med*. 2012;54(2):103-114. doi:10.1016/j.artmed.2011.11.001

221. Chan LWC, Chan T, Cheng LF, Mak WS. Machine learning of patient similarity: A case study on predicting survival in cancer patient after locoregional chemotherapy. In: *2010 IEEE International Conference on Bioinformatics and Biomedicine Workshops (BIBMW)*.; 2010:467-470. doi:10.1109/BIBMW.2010.5703846

222. Duvenaud D, Lloyd J, Grosse R, Tenenbaum J, Zoubin G. Structure discovery in nonparametric regression through compositional kernel search. In: *Proceedings of the 30th International Conference on Machine Learning*. PMLR; 2013:1166-1174. https://proceedings.mlr.press/v28/duvenaud13.html. Accessed June 3, 2022.

223. Schaid DJ. Genomic similarity and kernel methods I: Advancements by building on mathematical and statistical foundations. *Hum Hered*. 2010;70:109-131. doi:10.1159/000312641

224. Schaid DJ. Genomic similarity and kernel methods II: Methods for genomic information. *Hum Hered*. 2010;70(2):132-140. doi:10.1159/000312643

225. Scalzo F, Liebeskind D, Hu X. Reducing false intracranial pressure alarms using morphological waveform features. *IEEE Trans Biomed Eng*. 2013;60(1):235-239. doi:10.1109/TBME.2012.2210042

226. Zhu J, Hastie T. Kernel logistic regression and the import vector machine. *J Comput Graph Stat*. 2005;14(1):185-205. doi:10.1198/106186005X25619

227. Diamond S, Boyd S. CVXPY: A python-embedded modeling language for convex optimization. *J Mach Learn Res*. 2016;17:1-5. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4927437/. Accessed July 15, 2021.

228. Agrawal A, Verschueren R, Diamond S, Boyd S. A rewriting system for convex optimization problems. *J Control Decis*. 2018;5(1):42-60. doi:10.1080/23307706.2017.1397554

229. Brookhart MA, Stürmer T, Glynn RJ, Rassen J, Schneeweiss S. Confounding control in healthcare database research: Challenges and potential approaches. *Med Care*. 2010;48(60):S114-S120. doi:10.1097/MLR.0b013e3181dbebe3

230. Greenland S, Morgenstern H. Confounding in health research. *Annu Rev Public Health*. 2001;22(1):189-212. doi:10.1146/annurev.publhealth.22.1.189

231. Shrier I, Platt RW. Reducing bias through directed acyclic graphs. *BMC Med Res Methodol*. 2008;8(1):70. doi:10.1186/1471-2288-8-70

232. Austin PC, Tu JV. Automated variable selection methods for logistic regression produced unstable models for predicting acute myocardial infarction mortality. *J Clin Epi*. 2004;57(11):1138-1146. doi:10.1016/j.jclinepi.2004.04.003

233. Morris TP, White IR, Crowther MJ. Using simulation studies to evaluate statistical methods. *Stat Med*. 2019;38(11):2074-2102. doi:10.1002/sim.8086

234. Thrun SB, Bala J, Bloedorn E, et al. *The MONK's Problems a Performance Comparison of Different Learning Algorithms*. Carnegi Mellon University; 1991.

235. Márquez CG. Multivariate kernel functions for categorical variables. 2014. https://upcommons.upc.edu/bitstream/handle/2099.1/24508/99930.pdf.

236. McKerns M, Aivazis M. *Pathos: A Framework for Heterogenous Computing.*; 2010. https://uqfoundation.github.io/project/pathos. Accessed June 23, 2021.

237. McKerns MM, Strand L, Sullivan T, Fang A, Aivazis MAG. Building a framework for predictive science. *arXiv:12021056 [cs]*. 2011. http://arxiv.org/abs/1202.1056. Accessed June 23, 2021.

238. Research. *Alliance for Healthier Communities*. 2020. https://www.allianceon.org/Research. Accessed May 6, 2020.

239. Altschul D, Iveson M, Deary IJ. Generational differences in loneliness and its psychological and sociodemographic predictors: An exploratory and confirmatory machine learning study. *Psychol Med*. 2021;51(6):991-1000. doi:10.1017/S0033291719003933

240. Holt-Lunstad J. The potential public health relevance of social isolation and loneliness: Prevalence, epidemiology, and risk factors. *Public Policy Aging Rep*. 2017;27(4):127-130. doi:10.1093/ppar/prx030

241. Koning JLD, Stathi A, Richards S. Predictors of loneliness and different types of social isolation of rural-living older adults in the united kingdom. *Ageing & Society*. 2017;37(10):2012-2043. doi:10.1017/S0144686X16000696

242. National Academies of Sciences, Engineering, and Medicine. Social isolation and loneliness in older adults: Opportunities for the health care system. In: *Risk and Protective Factors for Social Isolation and Loneliness*. Washington, DC, USA: The National Academies Press; 2020. https://doi.org/10.17226/25663. Accessed February 22, 2022.

243. Nevarez-Flores AG, Breslin M, Carr VJ, et al. Proposing a causal pathway for health-related quality of life in people with psychotic disorders. *J Psychiatr Res*. 2021;138:550-559. doi:10.1016/j.jpsychires.2021.05.016

244. Nicholson NR. A review of social isolation: An important but underassessed condition in older adults. *J Primary Prevent*. 2012;33(2):137-152. doi:10.1007/s10935-012-0271-2

245. World Health Organization. *Social Isolation and Loneliness Among Older People: Advocacy Brief*. Geneva: World Health Organization; 2021. https://apps.who.int/iris/handle/10665/343206. Accessed February 22, 2022.

246. Lim MH, Eres R, Vasan S. Understanding loneliness in the twenty-first century: An update on correlates, risk factors, and potential solutions. *Soc Psychiatry Psychiatr Epidemiol*. 2020;55(7):793-810. doi:10.1007/s00127-020-01889-7

247. Social prescribing. Alliance for healthier communities. https://www.allianceon.org/Social-Prescribing. Published 2022. Accessed April 2, 2022.

248. Nowak DA, Mulligan K. Social prescribing. *Can Fam Physician*. 2021;67(2):88-91. doi:10.46747/cfp.670288

249. Mulligan K, Bhatti S, Rayner J, Hsiung S. Social prescribing: Creating pathways towards better health and wellness. *J Am Geriatr Soc*. 2020;68(2):426-428. doi:10.1111/jgs.16249

250. Kueper JK, Rayner J, Zwarenstein M, Lizotte DJ. Describing a complex primary health care population in a learning health system to support future decision support and artificial intelligence initiatives. *medRxiv*. 2022. doi:10.1101/2022.03.01.22271714

251. Doryab A, Villalba DK, Chikersal P, et al. Identifying behavioral phenotypes of loneliness and social isolation with passive sensing: Statistical analysis, data mining and machine learning of smartphone and fitbit data. *JMIR Mhealth Uhealth*. 2019;7(7):e13209. doi:10.2196/13209

252. Gower JC. A general coefficient of similarity and some of its properties. *Biometrics*. 1971;27(4):857-871. doi:10.2307/2528823

253. Tennant PWG, Murray EJ, Arnold KF, et al. Use of directed acyclic graphs (DAGs) to identify confounders in applied health research: Review and recommendations. *Int J Epidemiol*. 2021;50(2):620-632. doi:10.1093/ije/dyaa213

254. Wang Y, Liang D, Charlin L, Blei DM. The deconfounded recommender: A causal inference approach to recommendation. *arXiv:180806581 [cs, stat]*. August 2018. http://arxiv.org/abs/1808.06581. Accessed November 14, 2018.

255. Wang Y, Blei DM. The blessings of multiple causes. *J Am Stat Assoc*. 2019;114(528):1574-1596. http://arxiv.org/abs/1805.06826. Accessed November 6, 2019.

256. Feng Y, Yang Y, Lv S-G, Zhao Y, Suykens JAK. Learning with kernelized elastic net regularization. 2016:30.

257. Primary health care. World health organization. https://www.who.int/westernpacific/health-topics/primary-health-care. Published 2021. Accessed June 16, 2021.

258. Bickel PJ, Sakov A. ON THE CHOICE OF m IN THE m OUT OF n BOOTSTRAP AND CONFIDENCE BOUNDS FOR EXTREMA. *Stat Sin*. 2008;18:967-985.

259. Chen YT, Jewell SW, Witten DM. More powerful selective inference for the graph fused lasso. *arXiv:210910451 [stat]*. February 2022. http://arxiv.org/abs/2109.10451. Accessed May 20, 2022.

260. Apitz S, Fontoura P. The primary care research landscape and its relationship with clinical practice: A scientometric analysis. *Acta Médica Portuguesa*. 2021;34(13). doi:10.20344/amp.15262

261. Glanville J, Kendrick T, McNally R, Campbell J, Hobbs FR. Research output on primary care in Australia, Canada, Germany, the Netherlands, the United Kingdom, and the United States: Bibliometric analysis. *BMJ*. 2011;342(mar08 2):d1028-d1028. doi:10.1136/bmj.d1028

262. Rayner J, Muldoon L, Bayoumi I, McMurchy D, Mulligan K, Tharao W. Delivering primary health care as envisioned: A model of health and well-being guiding community-governed primary care organizations. *Journal Integr Care*. 2018;26(3):231-241. doi:10.1108/JICA-02-2018-0014

263. Wingrove P, Liaw W, Weiss J, Petterson S, Maier J, Bazemore A. Using machine learning to predict primary care and advance workforce research. *Ann Fam Med*. 2020;18(4):334-340. doi:10.1370/afm.2550

264. Aponte-Hao S, Wong ST, Thandi M, et al. Machine learning for identification of frailty in Canadian primary care practices. *Int J Popul Data Sci*. 2021;6(1). doi:10.23889/ijpds.v6i1.1650

265. Rahimi SA, Légaré F, Sharma G, et al. Application of artificial intelligence in community-based primary health care: Systematic scoping review and critical appraisal. *J Med Internet Res*. 2021;23(9):e29839. doi:10.2196/29839

266. Ronquillo CE, Peltonen L-M, Pruinelli L, et al. Artificial intelligence in nursing: Priorities and opportunities from an international invitational think-tank of the Nursing and Artificial Intelligence Leadership Collaborative. *J Adv Nursing*. 2021;n/a(n/a). doi:https://doi.org/10.1111/jan.14855

267. Yang Z, Silcox C, Sendak M, et al. Advancing primary care with artificial intelligence and machine learning. *Healthcare*. 2022;10(1):100594. doi:10.1016/j.hjdsi.2021.100594

268. Kueper JK, Terry A, Bahniwal R, et al. Connecting artificial intelligence and primary care challenges: Findings from a multi stakeholder collaborative consultation. *BMJ Health Care Inform*. 2022;29(1):e100493. doi:10.1136/bmjhci-2021-100493

269. About Community Health Centres. *International Federation of Community Health Centres*. 2022. https://www.ifchc.org/what-are-chcs/. Accessed May 17, 2022.

270. Yang S, Varghese P, Stephenson E, Tu K, Gronsbell J. *Machine Learning Approaches for Electronic Health Records Phenotyping: A Methodical Review*. medRxiv; 2022:2022.04.23.22274218. doi:10.1101/2022.04.23.22274218

271. Y GDharani, Nair NG, Satpathy P, Christopher J. Covariate shift: A review and analysis on classifiers. In: *2019 Global Conference for Advancement in Technology (GCAT)*.; 2019:1-6. doi:10.1109/GCAT47503.2019.8978471

272. Streiner DL, Norman GR, Cairney J. *Health Measurement Scales: A Practical Guide to Their Development and Use*. Fifth edition. Oxford: Oxford University Press; 2015.

273. Kellogg KC, Sendak M, Balu S. AI on the front lines. *MIT Sloan Management Review*. May 2022. https://sloanreview.mit.edu/article/ai-on-the-front-lines/. Accessed May 17, 2022.

274. Nash DM, Brown JB, Thorpe C, Rayner J, Zwarenstein M. The Alliance for Healthier Communities as a learning health system for primary care: A qualitative analysis in Ontario, Canada. *J Eval Clin Pract*. April 2022:jep.13692. doi:10.1111/jep.13692

275. Bi Q, Goodman KE, Kaminsky J, Lessler J. What is machine learning? A primer for the epidemiologist. *Am J Epidemiol*. 2019;188(12):2222-2239. doi:10.1093/aje/kwz189

276. About WorldPop. *WorldPop*. https://www.worldpop.org/about. Accessed May 17, 2022.

277. Abdeldayem OM, Dabbish AM, Habashy MM, et al. Viral outbreaks detection and surveillance using wastewater-based epidemiology, viral air sampling, and machine learning techniques: A comprehensive review and outlook. *Sci Total Environ*. 2022;803:149834. doi:10.1016/j.scitotenv.2021.149834

278. McCoy LG, Banja JD, Ghassemi M, Celi LA. Ensuring machine learning for healthcare works for all. *BMJ Health Care Inform*. 2020;27(3):e100237. doi:10.1136/bmjhci-2020-100237

279. Gichoya JW, McCoy LG, Celi LA, Ghassemi M. Equity in essence: A call for operationalising fairness in machine learning for healthcare. *BMJ Health Care Inform*. 2021;28(1):e100289. doi:10.1136/bmjhci-2020-100289

280. Rajkomar A, Hardt M, Howell MD, Corrado G, Chin MH. Ensuring fairness in machine learning to advance health equity. *Ann Intern Med*. 2018;169(12):866-872. doi:10.7326/M18-1990

281. Fox MP, Murray EJ, Lesko CR, Sealy-Jefferson S. On the need to revitalize descriptive epidemiology. *Am J Epidemiol*. March 2022:kwac056. doi:10.1093/aje/kwac056

282. Jefferies S, French N, Gilkison C, et al. COVID-19 in New Zealand and the impact of the national response: A descriptive epidemiological study. *Lancet Public Health*. 2020;5(11):e612-e623. doi:10.1016/S2468-2667(20)30225-5

283. Health TP. *Covid-19 and the Social Determinants of Health: What Do We Know?* Toronto Public Health; 2020:4. https://www.toronto.ca/wp-content/uploads/2020/07/956b-SDOHandCOVID19_Summary_2020July1.pdf. Accessed May 17, 2022.

284. AAAI. Unsupervised or Indirectly Supervised Learning. *AITopics*. 2022. https://aitopics.org/search?filters=taxnodes%3ATechnology%7CInformation+Technology%7CArtificial+Intelligence%7CMachine+Learning%40%40taxnodes%3AGenre%7COverview%40%40taxnodes%3ATechnology%7CInformation+Technology%7CArtificial+Intelligence%7CMachine+Learning%7CUnsupervised+or+Indirectly+Supervised+Learning. Accessed May 17, 2022.

285. Pearl J. Theoretical impediments to machine learning with seven sparks from the causal revolution. In: *Proceedings of the Eleventh ACM International Conference on Web Search and Data Mining - WSDM '18*. Marina Del Rey, CA, USA: ACM Press; 2018:3-3. doi:10.1145/3159652.3176182

286. Blakely T, Lynch J, Simons K, Bentley R, Rose S. Reflection on modern methods: When worlds collide—prediction, machine learning and causal inference. *Int J Epidemiol*. 2020;49(6):2058-2064. doi:10.1093/ije/dyz132

287. Rothman KJ, Greenland S. Causation and causal inference in epidemiology. *Am J Public Health*. 2005;95(S1):S144-S150. doi:10.2105/AJPH.2004.059204

288. Singer M, Bulled N, Ostrach B, Mendenhall E. Syndemics and the biosocial conception of health. *Lancet*. 2017;389(10072):941-950. doi:10.1016/S0140-6736(17)30003-X

289. Hastie T, Tibshirani R, Wainwright J. *Statistical Learning with Sparsity. The Lasso and Generalizations*. CRC Press; 2016.

290. Allauzen C, Mohri M, Talwalkar A. Sequence kernels for predicting protein essentiality. In: *Proceedings of the 25th International Conference on Machine Learning - ICML '08*. Helsinki, Finland: ACM Press; 2008:9-16. doi:10.1145/1390156.1390158

291. Nojoomi S, Koehl P. String kernels for protein sequence comparisons: Improved fold recognition. *BMC Bioinform*. 2017;18(1):137. doi:10.1186/s12859-017-1560-9

APPENDICES

Appendix A

List of Abbreviations

AI: Artificial Intelligence

AUPRC: Area Under the Precision Recall Curve

AUROC: Area Under the Receiver Operating Characteristic Curve

CHC: Community Health Centre

CV: Cross Validation

EHR: Electronic Health Record

ENCODE-FM: Electronic Nomenclature and Classification Of Disorders and Encounters for Family Medicine

HFSM: Hybrid Feature- and Similarity-Based Model

HFSM-Seq: Hybrid Feature- and Similarity-Based Model - Sequential

HFSM-Sim: Hybrid Feature- and Similarity-Based Model - Simultaneous

ICD-10: International Classification of Disease - Version 10

KLR: Kernel Logistic Regression

LHS: Learning Health System

LR: Logistic Regression

NMF: Non-Negative Matrix Factorization

PCP: Primary Care Provider

TEM: Technology, Engineering, and Math

UAR: Urban At-Risk

XAI: Explainable Artificial Intelligence

Appendix B

Combined PhD Agreement

Combined PhD degree template - *Epidemiology & Biostatistics* and *Computer Science*

	Program A	Program B	Combined program	comments
Name	Epidemiology & Biostatistics	Computer Science	*Epidemiology & Biostatistics* and *Computer Science*	Transcript will read under "Academic Program History": Program: *Epidemiology & Biostatistics* and *Computer Science* Plan: Combined Doctor of Philosophy *Epidemiology & Biostatistics* and *Computer Science* Status: Active in Program (or later, "Completed Program") Final degree awarded: Combined PhD in *Epidemiology & Biostatistics* and *Computer Science*.; **However, the first student to complete this will receive 2 PhDs.**
Home program	Epidemiology & Biostatistics	Computer Science	Admitting program	Home program is the fall-back program in case combined degree is not working. Choice of home program may also have funding consequences.
Entrance requirements	Thesis-based Master's degree (not necessarily in epidemiology) or equivalent from an accredited university and provide evidence of research potential. Minimum 80% average for all graduate-level coursework and strong performance in a recent statistics course.	Undergraduate or master's degree in computer science	Master's degree or equivalent in one of the two disciplines. Requires wide coverage of CS which may be supplemented with an appropriate portion of the Postgraduate Diploma Program when needed; decided in consultation with advisory committee.	Accepted into home program and then transfer into combined degree program.

Duration of degree	4 years	4 years	4 years, 5 years if additional background required (e.g. via portion of Postgraduate Diploma courses.)	WGRS is paid by the home department. Computer Science will provide two TAs and the supervisor will provide research funding.
Advisory committee structure	Supervisor to be assigned at acceptance. Committee formed by supervisor and student; size dependent on need but requires at least one member of the Department of *Epidemiology & Biostatistics* who holds SGPS Teaching and Advisory Membership.	2 members	At least 3 members. At least one committee member from *Computer Science* and at least one from *Epidemiology & Biostatistics*. (See below for supervisory requirements.)	
Course requirements	Each PhD candidate must complete a total of 4.0 credits, or 2.5 credits if they have previously completed a Master's within the department. For external candidates, of the 4.0 credits, 3.0 are specified required courses, and 1.0 are electives within the department (or	2.0 credits	Satisfy the specified required courses for EpiBio, plus 1.5 graduate credits from CS.	Note that there is typically no undergraduate education in *Epidemiology & Biostatistics*, hence more courses are typically required in the subject. (Note however that UWO has one of the few undergraduate epidemiology programs worldwide.)

	elsewhere with permission.) For candidates with UWO EpiBio Masters, of the 2.5 credits, 0.5 are specified required courses and 2.0 are electives approved by the graduate chair.			
Comprehensive examination(s)	Two-day written examination typically undertaken in Fall of second year. Thesis proposal document and examination.	None; if a student has a gap in knowledge they must take undergrad courses A PhD Research Topics Survey/Proposal (TSP)	*Epidemiology & Biostatistics* comprehensive examination. *Epidemiology & Biostatistics* thesis proposal document and examination. The format, length, and learning objectives of the EpiBio thesis proposal and examination are equivalent to the CS TSP. At least 1 proposal examiner from CS, and at least 1 from EpiBio, plus one additional for total of 3 examiners.	Although format would follow the EpiBio thesis proposal in terms of form and timing, the expectation is that there will be sufficient Computer Science content at the PhD level.
Additional program requirements	English Language Proficiency Certification - TOEFL, IELTS, etc. (program exit requirement; before thesis)	Must present in a public forum at least once a year	Must present in a public forum once a year. EpiBio seminar milestone. All modules required by the School of Graduate Studies and by individual programs	*Epidemiology & Biostatistics* thesis proposal public lecture may count as one of the CS public forum presentations.

	Seminar milestone. Progress reports (end of terms 3 and 5 and every Winter term after term 6.) Thesis proposal public lecture. Health and Safety Awareness Training (Worker), Safe Campus Community, Accessibility in Service or Accessibility in Teaching), introduction to Biomedical research milestone		must be completed, but will not need to be repeated separately for each program.	
Thesis requirements	By convention, the thesis will typically contain content that is equivalent to 2 - 5 research papers in reputable epidemiology, biostatistics, or public health journals. Must meet official SGPS thesis guidelines.	Makes a substantial and novel contribution to research in Computer Science.	Thesis will contain content equivalent to at least 3 research papers and make contributions to both the fields of *Epidemiology & Biostatistics* and *Computer Science*.	Thesis Intent to Submit forms will be submitted to the home department. Coordination of defense examiners, chair, public lecture, and date will be organized by the Academic Programs Coordinator of the home department. A single Certificate of Exam will be issued and signed by the supervisor(s), all thesis examiners, and the chair of the defense.

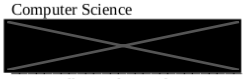
	Monograph or integrated article format. Contain a critical literature review chapter (may count as one of the research papers.)			
Thesis examination committee	Two members of the Program + 1 university examiner + 1 external to Western examiner	Two members of the Program + 1 university examiner + 1 external to Western examiner	3 internal – one from *CS*, one from *Epidemiology & Biostatistics*, one from outside of CS/EpiBio; The (University Examiner) and one external	The U. examiner can be from CS or EpiBio or from a third unit. The external examiner may be from CS or EpiBio or a third discipline provided they have qualifications to examine the thesis.
Supervisor(s)	SGPS PhD supervisory credentials in Epidemiology & Biostatistics department.	SGPS PhD supervisory credentials in Computer Science department.	One from each program or a single supervisor with SGPS supervisory credentials in both departments.	.
Funding	Minimum: Tuition & Fees – WGRS. GRA – supervisor. Additional funding acquired through scholarships and/or GTA positions may reduce the amount of funding from the above sources.	Standard department funding for PhD GRA – supervisor.	Primary funding to come from Home Department as per usual, with non-Home Department offering TA-ships for additional support.	
Teaching assistantship	Optional.	TAs available	TAs available	According to the GTA collective agreement, once *Epidemiology & Biostatistics* PhD students complete a TA-ship, *Epidemiology & Biostatistics* is required to give them the same

				number of hours in all remaining years unless the student turns down a TA position, then the department is not required to offer them TA positions again. Partnership with CS to allow for continuity of TAs and guarantee of funding may be beneficial. (E.g. student does TA in EpiBio, then in CS, would be guaranteed continuing TA funding from one department or the other.)

THE UNIVERSITY OF WESTERN ONTARIO

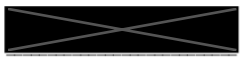
Marc Moreno Maza, Graduate Chair
Computer Science

Date: April 28, 2020



Piotr Wilk, Graduate Chair
Epidemiology & Biostatistics

Date: April 15, 2020



Dr. Linda Miller, Vice-Provost
School of Graduate and Postdoctoral Studies

Date: **April 15, 2020**

Appendix C

Objective 1 Extended Information[a]

[a] A version of this appendix has been published as supplementary material for: Kueper JK, Terry AL, Zwarenstein M, Lizotte DJ. Artificial intelligence and primary care research: a scoping review. *Annals of Family Medicine.* 2020;18(3):250-258. doi:10.1370/afm.2518

C.1 PRISMA-ScR Checklist

SECTION	ITEM	PRISMA-ScR CHECKLIST ITEM	REPORTED ON PAGE #
TITLE			
Title	1	Identify the report as a scoping review.	1
ABSTRACT			
Structured summary	2	Provide a structured summary that includes (as applicable): background, objectives, eligibility criteria, sources of evidence, charting methods, results, and conclusions that relate to the review questions and objectives.	2
INTRODUCTION			
Rationale	3	Describe the rationale for the review in the context of what is already known. Explain why the review questions/objectives lend themselves to a scoping review approach.	1-2
Objectives	4	Provide an explicit statement of the questions and objectives being addressed with reference to their key elements (e.g., population or participants, concepts, and context) or other relevant key elements used to conceptualize the review questions and/or objectives.	2
METHODS			
Protocol and registration	5	Indicate whether a review protocol exists; state if and where it can be accessed (e.g., a Web address); and if available, provide registration information, including the registration number.	2
Eligibility criteria	6	Specify characteristics of the sources of evidence used as eligibility criteria (e.g., years considered, language, and publication status), and provide a rationale.	2,3
Information sources*	7	Describe all information sources in the search (e.g., databases with dates of coverage and contact with authors to identify additional sources), as well as the date the most recent search was executed.	2
Search	8	Present the full electronic search strategy for at least 1 database, including any limits used, such that it could be repeated.	Supplemental Appendix 2
Selection of sources of evidence†	9	State the process for selecting sources of evidence (i.e., screening and eligibility) included in the scoping review.	2,3
Data charting process‡	10	Describe the methods of charting data from the included sources of evidence (e.g., calibrated forms or forms that have been tested by the team before their use, and whether data charting was done independently or in duplicate) and any processes for obtaining and confirming data from investigators.	3
Data items	11	List and define all variables for which data were sought and any assumptions and simplifications made.	3, Supplemental Appendix 3
Critical appraisal of individual	12	If done, provide a rationale for conducting a critical appraisal of included sources of evidence; describe the methods used and how this	NA

SECTION	ITEM	PRISMA-ScR CHECKLIST ITEM	REPORTED ON PAGE #
sources of evidence§		information was used in any data synthesis (if appropriate).	
Synthesis of results	13	Describe the methods of handling and summarizing the data that were charted.	3
RESULTS			
Selection of sources of evidence	14	Give numbers of sources of evidence screened, assessed for eligibility, and included in the review, with reasons for exclusions at each stage, ideally using a flow diagram.	3, Figure 1
Characteristics of sources of evidence	15	For each source of evidence, present characteristics for which data were charted and provide the citations.	3-5; Table 1; Figures 2 – 4, Supplemental Appendix 3
Critical appraisal within sources of evidence	16	If done, present data on critical appraisal of included sources of evidence (see item 12).	NA
Results of individual sources of evidence	17	For each included source of evidence, present the relevant data that were charted that relate to the review questions and objectives.	NA (too many studies to do in a meaningfully interpretable way)
Synthesis of results	18	Summarize and/or present the charting results as they relate to the review questions and objectives.	3-5; Table 1; Figures 2 – 4; Supplemental Appendix 3
DISCUSSION			
Summary of evidence	19	Summarize the main results (including an overview of concepts, themes, and types of evidence available), link to the review questions and objectives, and consider the relevance to key groups.	5,6
Limitations	20	Discuss the limitations of the scoping review process.	6
Conclusions	21	Provide a general interpretation of the results with respect to the review questions and objectives, as well as potential implications and/or next steps.	5,6,7
FUNDING			
Funding	22	Describe sources of funding for the included sources of evidence, as well as sources of funding for the scoping review. Describe the role of the funders of the scoping review.	7

JBI = Joanna Briggs Institute; PRISMA-ScR = Preferred Reporting Items for Systematic reviews and Meta-Analyses extension for Scoping Reviews. * Where *sources of evidence* (see second footnote) are compiled from, such as bibliographic databases, social media platforms, and Web sites. † A more inclusive/heterogeneous term used to account for the different types of evidence or data sources (e.g., quantitative and/or qualitative research, expert opinion, and policy documents) that may be eligible in a scoping review as opposed to only studies. This is not to be confused with *information sources* (see first footnote). ‡ The frameworks by Arksey and O'Malley (6) and Levac and colleagues (7) and the JBI guidance (4, 5) refer to the process of data extraction in a scoping review as data charting. § The process of systematically examining research evidence to assess its validity, results, and relevance before using it to inform a decision. This term is used for items 12 and 19 instead of "risk of bias" (which is more applicable to systematic reviews of interventions) to include and acknowledge the various sources of evidence that may be used in a scoping review (e.g., quantitative and/or qualitative research, expert opinion, and policy document).

From: Tricco AC, Lillie E, Zarin W, O'Brien KK, Colquhoun H, Levac D, et al. PRISMA Extension for Scoping Reviews (PRISMA-ScR): Checklist and Explanation. Ann Intern Med. ;169:467–473. doi: 10.7326/M18-0850

C.2 Search Strategies

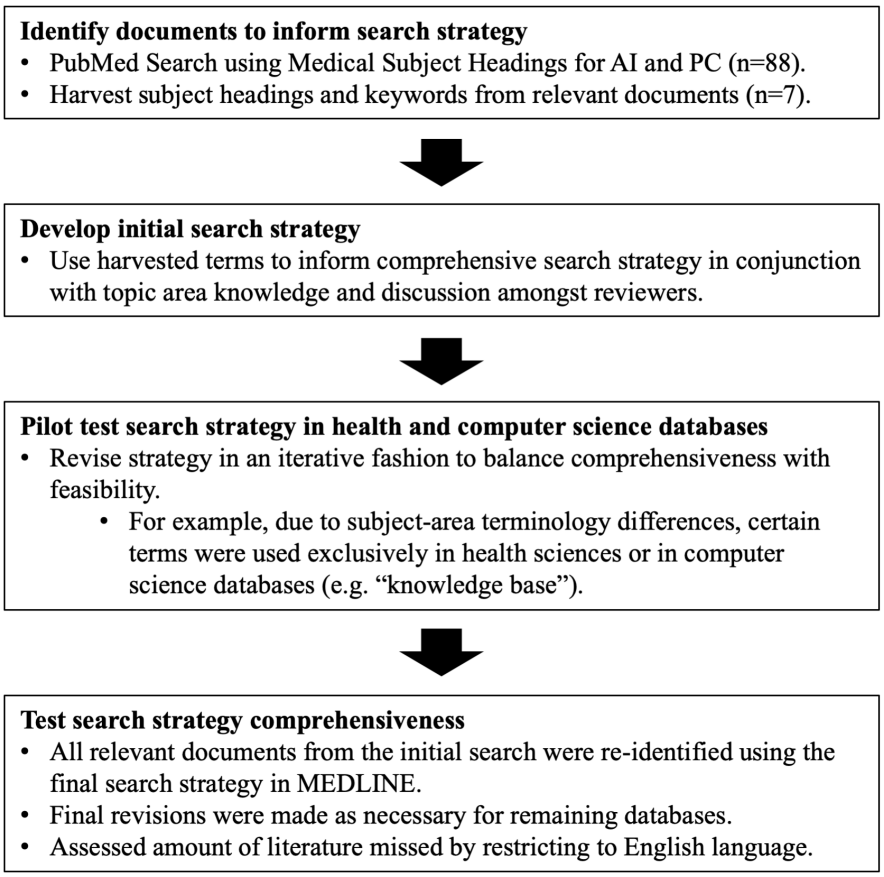


Figure C.1: Development of search strategies.
Legend: AI = Artificial Intelligence; PC = Primary Care.

Databases used:

i. Medline-OVID
ii. EMBASE
iii. Cinahl
iv. Cochrane Library
v. Web of Science
vi. Scopus
vii. IEEE Xplore
viii. ACM Digital Library
ix. MathSciNet
x. AAAI (https://aaai.org/ocs/index.php/index/index/search/advanced)
xi. arXiv

Database searching notes:

The databases listed above have different search capabilities in terms of keywords and

subject headings. We used the most rigorous approach possible for each database, whereby approaches can be broken down into three general categories:

1) Search with keywords and subject headings:

- Medline-OVID
  - Keywords were used to search title, abstract, and author keywords.
- Embase
  - Keywords were used to search title, abstract, and author keywords.
- Cinahl
  - Keywords were used to search title and abstract.
- Cochrane
  - Keywords were used to search title, abstract, and keywords.
- ACM Digital Library
  - No wildcard (*), use full spellings.
  - Only used artificial intelligence subject headings (CCS); health related headings are too broad and captured too many irrelevant documents to maintain review feasibility.
  - Search "The ACM Full-Text Collection."

Search strategy for category 1: i) Keywords and subject headings for artificial intelligence concept were searched with OR. ii) Keywords and subject headings for primary care concept were searched with OR. iii) i) and ii) were combined with AND.

2) Search with keywords only:

- Web of Science
  - Keywords in "Topic" field were used to search title, abstract, author keywords, and keywords plus.
- Scopus
  - Keywords were used to search title, abstract, and keywords.
- MathSciNet
  - Keywords in 'Anywhere' field were used to search author, author/related, title, review text, journal, institution code, series, MSC primary/secondary, MSC primary, MR number, and reviewer.
- arXiv
  - arXiv API was accessed using python.
  - Keywords in 'all' field were used to search title, author, abstract, comment, journal reference, subject category, report number, and id.

Search strategy for category 2: i) Keywords for artificial intelligence concept were searched with OR. ii) Keywords for primary care concept were searched with OR. iii) i) and ii) were combined with AND.

3) Search with limited keywords only:

- IEEE Xplore
  - Limited to 12 keywords.

  - Keywords were used to search metadata (abstract, index terms, bibliographic citation data.)
  - No wildcard (*) within phrase searching, so we wrote out the 12 terms in full.
- AAAI
  - Limited to 254 characters.
  - Used only primary care concept keywords because redundant to search artificial intelligence terms in artificial intelligence proceedings.
  - Case sensitive; spelled out keywords that are most important and most likely to be capitalized differentially by different authors using upper and lower case first letter(s) and spelled out less important or less capitalization-ambiguous keywords using only lower case.

Search strategy for IEEE Xplore: i) Keywords for artificial intelligence concept were searched with OR. ii) Keywords for primary care concept were searched with OR. iii) i) and ii) were combined with AND.

Search strategy for AAAI: i) Keywords for primary care concept were searched with OR.

Table C.1: Search terms for health sciences databases.

Concept	**Key Words** (syntax for Medline-OVID and EMBASE)	**Medline-Ovid**	**EMBASE**	**CINAHL**	**Cochrane Library** (default explods subject headings)
Artificial Intelligence	(Artificial Intelligence OR Computer Heuristics OR Expert System* OR Fuzzy Logic OR Machine Learning OR Support Vector Machine OR Natural Language Processing OR Neural Network* OR Robotic* OR Deep Learning OR Knowledge Representation OR Automated Reasoning OR Computer Vision OR Data Mining OR Bayesian Network* OR Bayes Network*).ti,ab,kw.	exp Artificial Intelligence/ OR Data Mining/ OR exp Decision Making, Computer Assisted/ OR exp Decision Support Techniques/	Exp Artificial Intelligence/ OR Expert System/ OR Fuzzy Logic/ OR Exp Machine Learning/ OR Natural Language Processing/ OR Robotics/ OR Computer Assisted Diagnosis/ OR Exp Computer Assisted Therapy/ OR Knowledge Base/ OR Knowledge Base/ OR Ontology Development/	(MH "Artificial Intelligence+") OR (MH "Data Mining") OR (MH "Decision Making, Computer Assisted") OR (MH "Diagnosis, Computer Assisted+") OR (MH "Therapy, Computer Assisted+") OR (MH "Decision Support Techniques+")	[mh "Artificial Intelligence"] OR [mh "Decision Making, Computer Assisted"] OR [mh "Decision Support Techniques"] OR [mh "Data Mining" not exploded; separate line]

Primary Care	(Primary Care OR Primary Health Care OR Primary Healthcare OR Primary Medical Care OR Family Medicine OR Family Healthcare OR Family Health Care OR Family Physician* OR Family Pract* OR General Practitioner* OR Nurse Practitioner* OR Family Doctor* OR Family Nurse* OR Community Medicine OR Community Pract* OR Ambulatory Care).ti,ab,kw.	Primary Health Care/ OR Physicians, Family/ OR Physicians, Primary Care/ OR General Practitioners/ OR exp General Practice/ OR Community Medicine/ OR Nurse Practitioners/ OR Family Nurse Practitioners/ OR Primary Care Nursing/ OR Nurses, Community Health/ OR Ambulatory Care/	Exp Primary Health Care/ OR Family Medicine/ OR Community Medicine/ OR Family Health/ OR General Practitioner/ OR General Practice/ OR Ambulatory Care/ OR Ambulatory Care Nursing/ OR Nurse Practitioner/ OR Family Nurse Practitioner/	(MH “Primary Health Care") OR (MH “Physicians, Family") OR (MH “Family Practice") OR (MH “Community Medicine") OR (MH “Community Health Centers") OR (MH “Nurse Practitioners") OR (MH “Family Nurse Practitioners") OR (MH “Ambulatory Care") OR (MH “Ambulatory Care Nursing") OR (MH “Ambulatory Care Facilities")	[mh “Primary Health Care"] OR [mh “Physicians, Primary Care"] OR [mh “Primary Care Nursing"] OR [mh “Physicians, Family"] OR [mh “General Practitioners"] OR [mh “General Practice"] OR [mh “Community Medicine"] OR [mh “Nurse Practitioners"] OR [mh “Family Nurse Practitioners"] OR [mh “Nurses, Community Health"] OR [mh “Ambulatory Care"]

Note: Keywords from the above “health databases" were used for Scopus and Web of Science.

Table C.2: Search terms for computer science databases.

Concept	**Key Words** (MathSciNew syntax)	**ACM Digital Library**
Artificial Intelligence	(“Artificial Intelligence" OR “Computer Heuristics" OR “Expert System" OR “Fuzzy Logic" OR “Knowledge Base" OR “Machine Learning" OR “Natural Language Processing" OR “Support Vector Machine" OR “Neural Network" OR “Robotic" OR “Deep Learning" OR “Knowledge Representation" OR “Automated Reasoning" OR “Computer Vision" OR “Data Mining" OR “Bayesian Network" OR “Bayes Network")	“Artificial Intelligence" “Robotic Planning" “Distributed Artificial Intelligence" “Computer Vision" “Machine Learning" “Machine Learning Algorithms"
Primary Care	(“Primary Care" OR “Primary Health Care" OR “Primary Healthcare" OR “Primary Medical Care" OR “Family Medicine" OR “Family Healthcare" OR “Family Health Care" OR “Family Physician" OR “Family Pract" OR “General Practitioner" OR “Nurse Practitioner" OR “Family Doctor" OR “Family Nurse" OR “Community Medicine" OR “Community Pract" OR “Ambulatory Care")	None.

Additional Search Strings:

CINAHL & Cochrane Library keyword syntax (all Table 1 keywords) to be combined with subject headings:

("Artificial Intelligence" OR "Computer Heuristics" OR "Expert System*" OR "Fuzzy Logic" OR "Machine Learning" OR "Support Vector Machine" OR "Natural Language Processing" OR "Neural Network*" OR "Robotic*" OR "Deep Learning" OR "Knowledge Representation" OR "Automated Reasoning" OR "Computer Vision" OR "Data Mining" OR "Bayesian Network*" OR "Bayes Network*")

("Primary Care" OR "Primary Health Care" OR "Primary Healthcare" OR "Primary Medical Care" OR "Family Medicine" OR "Family Healthcare" OR "Family Health Care" OR "Family Physician*" OR "Family Pract*" OR "General Practitioner*" OR "Nurse Practitioner*" OR "Family Doctor*" OR "Family Nurse*" OR "Community Medicine" OR "Community Pract*" OR "Ambulatory Care")

Web of Science syntax (use advanced search page; all Table 1 keywords):

Line 1: TS=("Artificial Intelligence" OR "Computer Heuristics" OR "Expert System*" OR "Fuzzy Logic" OR "Machine Learning" OR "Support Vector Machine" OR "Natural Language Processing" OR "Neural Network*" OR "Robotic*" OR "Deep Learning" OR "Knowledge Representation" OR "Automated Reasoning" OR "Computer Vision" OR "Data Mining" OR "Bayesian Network*" OR "Bayes Network*") AND LANGUAGE: (English)

Line 2: TS=("Primary Care" OR "Primary Health Care" OR "Primary Healthcare" OR "Primary Medical Care" OR "Family Medicine" OR "Family Healthcare" OR "Family Health Care" OR "Family Physician*" OR "Family Pract*" OR "General Practitioner*" OR "Nurse Practitioner*" OR "Family Doctor*" OR "Family Nurse*" OR "Community Medicine" OR "Community Pract*" OR "Ambulatory Care") AND LANGUAGE: (English)

Line 3: #2 AND #1

Scopus search syntax (all Table 1 keywords):

(TITLE-ABS-KEY (("Artificial Intelligence" OR "Computer Heuristics" OR "Expert System*" OR "Fuzzy Logic" OR "Machine Learning" OR "Support Vector Machine" OR "Natural Language Processing" OR "Neural Network*" OR "Robotic*" OR "Deep Learning" OR "Knowledge Representation" OR "Automated Reasoning" OR "Computer Vision" OR "Data Mining" OR "Bayesian Network*" OR "Bayes Network*") AND ("Primary Care" OR "Primary Health Care" OR "Primary Healthcare" OR "Primary Medical Care" OR "Family Medicine" OR "Family Healthcare" OR "Family Health Care" OR "Family Physician*" OR "Family Pract*" OR "General Practitioner*" OR "Nurse Practitioner*" OR "Family Doctor*" OR "Family Nurse*" OR "Community Medicine" OR "Community Pract*" OR "Ambulatory Care"))) AND (LIMIT-TO (LANGUAGE, "English"))

ACM Digital Libraries syntax (all Table 2 keywords and subject headings):

+("Artificial Intelligence" "Computer Heuristics" "Expert Systems" "Fuzzy Logic" "Knowledge Base" "Machine Learning" "Natural Language Processing" "Support Vector Machine" "Neural Network" "Robotic" "Deep Learning" "Knowledge Representation" "Automated Reasoning" "Computer Vision" "Bayesian Network" "Bayes Network" (+acmdlCCS:("Artificial Intelligence" "Robotic planning" "Distributed Artificial Intelligence" "Computer Vision" "Machine Learning" "Machine Learning Algorithms"))) +("Primary Care" "Primary Health Care" "Primary Healthcare" "Primary Medical Care" "Family Medicine" "Family Healthcare" "Family Health Care" "Family Physician" "Family Practice" "Family Practitioner" "General Practitioner" "Nurse Practitioner" "Community Medicine" "Community Practice" "Ambulatory Care" "Family Doctor" "Family Nurse")

IEEE syntax (use Command Search, metadata only; subset of Table 2 keywords (database limit is 12)):

((“Artificial Intelligence” OR “Machine Learning” OR “Data Mining” OR “Natural Language Processing”) AND (“Primary Care” OR “Primary Health Care” OR “Primary Healthcare” OR “Family Physician” OR “General Practitioner” OR “Family Doctor” OR “Nurse Practitioner” OR “Family Medicine”))

AAAI syntax (use ’search all categories for” line at https://aaai.org/ocs/index.php/index/index/search/advanced; subset of Table 2 keywords (254 character limit)):

“Primary Care” OR “primary care” OR “Primary Health Care” OR “primary health care” OR “Primary Healthcare” OR “primary healthcare” OR “family physician” OR “general practitioner” OR “family doctor” OR “nurse practitioner” OR “family medicine”

arXiv API access python code (adapted from https://arxiv.org/help/api/user-manual#Architecture; all Table 2 keywords):

```
import urllib

url = 'https://export.arxiv.org/api/query?search_query=all:%28%22artificial+intelligence
    %22+OR+%22computer+heuristics%22+OR+%22␣expert+system\*%22+OR+%22fuzzy+logic%22+OR+%22
    knowledge+base%22+OR+%22machine+learning%22+OR+%22natural+la␣nguage+processing%22+OR
    +%22support+vector+machine%22+OR+%22neural+network\*%22+OR+%22robotic\*%22+OR+%22deep+
    learning%22+OR+%22knowledge+representation%22+OR+%22automated+reasoning%22+OR+%22
    computer+vision%22+OR+%22data+mining%22+OR+%22bayesian+network\*%22+OR+%22bayes+
    network\*%22%29+AND+all:%28%22primary+care%22+OR+%22primary+health+care%22+OR+%22
    primary+healthcare%22+OR+%22primary+medical+care%22+OR+%22family+medicine%22+OR+%22
    family+healthcare%22+OR+%22family+health+care%22+OR+%22family+physician\*%22+OR+%22
    family+pract\*␣%22+OR+%22general+practitioner\*%22+OR+%22nurse+practitioner\*%22+OR
    +%22family+doctor\*%22+OR+%22family+nurse\*␣%22+OR+%22community+medicine%22+OR+%22
    community+pract\*%22+OR+%22ambulatory+care%22%29&start=0&max_resul␣ts=2000'

data = urllib.urlopen(url).read()

print data
```

C.3 Additional Methods and Results

Table C.3: Data extraction field characterizations.

Field	Definition and Subfields
Citation Information	Last and first name of first author, year of publication, and title
Study Purpose(s)	Three mutually exclusive overall research purposes for AI in the study: 1. *Method Development/Adaptation*: Research that created novel AI methods or modified existing AI methods to accomplish a task relevant to PC. For example, developing a new supervised machine learning algorithm to learn a model that will predict the probability of pathological heart murmurs using digital heart sound recording data [1]. This category includes studies that compare the performance of AI methods to the performance of humans or that include consultation with end users to inform tool development, as this is considered part of model testing, which may lead to further modifications before evaluating performance in the setting that the AI is intended to support (e.g. clinical practice.) 2. *Data Analysis*: Existing AI methods were used to analyze and/or extract information from data. For example, using natural language processing algorithms to identify cases of familial hypercholesteremia from electronic health records [2]. 3. *Evaluation*: Research that included AI implemented in its intended setting, possibly as part of a pilot study to assess impact or usability characteristics of a tool. For example, assessing the impact of a machine learning-derived diagnosis model on reducing cervical intraepithelial neoplasia overdiagnosis in a Dutch national population screening program [3].
Author Appointment(s)	Author affiliations as presented on the manuscript, divided into 16 categories: *Biological and Biomedical Sciences, Company, Computer Science, Engineering, Epidemiology and Biostatistics, Health Sciences, Informatics, Mathematics, Medicine – Unspecified, Medicine – Specialty, Family Medicine and Primary Care, Nursing, Public Health, Statistics, Other (specified)*, and *Unknown* when not enough affiliation information was provided to identify a broad discipline. When an author had multiple affiliations, all were recorded.

Primary Care Function(s)	Nine categories of PC functions or tasks that the researched AI supported or is intended to support in the future: 1. *Diagnostic Decision Support*: AI provided information to inform diagnosis, such as the probability that a patient has a particular condition. 2. *Treatment Decision Support*: AI provided information to inform treatment decisions, whereby treatment was interpreted broadly to include any management or care provided (or absence of unnecessary actions) to someone with the health condition(s) or symptom(s) of interest. 3. *Referral Support*: AI provided information to support decisions about referring patients to specialist services or AI assisted with technical aspects of the referral process. 4. *Future State Prediction*: AI provided predictions towards future events, for example utilization of emergency department, development of a health condition, or prognosis for an existing condition. 5. *Health Care Utilization Analyses*: AI provided information about interactions with or processes within health care systems, for example frequency or quantity of patient visits. 6. *Knowledge Base and Ontology Construction or Use*: Construction or use of knowledge bases or ontologies including PC concepts. 7. *Information Extraction*: AI used to extract knowledge from structured or unstructured data (e.g. electronic medical records) for further use. 8. *Descriptive Information Provision*: AI used to summarize data in a meaningful way for human interpretation, for example prevalence of a condition or patterns of patient profiles. 9. *Other (specified)*: The PC function was not represented by the above categories; specifics were recorded.
Author Reported Intended End-User(s)	People who the research or research end-product was stated as intended for, regardless of whether those intended end users were involved with the research or how close the research was to being applicable for those users in practice setting: *Patient, Physician, Nurse, Nurse Practitioner, Administrator, Researcher, Other (specified), or Unknown.* If the study was developing a deployable AI method or tool (broadly defined) but more research was needed before the AI method of interest would be ready to implement or be utilized by its intended end user, *Researcher* was included as a target end user.
Target Health Condition(s)	The health condition of interest as stated by the study authors or inferred by reviewers, or *Unknown* if no condition was stated or inferable. Conditions were extracted in full form and MZ later organized them into 27 and 10 category formats. When a study intended for AI to be applicable for all health conditions *"General"* was used; specifics about any test conditions were also extracted.
Location of Data Source(s) or intended location of implementation	Country or next level of granularity where data were collected, or the geographical location where the study stated implementation would occur. *Unknown* was used when the location of data source was not stated or when all data were simulated.

Subfield(s) of Artificial Intelligence

Artificial Intelligence methods were organized according to 10 subfields; a single study may include one or more subfields: 1. *Bayesian Network*: Graphical models (directed acyclic graphs) used to describe dependency relationships among variables that enable the efficient representation of multivariate probability distributions. The resulting distributions can be queried to find the probability of an event occurring given a particular set of evidence. Bayesian networks can be developed manually, such as from physician input, learned from data, or created using a combination of the two. For example, Teles et al. (2015) use a Bayesian Network to assist the diagnosis of dengue fever disease. The model includes variables for dengue fever risk factors, such as 'Respiratory Distress'. For prediction, a person's current risk factor variable values are inputted and the conditional probability they have dengue fever is outputted [4]. 2. *Computer Vision*: Includes extracting visual information and understanding it. Computer vision is distinct from image processing, which includes modifying an existing or creating a new image without focusing on the meaning of the image. For example, Zouridakis et al. (2015) present a smartphone app whereby a picture of a skin lesion is taken and computer vision is used to interpret the image and assess the likelihood of malignancy [5]. 3. *Data Mining*: The process of eliciting information from collections of data, such as by finding and counting pattern occurrences using inferential algorithms; humans may then interpret these patterns. For example, Soler et al. (2015) used data mining on electronic medical records to identify relationships between reasons for encounter and diagnoses recorded for the corresponding visit [6]. We did not consider extracting information in a structured way, such as using a database query to get a basic count of disease X diagnoses, to be the type of data mining that falls under the umbrella of artificial intelligence. 4. *Expert System*: Consists of two parts: 1) a knowledge base that contains facts and rules, such as if-then statements derived from medical guidelines and 2) an inference engine that uses the knowledge base to arrive at conclusions or answers to questions. For example, Lange et al. (1997) demonstrate the use of an expert system called Iliad for teaching diagnostic reasoning to Nurse Practitioner students [7]. Illiad's knowledge base is made up of medical facts and relationships. Bayesian or probabilistic and Boolean or deterministic reasoning may be used with the knowledge base to arrive at a level of confidence about a diagnosis [7]. 5. *Fuzzy Models*: Rely on fuzzy logic and fuzzy sets to represent problems with uncertainty. They are often used to provide more flexibility to outcomes instead of requiring strict classification into pre-defined groups. For example, Katigari et al. (2017) used a fuzzy model as the inference engine for an expert system designed to support diagnosis of diabetic neuropathy.85 Model input includes parameters such as time with diabetes, symptom severity, and laboratory blood test values; model output is an estimate of diabetic neuropathy severity [8]. 6. *Natural Language Processing*: The ability to read language used by humans and interpret it in a meaningful way; this is often accomplished by analysing syntactic and semantic characteristics of language. The input language may be audio or written. For example, Koeling et al. (2011) used natural language processing to analyse free text portions of medical records and enhance the accuracy of ovarian cancer symptom detection compared to only using the structured portion of medical records [9]. 7. *Robotics*: Robotics within artificial intelligence refers to machines that can act autonomously to navigate and alter their environment. A robot may rely on other types of artificial intelligence, such as computer vision and natural language processing, to accomplish this. Robotics outside of artificial intelligence include machines that are programmed by humans to perform a defined set of actions. No examples of robotics were captured by our review.

Subfield(s) of Artificial Intelligence (cont.)	8. *Supervised Machine Learning*: Involves an algorithm learning to associate labels with observations. In the context of health, the label is often an outcome, e.g. a disease state or outcome, and the observations are often patient variables. Labels may be numeric values or categorical classifications. Supervised machine learning uses existing labelled data which contain a collection of observations together with their correct label to produce a model that is able to assign a label to new, previously unseen observations. Supervised machine learning techniques include Support Vector Machines, K-Nearest Neighbours, Naïve Bayes Class ifier, and Random Forest Decision Trees. For example, Cox et al. (2016) used supervised machine learning to help identify undiagnosed post-stroke spasticity [10]. A model was trained using a large PC database that included the outcome of interest, post stroke spasticity events, and 72 candidate variables to predict the outcome, such as demographic information, prescriptions, and medical diagnoses [10]. They then used the model to identify people who had a high probability of post-stroke spasticity and checked whether the event was recorded in their records; the results of this study suggested an under recording of post stroke spasticity in PC records [10]. 9. *Unsupervised Machine Learning*: Algorithms learn patterns from unlabelled data (unlabelled meaning there are not defined, known outcome categories as was the case for supervised machine learning). Common unsupervised machine learning techniques include clustering data items into groups based on their similarity, association mining to identify observations that tend to occur together, autoencoders to condense data while maintaining adequate fidelity, and feature separation to examine different aspects of a dataset independently. For example, Newcomer et al. (2011) used cluster analysis on data from a health care organization to identify groups of complex patients who may benefit from targeted care strategies [11]. 10. *Other (specified)*: There are additional types of AI not captured by the above, such as multi-agent systems. We did not expect a high prevalence of these methods so did not create distinct categories, but recorded details when they arose. Other was also used to classify studies that did not focus on any specific technique of artificial intelligence. For example, Sola et al. (2018) studied physician perceptions of artificial intelligence in general without isolating any particular artificial techniques [12].
Reviewer who extracted the data	Initials of the person who (re)read the full text article and assigned values for the seven key characteristics outlined above: JKK, ALT, or DJL.
Reviewer notes	Optional free form notes from the person extracting the data.

Note: Subfields are ordered according to appearance in the results section of the manuscript. Cited examples are from studies captured by our scoping review. *Legend:* AI = Artificial Intelligence; PC = Primary Care. *References:* 1. Andrisevic N, Ejaz K, Rios-Gutierrez F, Alba-Flores R, Nordehn G, Burns S. Detection of heart murmurs using wavelet analysis and artificial neural networks. *J Biomech Eng.* 2005;127(6):899-904. 2. Safarova MS, Liu H, Kullo IJ. Rapid identification of familial hypercholesterolemia from electronic health records: The SEARCH study. *J Clin Lipidol.* 2016;10(5):1230-1239. 3. Kok MR, Boon ME, Schreiner-Kok PG, Hermans J, Grobbee DE, Kok LP. Less medical intervention after sharp demarcation of grade 1-2 cervical intraepithelial neoplasia smears by neural network screening. *Cancer.* 2001;93(3):173-178. 4. Teles G, Oliveira C, Braga R, et al. Using Bayesian networks to improve the decision- making process in public health systems. In *2014 IEEE 16th International Conference on e-Health Networking, Applications and Services (Healthcom).* 2014:565-570. 5. Zouridakis G, Wadhawan T, Situ N, et al. Melanoma

and other skin lesion detection using smart handheld devices. *Methods Mol Bio.* 2015;1256(bu3, 9214969):459-496. 6. Soler JK, Corrigan D, Kazienko P, et al. Evidence-based rules from family practice to inform family practice; The learning healthcare system case study on urinary tract infections. *BMC Fam Pract.* 2015;16(1). 7. Lange LL, Haak SW, Lincoln MJ, et al. Use of Iliad to Improve Diagnostic Performance of Nurse Practitioner Students. *J Nurs Educ.* 1997;36(1):36-45. 8. Katigari MR, Ayatollahi H, Malek M, Haghighi MK. Fuzzy expert system for diagnosing diabetic neuropathy. *World J Diabetes.* 2017;8(2):80-88. 9. Koeling R, Tate AR, Carroll JA. Automatically estimating the incidence of symptoms recorded in GP free text notes. In: *Proceedings of the first international workshop on Managing interoperability and complexity in health systems.* 2011. 10. Cox AP, Raluy-Callado M, Wang M, Bakheit AM, Moore AP, Dinet J. Predictive analysis for identifying potentially undiagnosed post-stroke spasticity patients in United Kingdom. *J Biomed Inform.* 2016;60:328-333. 11. Newcomer SR, Steiner JF, Bayliss EA. Identifying Subgroups of Complex Patients With Cluster Analysis. *Am J Manag Care.* 2011;17(8):E324-E332. 12. Sola D, Borioli GS, Quaglia R. Predicting GPs' engagement with artificial intelligence. *Br J Health Care Manag.* 2018;24(3):134-140.

Table C.4: Complete author appointment counts.

Appointment Type	Number of studies with at least one author with the corresponding appointment
Biological and Medical Sciences	29
Company	49
Computer Science	97
Engineering	71
Epidemiology and Biostatistics	23
Health Sciences	33
Informatics	63
Mathematics	16
Medicine – Other	94
Medicine – Specialty	99
Medicine – Family or Primary Care	57
Nursing	4
Public Health	20
Statistics	15
Other	132
Unknown	110

Note: Each study fulfills one or more appointment type categories; each category is counted a maximum of one time for any given study.

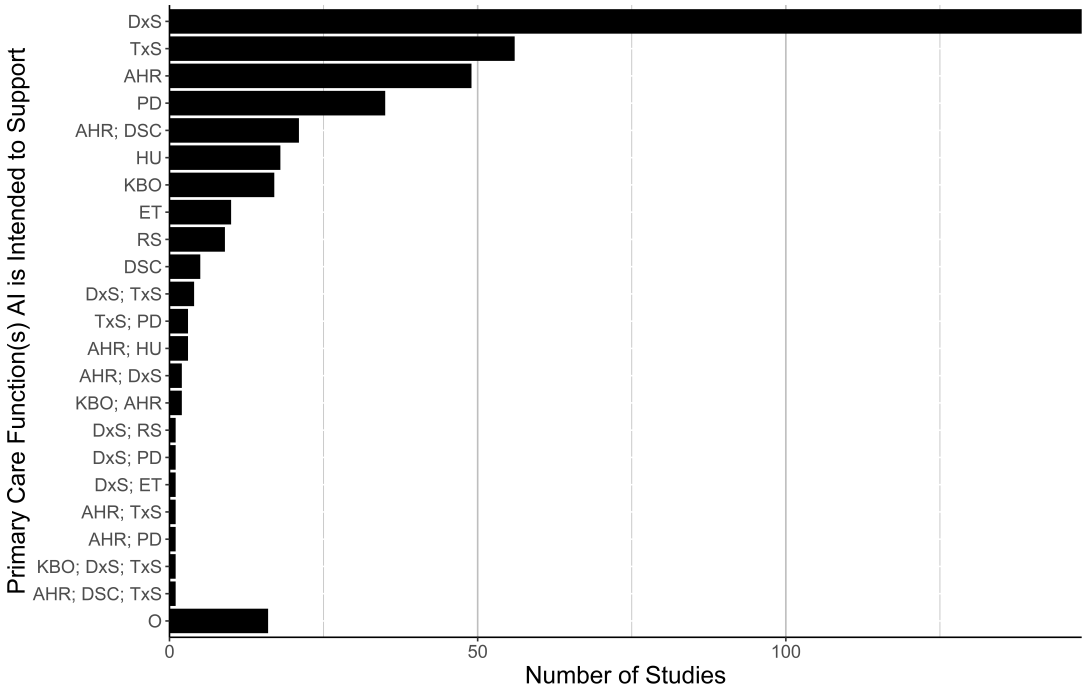


Figure C.2: Detailed breakdown of primary care functions.

Legend: AHR = Analyze Health Records; DCS = Descriptives; ET = Education or Training; KBO = Knowledge Base or Ontology Construction; PD = Prediction; DxS = Diagnostic Decision Support; RS = Referral Support; TxS = Treatment Decision Support; HU = Health Care Use; O = Other.

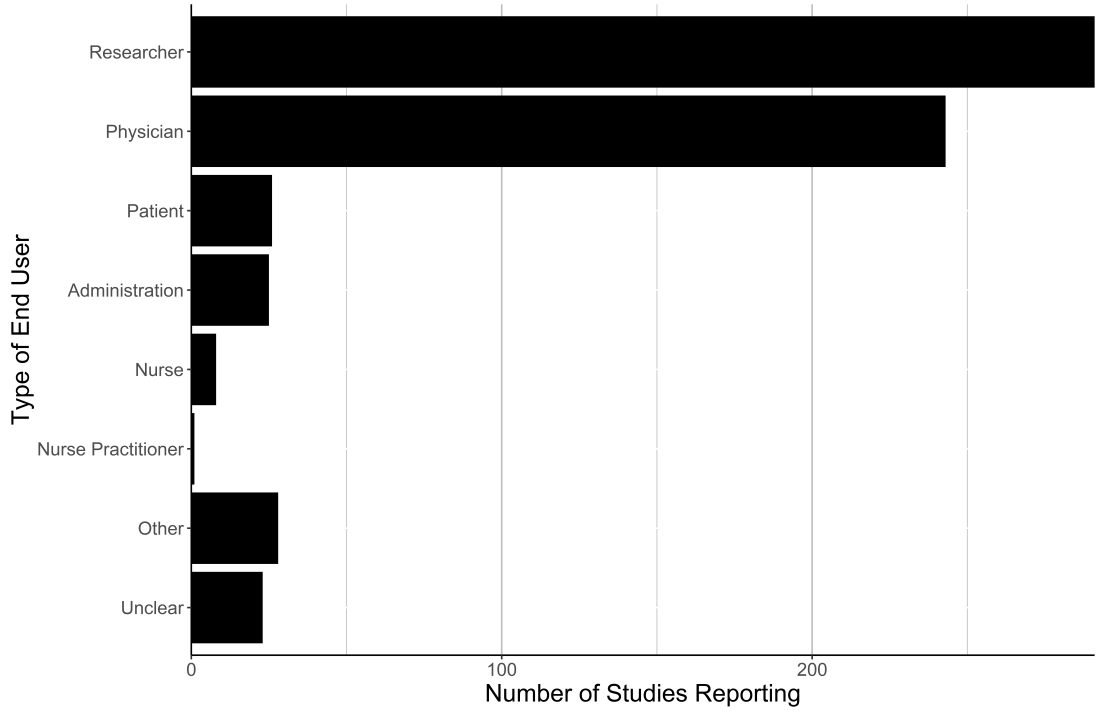


Figure C.3: Author reported end user total counts.
Note: A single study may contribute towards the count for one or more end user categories.

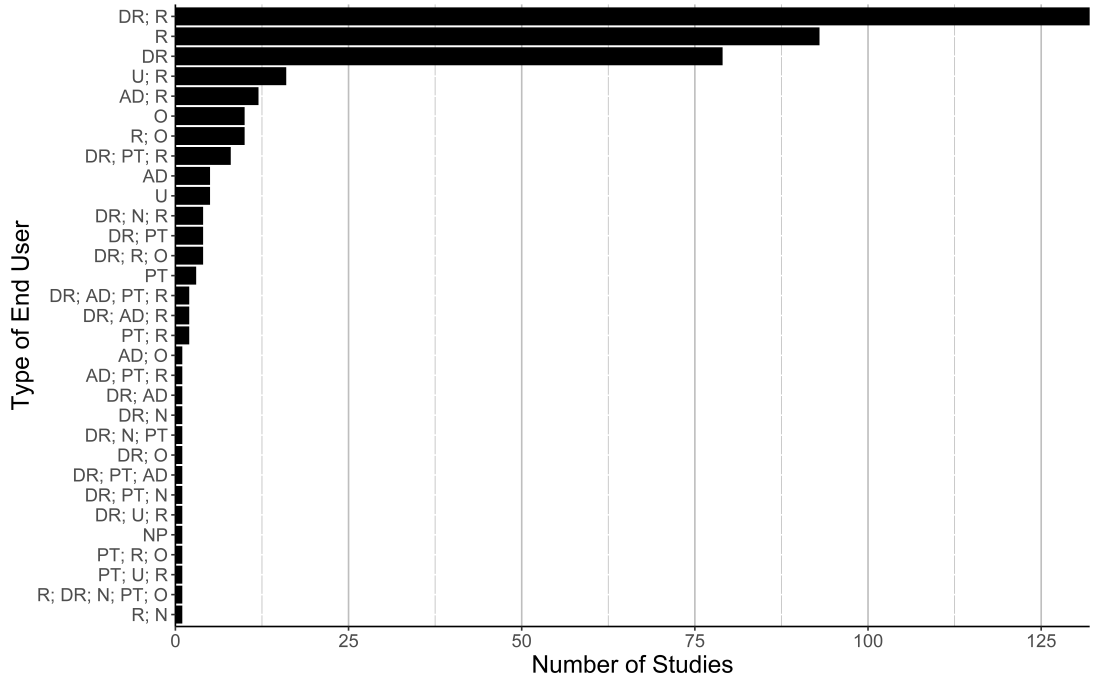


Figure C.4: Detailed breakdown of author reported intended end user combinations by study.
Legend: DR = Physician; R = Researcher; U = Unknown; AD = Administrator; PT = Patient; N = Nurse; NP = Nurse Practitioner; O = Other.

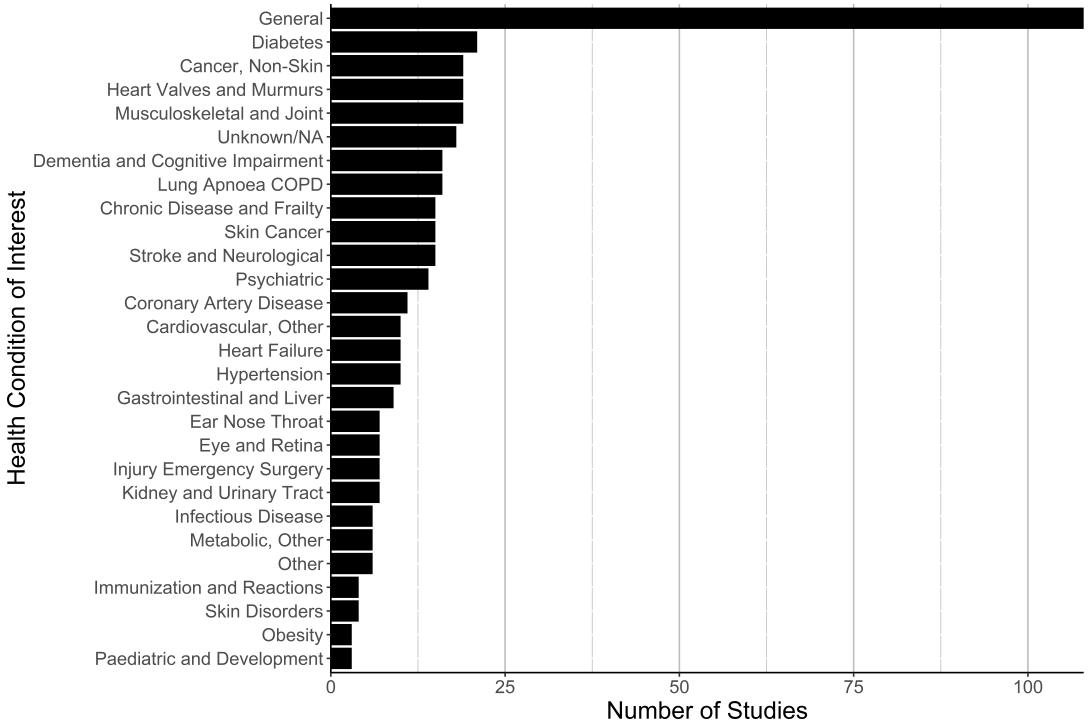


Figure C.5: Detailed breakdown of health conditions.

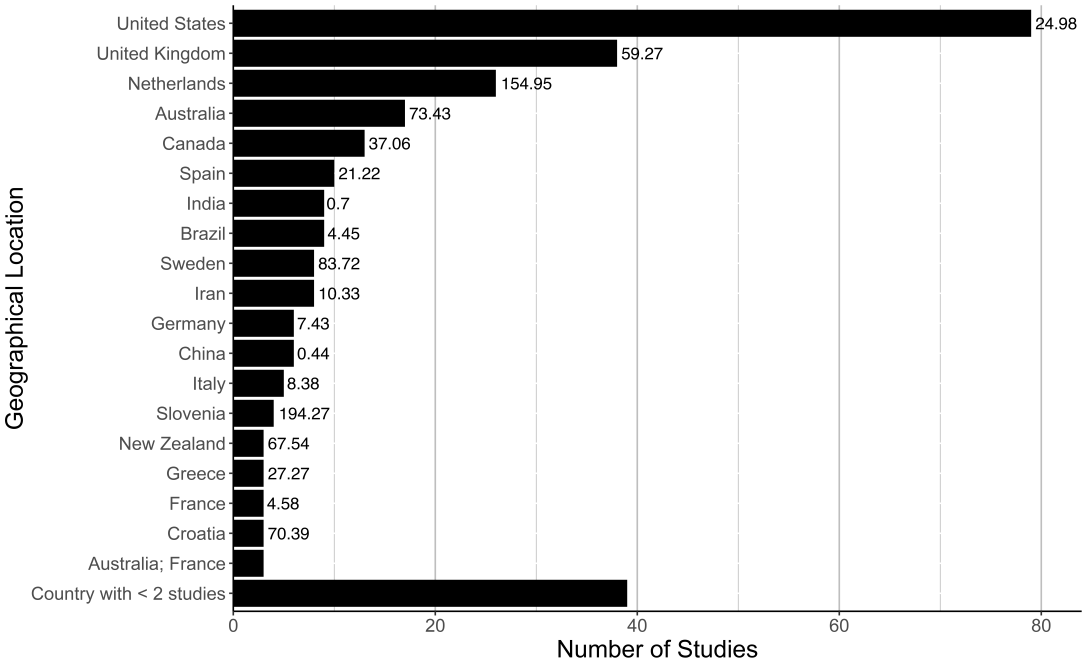


Figure C.6: Most frequent locations of data source or intended implementation with per capita rates.

Notes: Only studies with location reported are included (n=292). Number at the end of each bar is the number of studies per 100,000,000 people, based on 2013 population estimates.

Table C.5: Detailed breakdown of location.

Location	**Number of studies**
Unknown or Not Applicable	113
United States	79
United Kingdom	38
Netherlands	26
Australia	17
Canada	13
Spain	10
Brazil	9
India	9
Iran	8
Sweden	8
China	6
Germany	6
Italy	5
Slovenia	4
Australia and France	3
Croatia	3
France	3
Greece	3
New Zealand	3
Belgium	2
Egypt	2
Finland	2
Ireland	2
Japan	2
Norway	2
Singapore	2
Taiwan	2
Austria	1
Barcelona	1
Bulgaria	1
Canada and United States and United Kingdom and Brazil and Netherlands and Australia	1
Colombia	1
Czech Republic	1
Denmark	1
Europe	1
Germany and Norway	1
Greece and Bulgaria and Albania and Fyrom and Turkey	1
Hong Kong	1
Israel	1
Kuwait	1
Malaysia	1
Malta and Netherlands	1
Portugal	1
Saudi Arabia	1
South Africa	1
Switzerland	1
Turkey	1
United Kingdom and Greece and Germany	1

United States and Panama	1
United States and United Kingdom	1

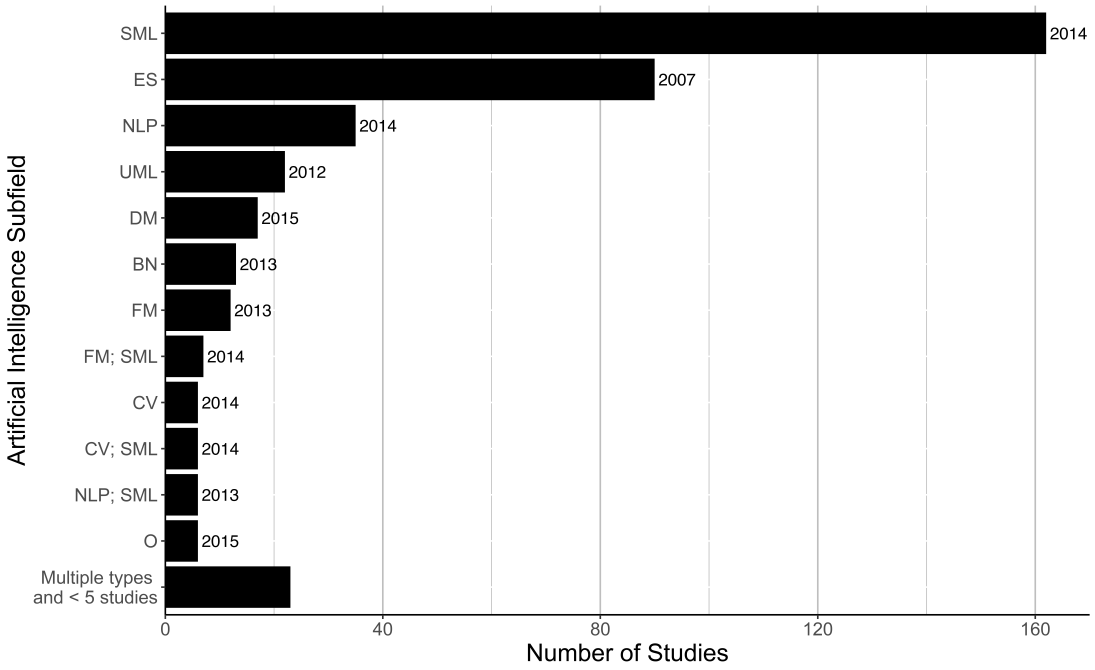


Figure C.7: Most frequent subfields of artificial intelligence with median year of publication.
Legend: SML = Supervised Machine Learning; ES = Expert System; NLP = Natural Language Processing; UML = Unsupervised Machine Learning; DM = Data Mining; BN = Bayesian Network; FM = Fuzzy Models; O = Other.

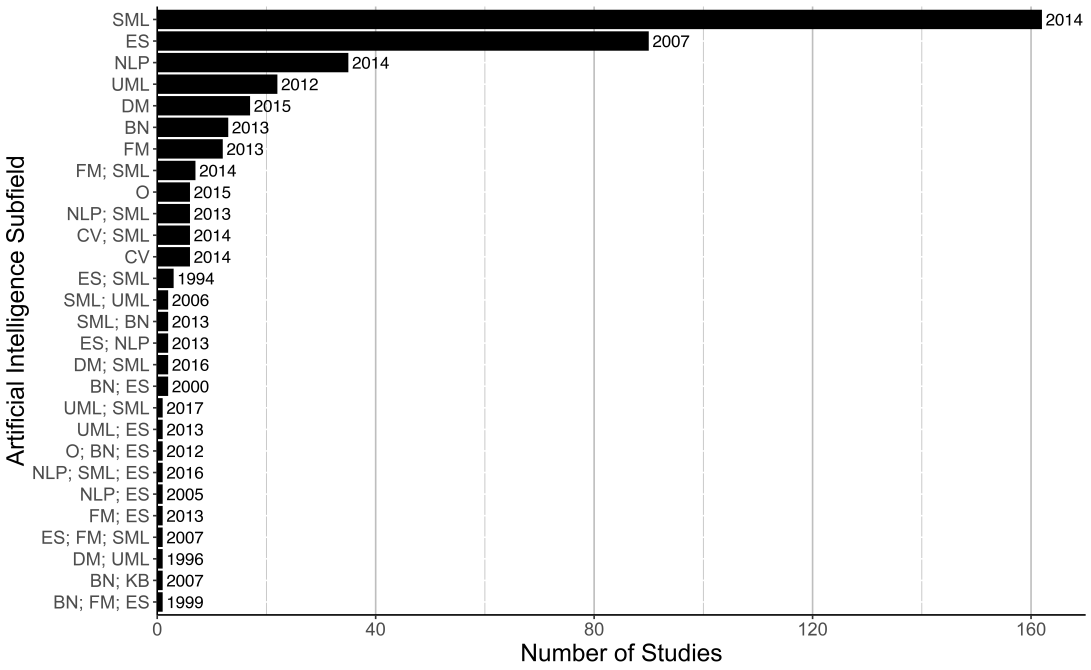


Figure C.8: Detailed breakdown of artificial intelligence subfields with median year of publication.
Legend: SML = Supervised Machine Learning; ES = Expert System; NLP = Natural Language Processing; UML = Unsupervised Machine Learning; DM = Data Mining; BN = Bayesian Network; FM = Fuzzy Models; O = Other.

C.4 References

1. Abdel-Aal RE, Mangoud AM. Modeling obesity using abductive networks. Computers and Biomedical Research. 1997;30(6):451-471. doi:10.1006/cbmr.1997.1460

2. Abe K, Takagi H, Minami M, Tian H. Features for measuring the congestive extent of internal hemorrhoids in endoscopic images. In: Proceedings of the Eleventh Australasian Data Mining Conference. Vol 146. Canberra, Australia: Conferences in Research and Practice in Information Technology Series; 2013:8.

3. Abidi S, Cox J, Shepherd M, Abidi SSR. Using OWL ontologies for clinical guidelines based comorbid decision support. In: Maui, HI, USA: IEEE Computer Society; 2012:3030-3038. doi:10.1109/HICSS.2012.629

4. Abidi SR, Abidi SS, Hussain S, Shepherd M. Ontology-based modeling of clinical practice guidelines: a clinical decision support system for breast cancer follow-up interventions at primary care settings. MEDINFO. 2007:845-849.

5. Abidi SR. Ontology-based modeling of breast cancer follow-up clinical practice guideline for providing clinical decision support. In: Twentieth IEEE International Symposium on Computer-Based Medical Systems (CBMS'07). Maribor, Slovenia: IEEE; 2007:542-547. doi:10.1109/CBMS.2007.80

6. Adjah ESO, Montvida O, Agbeve J, Paul SK. Data mining approach to identify disease cohorts from primary care electronic medical records: a case of diabetes mellitus. The Open Bioinformatics Journal. 2017;10(1):16-27. doi:10.2174/1875036201710010016

7. Aerts M, Minalu G, Bösner S, et al. Pooled individual patient data from five countries were used to derive a clinical prediction rule for coronary artery disease in primary care. Journal of Clinical Epidemiology. 2017;81:120-128. doi:10.1016/j.jclinepi.2016.09.011

8. af Klercker T, Zetraeus S. Dilemmas in introducing World Wide Web-based information technology in primary care: a focus group study. Family Practice. 1988;15(3):205-210.

9. af Klercker T. Effects of pruning of a decision-tree for the ear, nose, and throat realm in primary health care based on case-notes. Journal of Medical Systems. 1996;20(4):215-226. doi:10.1007/BF02263393

10. Afifi S, GholamHosseini H, Sinha R. A low-cost FPGA-based SVM classifier for melanoma detection. In: Malaysia: IEEE; 2016:631-636. doi:10.1109/IECBES.2016.7843526

11. Afifi S, GholamHosseini H, Sinha R. SVM classifier on chip for melanoma detection. In: Seogwipo: IEEE; 2017:270-274. doi:10.1109/EMBC.2017.8036814

12. Afzal Z, Kors JA, Sturkenboom MC, Schuemie MJ. Identifying drug-safety signals in electronic health records: An evaluation of automated case-detection algorithms with different sensitivity and specificity. Pharmacoepidemiology and Drug Safety. 2013;22(s1):285. doi:10.1002/pds.3512

13. Afzal Z, Schuemie MJ, Sen E, Jong G, Sturkenboom MC, Kors JA. Automatic generation of a case-detection algorithm for hepatobiliary disease using machine learning on free-text electronic health records. In: Vol 21. Pharmacoepidemiology and Drug Safety; 2012:194.

14. Afzal Z, Engelkes M, Verhamme KMC, et al. Automatic generation of case-detection algorithms to identify children with asthma from large electronic health record databases: automated case-detection algorithms. Pharmacoepidemiology and Drug Safety. 2013;22(8):826-833. doi:10.1002/pds.3438

15. Afzal Z, Pons E, Kang N, Sturkenboom MC, Schuemie MJ, Kors JA. ContextD: an algorithm to identify contextual properties of medical terms in a Dutch clinical corpus. BMC Bioinformatics. 2014;15(1). doi:10.1186/s12859-014-0373-3

16. Ahlstrom C, Hult P, Rask P, et al. Feature extraction for systolic heart murmur classification. Annals of Biomedical Engineering. 2006;34(11):1666-1677. doi:10.1007/s10439-006-9187-4

17. Al-Hamadani B. CardioOWL: An ontology-driven expert system for diagnosing coronary artery diseases. In: Subang, Selangor, Malaysia: IEEE; 2014:128-132. doi:10.1109/ICOS.2014.7042642

18. Alendahl K, Timpka T, Sjoberg C. Computerized knowledge bases in primary health care-a curse or a blessing for health promotion, prevention and patient quality. MEDINFO. 1995;95:917-921.

19. Alhashmi SM. Design of an internet-based advisory system: a multi-agent approach. In: Bui TD, Ho TV, Ha QT, eds. Intelligent Agents and Multi-Agent Systems. Vol 5357. Berlin, Heidelberg: Springer Berlin Heidelberg; 2008:14-25. doi:10.1007/978-3-540-89674-6_5

20. Ali Z, Alsulaiman M, Muhammad G, et al. Intra- and inter-database study for Arabic, English, and German databases: do conventional speech features detect voice pathology? Journal of Voice. 2017;31(3):386.e1-386.e8. doi:10.1016/j.jvoice.2016.09.009

21. Almansoori W, Zarour O, Jarada TN, Karampales P, Rokne J, Alhajj R. Applications of social network construction and analysis in the medical referral process. In: 2011 IEEE 9th International Conference on Dependable, Autonomic and Secure Computing. Sydney, Australia: IEEE; 2011:816-823. doi:10.1109/DASC.2011.140

22. Alty SR, Angarita-Jaimes N, Millasseau SC, Chowienczyk PJ. Predicting arterial stiffness from the digital volume pulse waveform. IEEE Transactions on Biomedical Engineering. 2007;54(12):2268-2275. doi:10.1109/TBME.2007.897805

23. Alvarez-Guisasola F, Conget I, Franch J, et al. Adding questions about cardiovascular risk factors improve the ability of the ADA questionnaire to identify unknown diabetic patients in Spain. Av Diabetol. 2010;26:347-352.

24. Alves R, Piñol M, Vilaplana J, et al. Computer-assisted initial diagnosis of rare diseases. PeerJ. 2016;4:e2211. doi:10.7717/peerj.2211

25. Anchala R, Di Angelantonio E, Prabhakaran D, Franco OH. Development and validation of a clinical and computerised decision support system for management of hypertension (DSS-HTN) at a primary health care (PHC) setting. Magrabi F, ed. PLoS ONE. 2013;8(11):e79638. doi:10.1371/journal.pone.0079638

26. Andrisevic N, Ejaz K, Rios-Gutierrez F, Alba-Flores R, Nordehn G, Burns S. Detection of heart murmurs using wavelet analysis and artificial neural networks. Journal of Biomechanical Engineering. 2005;127(6):899. doi:10.1115/1.2049327

27. Anwer A, Prasinos M, Bamiou D-E, et al. EMBalance data repository modeling and clinical application. In: IEEE 15th International Conference on Bioinformatics and Bioengineering. Belgrade, Serbia: IEEE; 2015. https://ieeexplore.ieee.org/document/7367677.

28. Aref-Eshghi E, Oake J, Godwin M, et al. Identification of dyslipidemic patients attending primary care clinics using electronic medical record (EMR) data from the Canadian Primary Care Sentinel Surveillance Network (CPCSSN) Database. Journal of Medical Systems. 2017;41(3). doi:10.1007/s10916-017-0694-7

29. Arnold CW, Oh A, Chen S, Speier W. Evaluating topic model interpretability from a primary care physician perspective. Computer Methods and Programs in Biomedicine. 2016;124:67-75. doi:10.1016/j.cmpb.2015.10.014

30. Astilean A, Avram C, Folea S, Silvasan I, Petreus D. Fuzzy Petri nets based decision support system for ambulatory treatment of non-severe acute diseases. In: 2010 IEEE International Conference on Automation, Quality and Testing, Robotics (AQTR). Cluj-Napoca, Romania: IEEE; 2010:1-6. doi:10.1109/AQTR.2010.5520706

31. Avila ML, Brandao LR, Wiliams S, et al. A simple pediatric post-thrombotic syndrome classification algorithm for the general practitioner. In: Vol 13. 2015:220.

32. Babic F, Vadovsky M, Muchova M, Paralic J, Majnaric L. Simple understandable analysis of medical data to support the diagnostic process. In: 2017 IEEE 15th International Symposium on Applied Machine Intelligence and Informatics (SAMI). Herl'any, Slovakia: IEEE; 2017:000153-000158. doi:10.1109/SAMI.2017.7880293

33. Baez GP, Viadero CF, Perez del Pino MA, Prochazka A, Suarez Araujo CP. HUMANN-based systems for differential diagnosis of dementia using neuropsychological tests. In: 2010 IEEE 14th International Conference on Intelligent Engineering Systems. Las Palmas: IEEE; 2010:67-72. doi:10.1109/INES.2010.5483872

34. Baez PG, Viadero CF, Espinosa NR, Perez del Pino MA, Suarez-Araujo CP. Detection of mild cognitive impairment using a counter propagation network based system. An e-health solution. In: 2015 International Workshop on Computational Intelligence for Multimedia Understanding (IWCIM). Prague, Czech Republic: IEEE; 2015:1-5. doi:10.1109/IWCIM.2015.7347094

35. Báez PG, Araujo CPS, Viadero CF, Procházka A. Differential diagnosis of dementia using HUMANN-S based ensembles. In: Fodor J, Klempous R, Suárez Araujo CP, eds. Recent Advances in Intelligent Engineering Systems. Vol 378. Berlin, Heidelberg: Springer Berlin Heidelberg; 2012:305-324. doi:10.1007/978-3-642-23229-9_14

36. Báez P, Viadero CF, García JR, Araujo CPS. An ensemble approach for the diagnosis of cognitive decline with missing data. In: Corchado E, Abraham A, Pedrycz W, eds. Hybrid Artificial Intelligence Systems. Vol 5271. Berlin, Heidelberg: Springer Berlin Heidelberg; 2008:353-360. doi:10.1007/978-3-540-87656-4_44

37. Balas EA, Li ZR, Spencer DC, Jaffrey F, Brent E, Mitchell JA. An expert system for performance-based direct delivery of published clinical evidence. Journal of the American Medical Informatics Association. 1996;3(1):56-65. doi:10.1136/jamia.1996.96342649

38. Barons MJ, Parsons N, Griffiths F, Thorogood M. A comparison of artificial neural network, latent class analysis and logistic regression for determining which patients benefit from a cognitive behavioural approach to treatment for non-specific low back pain. In: 2013 IEEE Symposium on Computational Intelligence in Healthcare and E-Health (CICARE). Singapore, Singapore: IEEE; 2013:7-12. doi:10.1109/CICARE.2013.6583061

39. Bekker J, Hommersom A, Lappenschaar M, Davis J. Measuring adverse drug effects on multimorbidity using tractable bayesian networks. 2016. http://arxiv.org/abs/1612.03055. Accessed July 3, 2019.

40. Beliakov G, Warren J. Fuzzy logic for decision support in chronic care. Artificial Intelligence in Medicine. 2001;21(1-3):209-213. doi:10.1016/S0933-3657(00)00087-7

41. Bermejo P, Lucas M, Rodríguez-Montes JA, et al. Single- and multi-label prediction of burden on families of schizophrenia patients. In: Peek N, Marín Morales R, Peleg M, eds. Artificial Intelligence in Medicine. Vol 7885. Berlin, Heidelberg: Springer Berlin Heidelberg; 2013:115-124. doi:10.1007/978-3-642-38326-7_18

42. Beveridge M, Fox J. Automatic generation of spoken dialogue from medical plans and ontologies. Journal of Biomedical Informatics. 2006;39(5):482-499. doi:10.1016/j.jbi.2005.12.008

43. Bhattacharya A, Mishra M, Singh A, Dutta MK. Machine learning based portable device for detection of cardiac abnormality. In: 2017 International Conference on Emerging Trends in Computing and Communication Technologies (ICETCCT). Dehradun: IEEE; 2017:1-4. doi:10.1109/ICETCCT.2017.8280307

44. Biermann E, Rihl J, Schenker M, Standl E. Semi-automatic generation of medical tele-expert opinion for primary care physician. Methods of Information in Medicine. 2003;42(03):212-219. doi:10.1055/s-0038-1634353

45. Biermans MCJ, Verheij RA, de Bakker DH, Zielhuis GA, de Vries Robbé PF. Estimating morbidity rates from electronic medical records in general practice: evaluation of a grouping system. Methods of Information in Medicine. 2008;47(02):98-106. doi:10.3414/ME0471

46. Biermans MCJ, de Bakker DH, Verheij RA, Gravestein JV, van der Linden MW, de Vries Robbé PF. Development of a case-based system for grouping diagnoses in general practice. International Journal of Medical Informatics. 2008;77(7):431-439. doi:10.1016/j.ijmedinf.2007.08.002

47. Bindels R, de Clercq PA, Winkens RAG, Hasman A. A test ordering system with automated reminders for primary care based on practice guidelines. International Journal of Medical Informatics. 2000;58-59:219-233. doi:10.1016/S1386-5056(00)00089-7

48. Bindels R, Hasman A, van Wersch JWJ, Talmon J, Winkens RAG. Evaluation of an automated test ordering and feedback system for general practitioners in daily practice. International Journal of Medical Informatics. 2004;73(9-10):705-712. doi:10.1016/j.ijmedinf.2004.06.001

49. Bland JDP, Weller P, Rudolfer S. Questionnaire tools for the diagnosis of carpal tunnel syndrome from the patient history. Muscle & Nerve. 2011;44(5):757-762. doi:10.1002/mus.22158

50. Boilot P, Hinesa E, John S, et al. Detection of bacteria causing eye infections using a neural network based electronic nose system. In: Section 4: Medical/Microbial. Electronic Noses and Olfaction; 2000:8.

51. Bourdès V, Ferrières J, Amar J, et al. Prediction of persistence of combined evidence-based cardiovascular medications in patients with acute coronary syndrome after hospital discharge using neural networks. Medical & Biological Engineering & Computing. 2011;49(8):947-955. doi:10.1007/s11517-011-0785-4

52. Brennan PF, Aronson AR. Towards linking patients and clinical information: detecting UMLS concepts in e-mail. Journal of Biomedical Informatics. 2003;36(4-5):334-341. doi:10.1016/j.jbi.2003.09.017

53. Bulegon H, Bortoleto S. Developing through a modeling of the database: an implementation of PEP for cardiovascular monitoring. In: World Congress on Nature and Biologically Inspired Computing. IEEE; 2009.

54. Bulegon H, Bortoleto S, Roman AC. Developing an implementation through a modeling of the database for cardiovascular monitoring. In: 2009 Ninth International Conference on Intelligent Systems Design and Applications. Pisa, Italy: IEEE; 2009:1156-1160. doi:10.1109/ISDA.2009.129

55. Buntinx F, Truyen J, Embrechts P, Moreel G, Peeters R. Evaluating patients with chest pain using classification and regression trees. Family Practice. 1992;9(2):149-153. doi:10.1093/fampra/9.2.149

56. Butler H, Baker M, Hegarty M, Finlayson D. A spectroscopic serum based blood test for brain tumours: optimisation for high-throughput sampling and the health economic impacts. In: Neuro-Oncology; 2018:i4.

57. Byrd RJ, Steinhubl SR, Sun J, Ebadollahi S, Stewart WF. Automatic identification of heart failure diagnostic criteria, using text analysis of clinical notes from electronic health records. International Journal of Medical Informatics. 2014;83(12):983-992. doi:10.1016/j.ijmedinf.2012.12.005

58. Cardillo E, Chiaravalloti MT, Pasceri E. Assessing ICD-9-CM and ICPC-2 use in primary care. An Italian case study. In: Proceedings of the 5th International Conference on Digital Health 2015 - DH '15. Florence, Italy: ACM Press; 2015:95-102. doi:10.1145/2750511.2750525

59. Cardillo E, Eccher C, Serafini L, Tamilin A. Logical analysis of mappings between medical classification systems. In: Dochev D, Pistore M, Traverso P, eds. Artificial Intelligence: Methodology, Systems, and Applications. Vol 5253. Berlin, Heidelberg: Springer Berlin Heidelberg; 2008:311-321. doi:10.1007/978-3-540-85776-1_26

60. Cardillo E, Serafini L, Tamilin A. A hybrid methodology for consumer-oriented healthcare knowledge acquisition. In: Springer-Verlag Berlin Heidelberg; 2010:38-49.

61. Carrara M, Bono A, Bartoli C, et al. Multispectral imaging and artificial neural network: mimicking the management decision of the clinician facing pigmented skin lesions. Physics in Medicine and Biology. 2007;52(9):2599-2613. doi:10.1088/0031-9155/52/9/018

62. Chamberlain D, Kodgule R, Ganelin D, Miglani V, Fletcher RR. Application of semi-supervised deep learning to lung sound analysis. In: 2016 38th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC). Orlando, FL, USA: IEEE; 2016:804-807. doi:10.1109/EMBC.2016.7590823

63. Chande PK, Katiyal S, Shrivastava M, Ramani AK. Expert based uroflow metering system. In: Proceedings IECON '91: 1991 International Conference on Industrial Electronics, Control and Instrumentation. Kobe, Japan: IEEE; 1991:1521-1524. doi:10.1109/IECON.1991.239110

64. Chang W-Y, Huang A, Chen Y-C, et al. The feasibility of using manual segmentation in a multifeature computer-aided diagnosis system for classification of skin lesions: a retrospective comparative study. BMJ Open. 2015;5(4):e007823-e007823. doi:10.1136/bmjopen-2015-007823

65. Chao J, Gu J, Zhang H, Chen H, Wu Z. The impact of the National Essential Medicines Policy on rational drug use in primary care institutions in Jiangsu Province of China. Iran J Public Health. 2018;47(1):24-32.

66. Chase HS, Mitrani LR, Lu GG, Fulgieri DJ. Early recognition of multiple sclerosis using natural language processing of the electronic health record. BMC Medical Informatics and Decision Making. 2017;17(1). doi:10.1186/s12911-017-0418-4

67. Chen J, Zheng X. A system architecture for smart health services and applications. In: Bikakis A, Zheng X, eds. Multi-Disciplinary Trends in Artificial Intelligence. Vol 9426. Cham: Springer International Publishing; 2015:449-456. doi:10.1007/978-3-319-26181-2_42

68. Chen J, Druhl E, Polepalli Ramesh B, et al. A natural language processing system that links medical terms in electronic health record notes to lay definitions: system development using physician reviews. Journal of Medical Internet Research. 2018;20(1):e26. doi:10.2196/jmir.8669

69. Chen S, Gutkowicz-Krusin D, Adrian J, Wells R. Performance of an adjuvant melanoma detection tool compared to physicians. In: Vol 64. J Am Acad Dermatol; 2011:AB121.

70. Choi E, Schuetz A, Stewart WF, Sun J. Using recurrent neural network models for early detection of heart failure onset. AMIA. 2017;24(2):361-370. doi:10.1093/jamia/ocw112

71. Chowdhury S, Linnarsson R, Wallgren A, Wallgren B, Wigertz O. Extracting knowledge from a large primary health care database using a knowledge-based statistical approach. Journal of Medical Systems. 1990;14(4):213-225. doi:10.1007/BF00999270

72. Çınar M, Engin M, Engin EZ, Ziya Ateşçi Y. Early prostate cancer diagnosis by using artificial neural networks and support vector machines. Expert Systems with Applications. 2009;36(3):6357-6361. doi:10.1016/j.eswa.2008.08.010

73. Closs VE, Ziegelmann PK, Flores JHF, Gomes I, Schwanke CHA. Anthropometric measures and frailty prediction in the elderly: an easy-to-use tool. Current Gerontology and Geriatrics Research. 2017;2017:1-8. doi:10.1155/2017/8703503

74. Cox A, Dinet J, Raluy-Callado M, MacLachlan S, Gabriel S. Resource utilisation and costs in patients with post-stroke spasticity in the United Kingdom. In: Vol 31. London: Clinical Rehabilitation; 2015:849. doi:10.1177/0269215517698029

75. Cox AP, Raluy-Callado M, Wang M, Bakheit AM, Moore AP, Dinet J. Predictive analysis for identifying potentially undiagnosed post-stroke spasticity patients in United Kingdom. Journal of Biomedical Informatics. 2016;60:328-333. doi:10.1016/j.jbi.2016.02.012

76. Cruz-Gutiérrez V, Posada-Zamora MA, Sánchez-López A. An Efficient Expert System for Diabetes with a Bayesian Inference Engine. In: Pichardo-Lagunas O, Miranda-Jiménez S, eds. Advances in Soft Computing. Vol 10062. Cham: Springer International Publishing; 2017:54-64. doi:10.1007/978-3-319-62428-0_5

77. Cubillas JJ, Ramos MI, Feito FR, Ureña T. An improvement in the appointment scheduling in primary health care centers using data mining. Journal of Medical Systems. 2014;38(8). doi:10.1007/s10916-014-0089-y

78. Das S, Sil J. Uncertainty management of health attributes for primary diagnosis. In: 2017 International Conference on Big Data Analytics and Computational Intelligence (ICBDAC). Chirala, Andhra Pradesh, India: IEEE; 2017:360-365. doi:10.1109/ICBDACI.2017.8070864

79. Dawson G, Dourish C, Goodwin G, et al. A precision medicine approach to antidepressant treatment in depression. In: Vol 40. Neuropsychopharmacology; 2015:S176. doi:10.1038/npp.2015.325

80. Dawson GR, Kingslake J., Dourish C.T., et al. A precision medicine approach to antidepressant treatment in depression. In: Assuring Replication: Problems and Solutions. Vol 79. Biological Psychiatry; 2016:S215. doi:10.1016/j.biopsych.2016.03.1054

81. de Groot H, Lous J, Peters L, Sandell A, Onell A. An allergy decision support system for non-allergy specialists: test and evaluation in routine use. In: New Findings in Allergy Diagnosis. Vol 64. Blackwell Publishing Ltd; 2009:533.

82. de Ipiña KL, Hernández MC, Graña M, Martínez E, Vaquero C. A computer-aided decision support system for shoulder pain pathology. In: Demazeau Y, Dignum F, Corchado JM, et al., eds. Trends in Practical Applications of Agents and Multiagent Systems. Vol 71. Berlin, Heidelberg: Springer Berlin Heidelberg; 2010:705-712. doi:10.1007/978-3-642-12433-4_83

83. de Ipiña KL, Hernández MC, Martínez E, Vaquero C. On the design of a CADS for shoulder pain pathology. In: Graña Romay M, Corchado E, Garcia Sebastian MT, eds. Hybrid Artificial Intelligence Systems. Vol 6076. Berlin, Heidelberg: Springer Berlin Heidelberg; 2010:508-515. doi:10.1007/978-3-642-13769-3_62

84. Denny JC, Peterson JF, Choma NN, et al. Development of a natural language processing system to identify timing and status of colonoscopy testing in electronic medical records. AMIA. 2009:141.

85. Denny JC, Choma NN, Peterson JF, et al. Natural language processing improves identification of colorectal cancer testing in the electronic medical record. Medical Decision Making. 2012;32(1):188-197. doi:10.1177/0272989X11400418

86. Dinet J, Cox AP, Raluy-Callado M, et al. Predictive analysis for identifying post stroke spasticity patients in UK primary care data. In: Vol 23. John Wiley & Sons, Ltd.; 2014:422. doi:10.1002/pds

87. Dixon BE, Simonaitis L, Goldberg HS, et al. A pilot study of distributed knowledge management and clinical decision support in the cloud. Artificial Intelligence in Medicine. 2013;59(1):45-53. doi:10.1016/j.artmed.2013.03.004

88. Dominguez-Morales JP, Jimenez-Fernandez AF, Dominguez-Morales MJ, Jimenez-Moreno G. Deep neural networks for the recognition and classification of heart murmurs using neuromorphic auditory sensors. IEEE Trans Biomed Circuits Syst. 2018;12(1):24-34. doi: 10.1109/TBCAS.2017.2751545

89. Dowell A, Darlow B, Macrae J, Stubbe M, Turner N, McBain L. Childhood respiratory illness presentation and service utilisation in primary care: a six-year cohort study in Wellington, New Zealand, using natural language processing (NLP) software. BMJ Open. 2017;7(7):e017146. doi:10.1136/bmjopen-2017-017146

90. Downs SM, Wallace MY. Mining association rules from a pediatric primary care decision support system. AMIA. 2000:200-204.

91. Dravenstott R, Kirchner HL, Strömblad C, Boris D, Leader J, Devapriya P. Applying predictive modeling to identify patients at risk to no-show. In: Proceedings of the 2014 Industrial and Systems Engineering Research Conference. ; :10.

92. Duarte PS, Mastrocolla LE, Farsky PS, et al. Selection of patients for myocardial perfusion scintigraphy based on fuzzy sets theory applied to clinical-epidemiological data and treadmill test results. Brazilian Journal of Medical and Biological Research. 2006;39(1):9-18. doi: 10.1590/S0100-879X2006000100002

93. Dubey AK. Using rough sets, neural networks, and logistic regression to predict compliance with cholesterol guidelines goals in patients with coronary artery disease. AMIA. 2003:834.

94. Dupuits FMHM. The use of the arden syntax for MLMs in HIOS +, a decision support system for general practitioners in the Netherlands. Comput Biol Med. 1994;24(5):405-410.

95. Dupuits FMHM. Evaluation of GPs' medical knowledge by a decision support system. Computer methods and programs in biomedicine. 1995;48:91-96.

96. Dupuits FMHM, Hasman A, Pop P. Computer-based assistance in family medicine. Computer Methods and Programs in Biomedicine. 1998;55(1):39-50. doi:10.1016/S0169-2607(97)00046-1

97. Ehrensperger MM, Taylor KI, Berres M, et al. BrainCheck - a very brief tool to detect incipient cognitive decline: optimized case-finding combining patient- and informant-based data. Alzheimer's Research & Therapy. 2014;6(9). doi:10.1186/s13195-014-0069-y

98. Eichler T, Thyrian JR, Fredrich D, et al. The benefits of implementing a computerized Intervention-Management-System (IMS) on delivering integrated dementia care in the primary care setting. International Psychogeriatrics. 2014;26(8):1377-1385. doi:10.1017/S1041610214000830

99. Emir B, Mardekian J, Masters ET, Clair A, Kuhn M, Silverman SL. Predictive modeling of a Fibromyalgia diagnosis: increasing the accuracy using real world data. In: Fibromyalgia, Soft Tissue Disorders, Regional and Specific Clinical Pain Syndromes: Research Focus. 2014:3.

100. Engelkes M, Afzal Z, Janssens HM, et al. Automated identification of asthma patients within an electronical medical record database using machine learning. In: Pharmacoepidemiology and Drug Safety. Vol 21. Wiley Online Library; 2012:12.

101. Ethier J-F, Curcin V, Barton A, et al. Clinical data integration model: core interoperability ontology for research using primary care data. Methods of Information in Medicine. 2015;54(01):16-23. doi:10.3414/ME13-02-0024

102. Exarchos TP, Rigas G, Bibas A, et al. Mining balance disorders' data for the development of diagnostic decision support systems. Computers in Biology and Medicine. 2016;77:240-248. doi:10.1016/j.compbiomed.2016.08.016

103. Fadhil A, Gabrielli S. Addressing challenges in promoting healthy lifestyles: the al-chatbot approach. In: Barcelona, Spain: Association for Computing Machinery; 2017.

104. Farmer N. An update and further testing of a knowledge-based diagnostic clinical decision support system for musculoskeletal disorders of the shoulder for use in a primary care setting: Musculoskeletal diagnostic CDSS. Journal of Evaluation in Clinical Practice. 2014;20(5):589-595. doi:10.1111/jep.12153

105. Farran B, Channanath AM, Behbehani K, Thanaraj TA. Predictive models to assess risk of type 2 diabetes, hypertension and comorbidity: machine-learning algorithms and validation using national health data from Kuwait—a cohort study. BMJ Open. 2013;3(5):e002457. doi:10.1136/bmjopen-2012-002457

106. Fathima SA, Hundewale N. Comparitive analysis of machine learning techniques for classification of arbovirus. In: Proceedings of 2012 IEEE-EMBS International Conference on Biomedical and Health Informatics. Hong Kong: IEEE; 2012:376-379. doi:10.1109/BHI.2012.6211593

107. Fernandez-Millan R, Medina-Merodio J-A, Plata R, Martinez-Herraiz J-J, Gutierrez-Martinez J-M. A laboratory test expert system for clinical diagnosis support in primary health care. Applied Sciences. 2015;5(3):222-240. doi:10.3390/app5030222

108. Ferreira-Santos D, Rodrigues PP. Improving diagnosis in obstructive sleep apnea with clinical data: a bayesian network approach. In: 2017 IEEE 30th International Symposium on Computer-Based Medical Systems (CBMS). Thessaloniki: IEEE; 2017:612-617. doi:10.1109/CBMS.2017.19

109. Flores CD, Marcelo FJ, Rosecler BM, Ana R, Helder C. Method for building a medical training simulator with bayesian networks: SimDeCS. In: 2nd KES International Conference on Innovation in Medicine and Healthcare. Vol 207. San Sebastian, Spain: Studies in Health Technology and Informatics; 2014:102–114. doi:10.3233/978-1-61499-474-9-102

110. Fok SC, Ng EYK, Thimm GL. Developing case-based reasoning for discovery of breast cancer. Journal of Mechanics in Medicine and Biology. 2003;3(3 & 4):231-245.

111. Font CR, Vas J, Rebello-Aguirre AC, et al. Subjective and objective (SVM) analysis of Brain perfusion SPECT in migraine patients treated with acupuncture. Preliminary results of a pragmatic randomised controlled trial. In: Neurosciences: Clinical Science. Vol 37. Eur J Nucl Med Mol Imagining; 2010:S391.

112. Forbes D, Wongthongtham P, Singh J. Development of patient-practitioner assistive communications (PPAC) ontology for type 2 diabetes management. In: Proceedings of The Workshop on New Trends of Computational Intelligence in Health Applications. Syndey, NSW: CEUR-WS; 2012:43-54.

113. Fourie D, Booysen MJ. Real time identification of heart sounds using selectional regional correlation of the time frequency domain. In: 2015 IEEE Symposium Series on Computational Intelligence. Cape Town, South Africa: IEEE; 2015:1326-1330. doi:10.1109/SSCI.2015.189

114. Freitas Junior MG, Nicolosi DEC. Computer system to aid in diagnosing acute coronary syndromes. Global Heart. 2014;9(1):e274. doi:10.1016/j.gheart.2014.03.2208

115. Frost DW, Vembu S, Wang J, Tu K, Morris Q, Abrams HB. Using the electronic medical record to identify patients at high risk for frequent emergency department visits and high system costs. The American Journal of Medicine. 2017;130(5):601.e17-601.e22. doi:10.1016/j.amjmed.2016.12.008

116. Gajjar K, Trevisan J, Owens G, et al. Fourier-transform infrared spectroscopy coupled with a classification machine for the analysis of blood plasma or serum: a novel diagnostic approach for ovarian cancer. The Analyst. 2013;138(14):3917. doi:10.1039/c3an36654e

117. Gamlyn L, Needham P, Sopher SM, Harris TJ. The development of a neural network-based ambulatory ECG monitor. Neural Computing & Applications. 1999;8(3):273-278. doi:10.1007/s005210050030

118. Gaonkar AP, Kulkarni R, Caytiles RD, Iyengar NChSN. Classification of lower back pain disorder using multiple machine learning techniques and identifying degree of importance of each parameter. International Journal of Advanced Science and Technology. 2017;105:11-24. doi:10.14257/ijast.2017.105.02

119. García-Altés A, Santín D, Barenys M. Applying artificial neural networks to the diagnosis of organic dyspepsia. Statistical Methods in Medical Research. 2007;16(4):331-346. doi:10.1177/0962280206071839

120. Gharehbaghi A, Dutoit T, Sepehri AA, Kocharian A, Lindén M. A novel method for screening children with isolated bicuspid aortic valve. Cardiovascular Engineering and Technology. 2015;6(4):546-556. doi:10.1007/s13239-015-0238-6

121. Gharehbaghi A, Linden M, Babic A. A decision support system for cardiac disease diagnosis based on machine learning methods. Studies in Health Technology and Informatics. 2017:43–47. doi:10.3233/978-1-61499-753-5-43

122. Gharehbaghi A, Sepehri AA, Kocharian A, Lindén M. An intelligent method for discrimination between aortic and pulmonary stenosis using phonocardiogram. In: Jaffray DA, ed. World Congress on Medical Physics and Biomedical Engineering, June 7-12, 2015, Toronto, Canada. Vol 51. Cham: Springer International Publishing; 2015:1010-1013. doi:10.1007/978-3-319-19387-8_246

123. Giabbanelli PJ, Torsney-Weir T, Mago VK. A fuzzy cognitive map of the psychosocial determinants of obesity. Applied Soft Computing. 2012;12(12):3711-3724. doi:10.1016/j.asoc.2012.02.006

124. Gibbons RD, Hooker G, Finkelman MD, et al. The computerized adaptive diagnostic test for major depressive disorder (CAD-MDD): a screening tool for depression. The Journal of Clinical Psychiatry. 2013;74(07):669-674. doi:10.4088/JCP.12m08338

125. Glasspool DW, Fox J, Coulson AS, Emery J. Risk assessment in genetics: a semi-quantitative approach. Stud Health Technol Inform. 2001;2001(84):459-463.

126. Goh WL, Lau KT. ESMP - expert system for medicine prescription. Expert Systems With Applications. 1991;3:457-462.

127. Goldstein MK, Hoffman BB, Coleman RW, et al. Implementing clinical practice guidelines while taking account of changing evidence: ATHENA DSS, an easily modifiable decision-support system for managing hypertension in primary care. AMIA. 2000:300-304.

128. Goldstein MK, Coleman RW, Tu SW, et al. Translating research into practice: organizational issues in implementing automated decision support for hypertension in three medical centers. Journal of the American Medical Informatics Association. 2004;11(5):368-376. doi:10.1197/jamia.M1534

129. Gonzalez C, Bernd B, Lopez DM. Adaptive intelligent systems for pHealth - an architectural approach. pHealth. 2012:170–175. doi:10.3233/978-1-61499-069-7-170

130. Greenberg JN, Toledano A, Adrian JS, Gutkowicz-Krusin D, Chen SC. Biopsy performance of a multi-spectral computer vision system for melanoma compared to physicians. In: Journal of Investigative Dermatology. Vol 132. Pigmentation & Melanoma; 2012:S131. doi:10.1038/jid.2012.90

131. Grigull L, Lechner W, Petri S, et al. Diagnostic support for selected neuromuscular diseases using answer-pattern recognition and data mining techniques: a proof of concept multicenter prospective trial. BMC Medical Informatics and Decision Making. 2016;16(1). doi:10.1186/s12911-016-0268-5

132. Gu Y, Kennelly J, Warren J, Nathani P, Boyce T. Automatic detection of skin and subcutaneous tissue infections from primary care electronic medical records. Studies in Health Technology and Informatics. 2015:74–80. doi:10.3233/978-1-61499-558-6-74

133. Guilbert TW, Arndt B, Temte J, et al. The theory and application of UW eHealth-PHINEX, a clinical electronic health record–public health information exchange. 2012;111(3):11.

134. Gums T, Carter B, Foster E. Medication adherence: algorithms for success (MED-ALS). In: Pharmacotherapy: The Journal of Human Pharmacology and Drug Therapy. Vol 37. Pharmacotherapy; 2017:e136.

135. Gunčar G, Kukar M, Notar M, et al. An application of machine learning to haematological diagnosis. Scientific Reports. 2018;8(1). doi:10.1038/s41598-017-18564-8

136. Hao S, Geng S, Fan L, Chen J, Zhang Q, Li L. Intelligent diagnosis of jaundice with dynamic uncertain causality graph model. Journal of Zhejiang University-SCIENCE B. 2017;18(5):393-401. doi:10.1631/jzus.B1600273

137. Harding SM. Prediction formulae for sleep-disordered breathing: Current Opinion in Pulmonary Medicine. 2001;7(6):381-385. doi:10.1097/00063198-200111000-00003

138. Hashemi J, Spina TV, Tepper M, et al. Computer vision tools for the non-invasive assessment of autism-related behavioral markers. 2012. http://arxiv.org/abs/1210.7014. Accessed July 3, 2019.

139. Hashemi J, Spina TV, Tepper M, et al. A computer vision approach for the assessment of autism-related behavioral markers. In: 2012 IEEE International Conference on Development and Learning and Epigenetic Robotics (ICDL). San Diego, CA, USA: IEEE; 2012:1-7. doi:10.1109/DevLrn.2012.6400865

140. Hashmi Z, Zrimec T, Hopkins A. A framework for automatic medical information processing of referral letters for query formulation. The Journal on Information Technology in Healthcare. 2009;7(6):327-344.

141. Hasman A, Pop P, Winkens RAG, Blom JL. To test or not to test, that is the question. Clinica Chimica Acta. 1993;222:49-56.

142. Hazlehurst BL, Lawrence JM, Donahoo WT, et al. Automating assessment of lifestyle counseling in electronic health records. American Journal of Preventive Medicine. 2014;46(5):457-464. doi:10.1016/j.amepre.2014.01.001

143. Hazlehurst B, Mullooly J, Naleway A, Crane B. Detecting possible vaccination reactions in clinical notes. AMIA. 2005:306-310.

144. He H, Wang J, Graco W, Hawkins S. Application of neural networks to detection of medical fraud. Expert Systems with Applications. 1997;13(4):329-336. doi:10.1016/S0957-4174(97)00045-6

145. Heckerling P, Canaris G, Flach S, Tape T, Wigton R, Gerber B. Predictors of urinary tract infection based on artificial neural networks and genetic algorithms. International Journal of Medical Informatics. 2007;76(4):289-296. doi:10.1016/j.ijmedinf.2006.01.005

146. Hejlesen OK, Olesen KG, Dessau R, Beltoft I, Trangeled M. Decision support for diagnosis of lyme disease. Connecting Medical Informatics and Bio-Informatics. 2005:205-210.

147. Hertroijs DFL, Elissen AMJ, Brouwers MCGJ, et al. A risk score including body mass index, glycated haemoglobin and triglycerides predicts future glycaemic control in people with type 2 diabetes. Diabetes, Obesity and Metabolism. 2018;20(3):681-688. doi:10.1111/dom.13148

148. Homa L, Rose J, Hovmand PS, et al. A participatory model of the paradox of primary care. The Annals of Family Medicine. 2015;13(5):456-465. doi:10.1370/afm.1841

149. Hommersom A, Verwer S, Lucas PJF. Discovering probabilistic structures of healthcare processes. In: Riaño D, Lenz R, Miksch S, Peleg M, Reichert M, ten Teije A, eds. Process Support and Knowledge Representation in Health Care. Vol 8268. Cham: Springer International Publishing; 2013:53-67. doi:10.1007/978-3-319-03916-9_5

150. Horowitz N, Moshkowitz M, Halpern Z, Leshno M. Applying data mining techniques in the development of a diagnostics questionnaire for GERD. Digestive Diseases and Sciences. 2007;52(8):1871-1878. doi:10.1007/s10620-006-9202-5

151. Hosomura N, Malmasi S, Timerman D, et al. Decline of insulin therapy and delays in insulin initiation in people with uncontrolled diabetes mellitus. Diabetic Medicine. 2017;34(11):1599-1602. doi:10.1111/dme.13454

152. Hosomura N, Malmasi S, Timerman D, et al. Studying insulin decline by patients with diabetes. In: What's New in Diabetes Epidemiology? American Diabetes Association; 2017:1.

153. Hrabak KM, Campbell JR, Tu SW, McClure R, Weida RT. Creating interoperable guidelines: requirements of vocabulary standards in immunization decision support. MEDINFO. 2007:930-934.

154. Hripcsak G, Soulakis ND, Li L, et al. Syndromic surveillance using ambulatory electronic health records. Journal of the American Medical Informatics Association. 2009;16(3):354-361. doi:10.1197/jamia.M2922

155. Hung J, Posey J, Freedman R, Thorton T. Electronic surveillance of disease states: a preliminary study in electronic detection of respiratory diseases in a primary care setting. AMIA. 1998:688-692.

156. Hunt DL, Haynes RB, Morgan D. Using old technology to implement modern computer-aided decision support for primary diabetes care. AMIA. 2001:274-278.

157. Huseth O, Rost TB. Developing an annotated corpus of patient histories from the primary care health record. In: 2007 IEEE International Conference on Bioinformatics and Biomedicine Workshops. San Jose, CA, USA: IEEE; 2007:165-173. doi:10.1109/BIBMW.2007.4425415

158. Hylan TR, Von Korff M, Saunders K, et al. Automated prediction of risk for problem opioid use in a primary care setting. The Journal of Pain. 2015;16(4):380-387. doi:10.1016/j.jpain.2015.01.011

159. Inbarani HH, Kumar SS, Azar AT, Hassanien AE. Soft rough sets for heart valve disease diagnosis. In: Hassanien AE, Tolba MF, Taher Azar A, eds. Advanced Machine Learning Technologies and Applications. Vol 488. Cham: Springer International Publishing; 2014:347-356. doi:10.1007/978-3-319-13461-1_33

160. Jagannathan V, Mullett CJ, Arbogast JG, et al. Assessment of commercial NLP engines for medication information extraction from dictated clinical notes. International Journal of Medical Informatics. 2009;78(4):284-291. doi:10.1016/j.ijmedinf.2008.08.006

161. Jammeh E, Carroll C, Pearson S, et al. Using NHS primary care data to identify undiagnosed dementia. In: ABN. Vol 86. Journal of Neurology, Neurosurgery & Psychiatry; 2015:e4. doi:10.1136/jnnp-2015-312379.44

162. Janssen KJM, Siccama I, Vergouwe Y, et al. Development and validation of clinical prediction models: marginal differences between logistic regression, penalized maximum likelihood estimation, and genetic programming. Journal of Clinical Epidemiology. 2012;65(4):404-412. doi:10.1016/j.jclinepi.2011.08.011

163. Jarvik JG, Gold LS, Tan K, et al. Long-term outcomes of a large, prospective observational cohort of older adults with back pain. The Spine Journal. 2018;18(9):1540-1551. doi:10.1016/j.spinee.2018.01.018

164. Jatobá A, Bellas HC, Koster I, et al. Supporting decision-making in patient risk assessment using a hierarchical fuzzy model. Cognition, Technology & Work. 2018;20(3):477-488. doi:10.1007/s10111-018-0475-1

165. Javed F, Venkatachalam PA, Hani AFM. Knowledge based system with embedded intelligent heart sound analyser for diagnosing cardiovascular disorders. Journal of Medical Engineering & Technology. 2007;31(5):341-350. doi:10.1080/03091900600887876

166. Johnson P, Tu S, Jones N. Achieving reuse of computable guideline systems. MEDINFO. 2001:99-103.

167. Johnson PD, Tu S, Booth N, Sugden B, Purves IN. Using scenarios in chronic disease management guidelines for primary care. AMIA. 2000:389-393.

168. Jordan P, Shedden-Mora MC, Löwe B. Predicting suicidal ideation in primary care: An approach to identify easily assessable key variables. General Hospital Psychiatry. 2018;51:106-111. doi:10.1016/j.genhosppsych.2018.02.002

169. Joudaki H, Rashidian A, Minaei-Bidgoli B, et al. Improving fraud and abuse detection in general physician claims: a data mining study. International Journal of Health Policy and Management. 2015;5(3):165-172. doi:10.15171/ijhpm.2015.196

170. Kabadi S, Lee A, Kuhn M, Gray D. Predictive model of Parkinson's Disease in large electronic health records database. Value in Health. 2016;19(3):A59. doi:10.1016/j.jval.2016.03.179

171. Kalankesh L, Weatherall J, Ba-Dhfari T, Buchan I, Brass A. Taming EHR data: using semantic similarity to reduce dimensionality. Studies in Health Technology and Informatics. 2013:52–56. doi:10.3233/978-1-61499-289-9-52

172. Kalatzis FG, Exarchos TP, Giannakeas N, et al. Point-of-care monitoring and diagnostics for rheumatoid arthritis and multiple sclerosis. In: Advanced Topics in Scattering Theory and Biomedical Engineering. Patras, Greece: WORLD SCIENTIFIC; 2010:98-105. doi:10.1142/9789814322034_0010

173. Kalina J, Seidl L, Zvara K, Grunfeldova H, Slovak D, Zvarova J. System for selecting relevant information for decision support. Data and Knowledge for Medical Decision Support. 2013:83-87. doi:10.24105/ejbi.2013.09.1.1

174. Katigari MR, Ayatollahi H, Malek M, Haghighi M. Fuzzy expert system for diagnosing diabetic neuropathy. World Journal of Diabetes. 2017;8(2):80. doi:10.4239/wjd.v8.i2.80

175. Kay E, Agarwal A. DropConnected neural networks trained on time-frequency and inter-beat features for classifying heart sounds. Physiological Measurement. 2017;38(8):1645-1657. doi:10.1088/1361-6579/aa6a3d

176. Kent RD, Preney PD, Snowdon AW, et al. Design and implementation of a primary health care services navigational system architecture. In: Watada J, Phillips-Wren G, Jain LC, Howlett RJ, eds. Intelligent Decision Technologies. Vol 10. Berlin, Heidelberg: Springer Berlin Heidelberg; 2011:743-752. doi:10.1007/978-3-642-22194-1_73

177. Kerr GS, Richards JS, Nunziato CA, et al. Measuring physician adherence with gout quality indicators: a role for natural language processing: natural language processing and gout quality indicators. Arthritis Care & Research. 2015;67(2):273-279. doi:10.1002/acr.22406

178. Khalili SP, LaNoue M. Population health at home: building a data ware-house and applying cluster analysis to identify patterns in healthcare utilization and multimorbidity among primary care patients. In: Abstracts from the 2017 Society of General Internal Medicine Annual Meeting. JGIM; 2017:S786.

179. Klann JG, Anand V, Downs SM. Patient-tailored prioritization for a pediatric care decision support system through machine learning. J Am Med Inform Assoc. 2013;20:e267-e274. doi:10.1136/amiajnl-2013-001865

180. Klausner S, Entacher K, Kranzer S, Flamm M, Sonnichsen A, Fritsch G. ProPath - a guideline based software for the implementation into the medical environment. In: IEEE Canada International Humanitarian Technology Conference. IEEE; 2014:6.

181. Kocharian A, Sepehri A-A, Janani A, Malakan-Rad E. Efficiency, sensitivity and specificity of automated auscultation diagnosis device for detection and discrimination of cardiac murmurs in children. Iran J Pediatr. 2013;23(4):6.

182. Koeling R, Tate AR, Carroll JA. Automatically estimating the incidence of symptoms recorded in GP free text notes. In: Proceedings of the First International Workshop on Managing Interoperability and Complexity in Health Systems - MIXHS '11. Glasgow, Scotland, UK: ACM Press; 2011:43. doi:10.1145/2064747.2064757

183. Kok MR, Boon ME, Schreiner-Kok PG, Hermans J, Grobbee DE, Kok LP. Less medical intervention after sharp demarcation of grade 1-2 cervical intraepithelial neoplasia smears by neural network screening. Cancer Cytopathology. 2001;93(3):6.

184. Kop R, Hoogendoorn M, Teije A ten, et al. Predictive modeling of colorectal cancer using a dedicated pre-processing pipeline on routine electronic medical records. Computers in Biology and Medicine. 2016;76:30-38. doi:10.1016/j.compbiomed.2016.06.019

185. Koskela T-H, Ryynanen O-P, Soini EJ. Risk factors for persistent frequent use of the primary health care services among frequent attenders: A Bayesian approach. Scandinavian Journal of Primary Health Care. 2010;28(1):55-61. doi:10.3109/02813431003690596

186. Kruse C, Eiken P, Vestergaard P. Clinical fracture risk evaluated by hierarchical agglomerative clustering. Osteoporosis International. 2017;28(3):819-832. doi:10.1007/s00198-016-3828-8

187. Kunhimangalam R, Ovallath S, Joseph PK. Computer aided diagnostic problem solving: Identification of peripheral nerve disorders. IRBM. 2013;34(3):244-251. doi:10.1016/j.irbm.2013.04.003

188. Kuntagod N, Mukherjee C. Mobile decision support system for outreach health worker. In: 2011 IEEE 13th International Conference on E-Health Networking, Applications and Services. Columbia, MO, USA: IEEE; 2011:56-59. doi:10.1109/HEALTH.2011.6026786

189. Kwiatkowska M, Atkins MS, Ayas NT, Ryan CF, Street H. Telemedicine system for early assessment of obstructive sleep apnea. In: Banff; 2005:147-155.

190. Lacagnina V, Leto-Barone MS, La Piana S, La Porta G, Pingitore G, Di Lorenzo G. Comparison between statistical and fuzzy approaches for improving diagnostic decision making in patients with chronic nasal symptoms. Fuzzy Sets and Systems. 2014;237:136-150. doi:10.1016/j.fss.2013.10.013

191. Lacson R, Harris K, Brawarsky P, et al. Evaluation of an automated information extraction tool for imaging data elements to populate a breast cancer screening registry. Journal of Digital Imaging. 2015;28(5):567-575. doi:10.1007/s10278-014-9762-4

192. LaFreniere D, Zulkernine F, Barber D, Martin K. Using machine learning to predict hypertension from a clinical dataset. In: 2016 IEEE Symposium Series on Computational Intelligence (SSCI). Athens, Greece: IEEE; 2016:1-7. doi:10.1109/SSCI.2016.7849886

193. Lähteenmäki J, Leppänen J, Orsama A-L, et al. Remote patient monitoring system with decision support. BioMed. 2010:5.

194. Lange LL, Haak SW, Lincoln MJ, et al. Use of Iliad to improve diagnostic performance of nurse practitioner students. Journal of Nursing Education. 1997;36(1):36-45.

195. Lappenschaar M, Hommersom A, Lucas PJF, Lagro J, Visscher S. Multilevel Bayesian networks for the analysis of hierarchical health care data. Artificial Intelligence in Medicine. 2013;57(3):171-183. doi:10.1016/j.artmed.2012.12.007

196. Lappenschaar M, Hommersom A, Lucas PJF, et al. Multilevel temporal Bayesian networks can model longitudinal change in multimorbidity. Journal of Clinical Epidemiology. 2013;66(12):1405-1416. doi:10.1016/j.jclinepi.2013.06.018

197. Lappenschaar M, Hommersom A, Visscher S, Lucas PJF. Modeling inter-practice variation of disease interactions using multilevel bayesian networks. In: AIME'11: The 13th Conference on Artificial Intelligence in Medicine, ProBiomed'11. Bled Slovenia; 2011:15.

198. Larkin A, Healy C, Warters M. Simulation-based medical education: strategies to improve T2D management? In: Endocrine Healthcare Delivery & Education II. San Diego: Endocrine Society; 2015. doi:10.1210/endo-meetings.2015.AHPAA.11

199. Lau K, Wilkinson J, Moorthy R. A web-based prediction score for head and neck cancer referrals. Clinical Otolaryngology. 2018;43(4):1043-1049. doi:10.1111/coa.13098

200. Lavesson N. Predicting the risk of future hospitalization. In: 2010 Workshops on Database and Expert Systems Applications. Bilbao, TBD, Spain: IEEE; 2010:120-124. doi:10.1109/DEXA.2010.43

201. Lavrač N, Bohanec M, Pur A, et al. Resource modeling and analysis of regional public health care data by means of knowledge technologies. In: Miksch S, Hunter J, Keravnou ET, eds. Artificial Intelligence in Medicine. Vol 3581. Berlin, Heidelberg: Springer Berlin Heidelberg; 2005:414-418. doi:10.1007/11527770_55

202. Leidy NK, Malley KG, Steenrod A, et al. Can we find undiagnosed, high-risk patients with COPD in primary care? Using random forests to identify best variable sets for COPD case identification. In: COPD: Screening and Diagnostic Tools. Vol 189. San Diego, CA: Am J Respir Crit Care Med; 2014:A2961.

203. Leidy NK, Malley KG, Steenrod AW, et al. Insight into best variables for COPD case identification: a random forests analysis. Chronic Obstructive Pulmonary Diseases: Journal of the COPD Foundation. 2016;3(1):406-418. doi:10.15326/jcopdf.3.1.2015.0144

204. Li C-M, Du Y-C, Wu J-X, et al. Synchronizing chaotification with support vector machine and wolf pack search algorithm for estimation of peripheral vascular occlusion in diabetes mellitus. Biomedical Signal Processing and Control. 2014;9:45-55. doi:10.1016/j.bspc.2013.10.001

205. Lidbury BA, Richardson AM, Badrick T. Assessment of machine-learning techniques on large pathology data sets to address assay redundancy in routine liver function test profiles. Diagnosis. 2015;2(1):41-51. doi:10.1515/dx-2014-0063

206. Lijun Pan, Xiaoting Fu, Fangfang Cai, Yu Meng, Changjiang Zhang. A compact electronic medical record system for regional clinics and health centers in China: Design and its application. In: 2016 IEEE International Conference on Bioinformatics and Biomedicine (BIBM). Shenzhen, China: IEEE; 2016:1010-1015. doi:10.1109/BIBM.2016.7822660

207. Lin C, Lin C-M, Li S-T, Kuo S-C. Intelligent physician segmentation and management based on KDD approach. Expert Systems with Applications. 2008;34(3):1963-1973. doi:10.1016/j.eswa.2007.02.038

208. Lin H-S, Talwar HS, Tarca AL, et al. Autoantibody approach for serum-based detection of head and neck cancer. Cancer Epidemiology Biomarkers & Prevention. 2007;16(11):2396-2405. doi:10.1158/1055-9965.EPI-07-0318

209. Linden A, Yarnold PR. Using data mining techniques to characterize participation in observational studies: Data mining in observational studies. Journal of Evaluation in Clinical Practice. 2016;22(6):839-847. doi:10.1111/jep.12515

210. Linnarsson R. Decision support for drug prescription integrated with computer-based patient records in primary care. Medical Informatics. 1993;18(2):131-142. doi:10.3109/14639239309034475

211. Lismont J, Janssens A-S, Odnoletkova I, vanden Broucke S, Caron F, Vanthienen J. A guide for the application of analytics on healthcare processes: A dynamic view on patient pathways. Computers in Biology and Medicine. 2016;77:125-134. doi:10.1016/j.compbiomed.2016.08.007

212. Liu F, Antieau LD, Yu H. Toward automated consumer question answering: Automatically separating consumer questions from professional questions in the healthcare domain. Journal of Biomedical Informatics. 2011;44(6):1032-1038. doi:10.1016/j.jbi.2011.08.008

213. Long J, Yuan MJ, Poonawala R. An observational study to evaluate the usability and intent to adopt an artificial intelligence–powered medication reconciliation tool. interactive Journal of Medical Research. 2016;5(2):e14. doi:10.2196/ijmr.5462

214. Luther SL, McCart JA, Berndt DJ, et al. Improving identification of fall-related injuries in ambulatory care using statistical text mining. American Journal of Public Health. 2015;105(6):1168-1173. doi:10.2105/AJPH.2014.302440

215. MacRae J, Darlow B, McBain L, et al. Accessing primary care big data: the development of a software algorithm to explore the rich content of consultation records. BMJ Open. 2015;5(8):e008160. doi:10.1136/bmjopen-2015-008160

216. Maglogiannis I, Loukis E, Zafiropoulos E, Stasis A. Support Vectors Machine-based identification of heart valve diseases using heart sounds. Computer Methods and Programs in Biomedicine. 2009;95(1):47-61. doi:10.1016/j.cmpb.2009.01.003

217. Mahesh V, Ramakrishnan S. Assessment and classification of normal and restrictive respiratory conditions through pulmonary function test and neural network. Journal of Medical Engineering & Technology. 2007;31(4):300-304. doi:10.1080/03091900701233962

218. Majnarić-Trtica L, Vitale B. Systems biology as a conceptual framework for research in family medicine; use in predicting response to influenza vaccination. Primary Health Care Research & Development. 2011;12(04):310-321. doi:10.1017/S1463423611000089

219. Maramba ID, Davey A, Elliott MN, et al. Web-based textual analysis of free-text patient experience comments from a survey in primary care. JMIR Medical Informatics. 2015;3(2):e20. doi:10.2196/medinform.3783

220. Markaki M, Germanakis I, Stylianou Y. Automatic classification of systolic heart murmurs. In: 2013 IEEE International Conference on Acoustics, Speech and Signal Processing. Vancouver, BC, Canada: IEEE; 2013:1301-1305. doi:10.1109/ICASSP.2013.6637861

221. Marling C, Wiley M, Bunescu R, Shubrook J, Schwartz F. Emerging applications for intelligent diabetes management. AI Magazine. 2012;33(2):67-78.

222. Martin CM, Vogel C, Grady D, et al. Implementation of complex adaptive chronic care: the Patient Journey Record system (PaJR): Complex adaptive chronic care - PaJR. Journal of Evaluation in Clinical Practice. 2012;18(6):1226-1234. doi:10.1111/j.1365-2753.2012.01880.x

223. Martin L. Knowledge acquisition and evaluation of an expert system for managing disorders of the outer eye. Computers in Nursing. 2001;19(3):114-117.

224. Martinez FJ, Mannino DM, Leidy NK, et al. A new approach for identifying patients with undiagnosed, clinically significant COPD in primary care. In: Not a Second Time: Solutions to COPD Readmissions and Care Management. Vol 191. American Journal of Respiratory and Critical Care Medicine; 2015:A6311.

225. Martinez F, Mannino D, Leidy N, et al. Identifying patients with undiagnosed clinically-significant COPD in primary care: what questions should we be asking? In: Obstructive Lung Diseases. Vol 148. Chest; 2015:696A. doi:10.1378/chest.2273164

226. Martinez F, Mannino DM, Leidy NK, et al. A novel approach for identifying patients with undiagnosed clinically-significant COPD. In: 1.6 General Practice and Primary Care. Vol 46. European Respiratory Society; 2015:OA3284. doi:10.1183/13993003.congress-2015.OA3284

227. Maslekar S, Gardiner AB, Monson JRT, Duthie GS. Artificial neural networks to predict presence of significant pathology in patients presenting to routine colorectal clinics: neural networks to predict colorectal pathology. Colorectal Disease. 2010;12(12):1254-1259. doi:10.1111/j.1463-1318.2009.02005.x

228. Masood A, Al-Jumaily A. SA-SVM based automated diagnostic system for skin cancer. In: Wang Y, Jiang X, Zhang D, eds. Sixth International Conference on Graphic and Image Processing (ICGIP 2014). Beijing, China; 2015:94432L. doi:10.1117/12.2179094

229. Masood A, Al-Jumaily AA, Adnan T. Development of automated diagnostic system for skin cancer: performance analysis of neural network learning algorithms for classification. In: Wermter S, Weber C, Duch W, et al., eds. Artificial Neural Networks and Machine Learning – ICANN 2014. Vol 8681. Cham: Springer International Publishing; 2014:837-844. doi:10.1007/978-3-319-11179-7_105

230. Masood A, Al-Jumaily A, Anam K. Texture analysis based automated decision support system for classification of skin cancer using SA-SVM. In: Loo CK, Yap KS, Wong KW, Teoh A, Huang K, eds. Neural Information Processing. Vol 8835. Cham: Springer International Publishing; 2014:101-109. doi:10.1007/978-3-319-12640-1_13

231. Masood A, Al-Jumaily A, Aung YM. Scaled conjugate gradient based decision support system for automated diagnosis of skin cancer. In: Biomedical Engineering / 817: Robotics Applications. Zurich, Switzerland: ACTAPRESS; 2014. doi:10.2316/P.2014.818-020

232. Matheny ME, FitzHenry F, Speroff T, et al. Detection of infectious symptoms from VA emergency department and primary care clinical documentation. International Journal of Medical Informatics. 2012;81(3):143-156. doi:10.1016/j.ijmedinf.2011.11.005

233. Mathias JS, Agrawal A, Feinglass J, Cooper AJ, Baker DW, Choudhary A. Development of a 5 year life expectancy index in older adults using predictive mining of electronic health record data. Journal of the American Medical Informatics Association. 2013;20(e1):e118-e124. doi:10.1136/amiajnl-2012-001360

234. Maurer A, Hanrahan S, Nedrud J, Hebb AO, Papandreou-Suppappola A. Suppression of neurostimulation artifacts and adaptive clustering of Parkinson's patients behavioral tasks using EEG. In: 2016 50th Asilomar Conference on Signals, Systems and Computers. Pacific Grove, CA, USA: IEEE; 2016:851-855. doi:10.1109/ACSSC.2016.7869169

235. McBride J, Zhao X, Munro N, Jicha G, Smith C, Jiang Y. Discrimination of mild cognitive impairment and alzheimer's disease using transfer entropy measures of scalp EEG. Journal of Healthcare Engineering. 2015;6(1):55-70. doi:10.1260/2040-2295.6.1.55

236. McCart JA, Berndt DJ, Jarman J, Finch DK, Luther SL. Finding falls in ambulatory care clinical documents using statistical text mining. Journal of the American Medical Informatics Association. 2013;20(5):906-914. doi:10.1136/amiajnl-2012-001334

237. McFadden P, Crim A. Comparison of the effectiveness of interactive didactic lecture versus online simulation-based CME programs directed at improving the diagnostic capabilities of primary care practitioners. Journal of Continuing Education in the Health Professions. 2016;36(1):32-37. doi:10.1097/CEH.0000000000000061

238. Messadi M, Bessaid A, Taleb-Ahmed A. Extraction of specific parameters for skin tumour classification. Journal of Medical Engineering & Technology. 2009;33(4):288-295. doi:10.1080/03091900802451315

239. Meulendijk MC, Spruit MR, Brinkkemper S. Risk mediation in association rules. In: ten Teije A, Popow C, Holmes JH, Sacchi L, eds. Artificial Intelligence in Medicine. Vol 10259. Cham: Springer International Publishing; 2017:327-331. doi:10.1007/978-3-319-59758-4_38

240. Meulendijk M, Spruit M, Lefebvre A, Brinkkemper S. To what extent can prescriptions be meaningfully exchanged between primary care terminologies? A case study of four western European classification systems. IET Software. 2017;11(5):256-264. doi:10.1049/iet-sen.2016.0301

241. Meyer KE, Sather-Levine B, Laurent-Bopp D, Gruenewald D, Nichol P, Kimmerle M. The impact of clinical information systems research on the future of advanced practice nursing. Adv Prac Nurs Q. 1996;2(3):58-64.

242. Mezghani N, Chav R, Humbert L, Parent S, Skalli W, de Guise JA. A computer-based classifier of three-dimensional spinal scoliosis severity. International Journal of Computer Assisted Radiology and Surgery. 2008;3(1-2):55-60. doi:10.1007/s11548-008-0163-3

243. Michalowski M, Michalowski W, O'Sullivan D, Wilk S, Carrier M. AFGuide system to support personalized management of atrial fibrillation. Association for the Advancement of Artificial Intelligence. 2017.

244. Michelson M, Minton SN, See K. A language-modeling approach to health data interoperability. Association for the Advancement of Artificial Intelligence. 2014:8.

245. Mikołajczyk A, Kwasigroch A, Grochowski M. Intelligent system supporting diagnosis of malignant melanoma. In: Mitkowski W, Kacprzyk J, Oprzędkiewicz K, Skruch P, eds. Trends in Advanced Intelligent Control, Optimization and Automation. Vol 577. Cham: Springer International Publishing; 2017:828-837. doi:10.1007/978-3-319-60699-6_79

246. Miller RE, Learned-Miller EG, Trainer P, Paisley A, Blanz V. Early diagnosis of acromegaly: computers vs clinicians: Computer diagnosis of acromegaly. Clinical Endocrinology. 2011;75(2):226-231. doi:10.1111/j.1365-2265.2011.04020.x

247. Millier A, Mejri S, Aballea S A, Toumi M. Health care organizations/services clustering: comparison of several techniques using UK data. In: Value in Health. Vol 15. 2012:A532. doi:10.1016/j.jval.2012.08.1852

248. Mishra S, Yamasaki T, Imaizumi H. Supervised classification of dermatological diseases by deep learning. 2018. http://arxiv.org/abs/1802.03752. Accessed July 3, 2019.

249. Modell M, Iliffe S, Austin A, Leaning MS. From guidelines to decision support in the management of asthma. Health Telematics for Clinical Guidelines and Protocols. 1995:105-113.

250. Moraes E, Brito K, Meira S. CompoPHC: An ontology-based component for primary health care. In: 2012 IEEE 13th International Conference on Information Reuse & Integration (IRI). Las Vegas, NV, USA: IEEE; 2012:592-599. doi:10.1109/IRI.2012.6303063

251. Moran W, Zhang MJ, Su Z, Taber D, Mauldin P, Srinivas T. Using predictive modeling to manage cardiovascular risk and improve kidney transplant outcomes. In: Cardiovascular Complications in Kidney Transplantation. 2016:208.

252. Moretz C, Zhou Y, Dhamane AD, et al. Development and validation of a predictive model to identify individuals likely to have undiagnosed chronic obstructive pulmonary disease using an administrative claims database. Journal of Managed Care & Specialty Pharmacy. 2015;21(12):1149-1159. doi:10.18553/jmcp.2015.21.12.1149

253. Murtaugh MA, Gibson BS, Redd D, Zeng-Treitler Q. Regular expression-based learning to extract bodyweight values from clinical notes. Journal of Biomedical Informatics. 2015;54:186-190. doi:10.1016/j.jbi.2015.02.009

254. Myburgh HC, Jose S, Swanepoel DW, Laurent C. Towards low cost automated smartphone- and cloud-based otitis media diagnosis. Biomedical Signal Processing and Control. 2018;39:34-52. doi:10.1016/j.bspc.2017.07.015

255. Myburgh HC, van Zijl WH, Swanepoel D, Hellström S, Laurent C. Otitis media diagnosis for developing countries using tympanic membrane image-analysis. EBioMedicine. 2016;5:156-160. doi:10.1016/j.ebiom.2016.02.017

256. Nahar J, Ali ABMS, Imam T, Tickle K, Chen P. Brain cancer diagnosis-association rule based computational intelligence approach. In: 2016 IEEE International Conference on Computer and Information Technology (CIT). Nadi, Fiji: IEEE; 2016:89-95. doi:10.1109/CIT.2016.106

257. Newcomer SR, Steiner JF, Bayliss EA. Identifying subgroups of complex patients with cluster analysis. The American Journal of Managed Care. 2011;17(8):e324-e332.

258. Ng K, Stewart WF, Defilippi C, et al. Data driven modeling of electronic health record data to detect pre-diagnostic heart failure in primary care. In: Health Tech. Vol 132. Circulation; 2015:A17713.

259. Ng K, Steinbuhl S, deFilippi C, Dey S, Stewart W. Early detection of heart failure using electronic health records: practical implications for time before diagnosis, data diversity, data quantity, and data density. Journal of Patient-Centered Research and Reviews. 2017;4(3):174-175. doi:10.17294/2330-0698.1523

260. Ni L, Lu C, Liu N, Liu J. MANDY: towards a smart primary care chatbot application. In: Chen J, Theeramunkong T, Supnithi T, Tang X, eds. Knowledge and Systems Sciences. Vol 780. Singapore: Springer Singapore; 2017:38-52. doi:10.1007/978-981-10-6989-5_4

261. Nijeweme-d'Hollosy W, van Velsen L, Poel M, Groothuis-Oudshoorn CGM, Soer R, Hermens H. Evaluation of three machine learning models for self-referral decision support on low back pain in primary care. International Journal of Medical Informatics. 2018;110:31-41. doi:10.1016/j.ijmedinf.2017.11.010

262. Nijeweme-dHollosy WO, Velsen L van, Groothuis-Oudshoorn KGM, Soer R, Hermens H. Should I see a healthcare professional or can I perform self-care: self-referral decision support for patients with low back pain. In: 2016 IEEE International Conference on Healthcare Informatics (ICHI). Chicago, IL, USA: IEEE; 2016:495-503. doi:10.1109/ICHI.2016.90

263. Nikolova I, Boytcheva S, Angelova G, Angelov Z. Combining structured and free textual data of diabetic patients' smoking status. In: Dichev C, Agre G, eds. Artificial Intelligence: Methodology, Systems, and Applications. Vol 9883. Cham: Springer International Publishing; 2016:57-67. doi:10.1007/978-3-319-44748-3_6

264. O'Bryant SE, Edwards M, Johnson L, et al. A blood screening test for Alzheimer's disease. Alzheimer's & Dementia: Diagnosis, Assessment & Disease Monitoring. 2016;3:83-90. doi:10.1016/j.dadm.2016.06.004

265. O'Bryant SE, Xiao G, Zhang F, et al. Validation of a serum screen for Alzheimer's Disease across assay platforms, species, and tissues. Journal of Alzheimer's Disease. 2014;42(4):1325-1335. doi:10.3233/JAD-141041

266. Odunukan OW, Elayavilli RK, Cha S, et al. Natural language processing identifies important side-effects In very elderly patients With hypertension. In: Vol 64. Hypertension; 2014:A239.

267. Orient JM. Evaluation of abdominal pain: clinicians' performance compared with three protocols. South Med J. 1986;79(7):793-799.

268. Orsi T, Araujo E, Simoes R. Fuzzy chest pain assessment for unstable angina based on Braunwald symptomatic and obesity clinical conditions. In: 2014 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE). Beijing, China: IEEE; 2014:1076-1082. doi:10.1109/FUZZ-IEEE.2014.6891836

269. Ou MH, West GAW, Lazarescu M, Clay CD. Evaluation of TELEDERM for dermatological services in rural and remote areas. Artificial Intelligence in Medicine. 2008;44(1):27-40. doi:10.1016/j.artmed.2008.04.006

270. Pakhomov SVS, Shah ND, Van Houten HK, Hanson PL, Smith SA. The role of the electronic medical record in the assessment of health related quality of life. AMIA Annu Symp Proc. 2011;2011:1080-1088.

271. Passi K, Zhao H. A decision support system (DSS) for colorectal cancer follow-up program via a semantic framework: International Journal of Healthcare Information Systems and Informatics. 2015;10(1):17-38. doi:10.4018/IJHISI.2015010102

272. Perlini S, Piepoli M, Marti G, et al. Treatment of chronic heart failure: an expert system advisor for general practitioners (*). Acta Cardiologica. 1990;XLV(5):365-378.

273. Perveen S, Shahbaz M, Guergachi A, Keshavjee K. Performance analysis of data mining classification techniques to predict diabetes. Procedia Computer Science. 2016;82:115-121. doi:10.1016/j.procs.2016.04.016

274. Pietilä A, El-Segaier M, Vigário R, Pesonen E. Blind source separation of cardiac murmurs from heart recordings. In: Rosca J, Erdogmus D, Príncipe JC, Haykin S, eds. Independent Component Analysis and Blind Signal Separation. Vol 3889. Berlin, Heidelberg: Springer Berlin Heidelberg; 2006:470-477. doi:10.1007/11679363_59

275. Piury J, Laita LM, Roanes-Lozano E, et al. A Gröbner bases-based rule based expert system for fibromyalgia diagnosis. Revista de la Real Academia de Ciencias Exactas, Fisicas y Naturales Serie A Matematicas. 2012;106(2):443-456. doi:10.1007/s13398-012-0064-8

276. Pollettini JT, Nicolas FP, Panico SG, et al. A software architecture-based framework supporting suggestion of medical surveillance level from classification of electronic patient records. In: 2009 International Conference on Computational Science and Engineering. Vancouver, BC, Canada: IEEE; 2009:166-173. doi:10.1109/CSE.2009.231

277. Poulin C, Shiner B, Thompson P, et al. Predicting the risk of suicide by analyzing the text of clinical notes. Brusic V, ed. PLoS ONE. 2014;9(1):e85733. doi:10.1371/journal.pone.0085733

278. Prasanna P, Jain S, Bhagat N, Madabhushi A. Decision support system for detection of diabetic retinopathy using smartphones. In: Proceedings of the ICTs for Improving Patients Rehabilitation Research Techniques. Venice, Italy: IEEE; 2013. doi:10.4108/icst.pervasivehealth.2013.252093

279. Ptochos D, Panopoulos D, Metaxiotis K, Askounis D, Psarras J. Using internet GIS technology for early warning, response and controlling the quality of the public health sector. International Journal of Electronic Healthcare. 2004;1(1):78. doi:10.1504/IJEH.2004.004661

280. Quiceno-Manrique AF, Godino-Llorente JI, Blanco-Velasco M, Castellanos-Dominguez G. Selection of dynamic features based on time–frequency representations for heart murmur detection from phonocardiographic signals. Annals of Biomedical Engineering. 2010;38(1):118-137. doi:10.1007/s10439-009-9838-3

281. Rahimi A, Liaw S-T, Taggart J, Ray P, Yu H. Validating an ontology-based algorithm to identify patients with Type 2 Diabetes Mellitus in electronic health records. International Journal of Medical Informatics. 2014;83(10):768-778. doi:10.1016/j.ijmedinf.2014.06.002

282. Rajan V, Bhattacharya S, Shetty R, Sitaram A, Vivek G. Clinical decision support for stroke using multi–view learning based models for NIHSS scores. In: Cao H, Li J, Wang R, eds. Trends and Applications in Knowledge Discovery and Data Mining. Vol 9794. Cham: Springer International Publishing; 2016:190-199. doi:10.1007/978-3-319-42996-0_16

283. Rajkomar A, Yim JWL, Grumbach K, Parekh A. Weighting primary care patient panel size: a novel electronic health record-derived measure using machine learning. JMIR Medical Informatics. 2016;4(4):e29. doi:10.2196/medinform.6530

284. Raluy-Callado M, Cox A, MacLachlan S, Gabriel S, Dinet J. Resource utilisation and costs in patients with post-stroke spasticity In the United Kingdom. Value in Health. 2015;18(7):A386. doi:10.1016/j.jval.2015.09.840

285. Raluy-Callado M, Cox A, MacLachlan S, et al. A retrospective study to assess resource utilization and costs in patients with post-stroke spasticity in the United Kingdom. Current Medical Research and Opinion. 2018;34(7):1317-1324. doi:10.1080/03007995.2018.1447449

286. Rao RR, Makkithaya K, Gupta N. Ontology based semantic representation for Public Health data integration. In: 2014 International Conference on Contemporary Computing and Informatics (IC3I). Mysore, India: IEEE; 2014:357-362. doi:10.1109/IC3I.2014.7019701

287. Ray HN, Boxwala AA, Anantraman V, Ohno-Machado L. Providing context-sensitive decision-support based on WHO guidelines. AMIA. 2002:637-641.

288. Reece AG, Danforth CM. Instagram photos reveal predictive markers of depression. EPJ Data Science. 2017;6(1). doi:10.1140/epjds/s13688-017-0110-z

289. Reed C, Boswell B, Neville R. Multi-agent patient representation in primary care. In: Miksch S, Hunter J, Keravnou ET, eds. Artificial Intelligence in Medicine. Vol 3581. Berlin, Heidelberg: Springer Berlin Heidelberg; 2005:375-384. doi:10.1007/11527770_50

290. Rezaeiahari M, Poranki S, Khasawneh M. Predicting emergency department visits among patients receiving primary care using risk scores. In: Proceedings of the 2017 Industrial and Systems Engineering Conference. ; 2017:6.

291. Riaño D, Fernández-Pérez A. Simulation-based episodes of care data synthetization for chronic disease patients. In: Riaño D, Lenz R, Reichert M, eds. Knowledge Representation for Health Care. Vol 10096. Cham: Springer International Publishing; 2017:36-50. doi:10.1007/978-3-319-55014-5_3

292. Rodbard D, Vigersky RA. Design and development of a computer assisted clinical decision support system to help physicians manage patients with type 2 diabetes mellitus. In: Tools for Improving Diabetes Control. Vol 53. Diabetologia; 2010:S414. doi:10.1007/s00125-010-1872-z

293. Rodríguez-Solano C, Laita LM, Roanes-Lozano E, López-Corral L, Laita L. A computational system for diagnosis of depressive situations. Expert Systems with Applications. 2006;31(1):47-55. doi:10.1016/j.eswa.2005.09.011

294. Romero-Córdoba R, Olivas JA, Romero FP, Alonso-Gonzalez F, Serrano-Guerrero J. An application of fuzzy prototypes to the diagnosis and treatment of fuzzy diseases: diagnosis and treatment of fuzzy diseases. International Journal of Intelligent Systems. 2017;32(2):194-210. doi:10.1002/int.21836

295. Røst TB, Sørby ID, Seland G. Development of a medication reconciliation tool for Norwegian primary care EPR systems: experiences from a user-initiated project. In: Proceedings of the 2nd European Workshop on Practical Aspects of Health Informatics. Trondheim, Norway: ceur-ws; 2014:10.

296. Rother A-K, Schwerk N, Brinkmann F, Klawonn F, Lechner W, Grigull L. Diagnostic support for selected paediatric pulmonary diseases using answer-pattern recognition in questionnaires based on combined data mining applications—a monocentric observational pilot study. Hartl D, ed. PLOS ONE. 2015;10(8):e0135180. doi:10.1371/journal.pone.0135180

297. Saadé RG, Tsoukas A, Tsoukas G. Prototyping a decision support system in the clinical environment: assessment of patients with osteoporosis OSTEODSS. Expert Systems with Applications. 2004;27(3):427-438. doi:10.1016/j.eswa.2004.05.019

298. Safarova MS, Liu H, Kullo IJ. Rapid identification of familial hypercholesterolemia from electronic health records: The SEARCH study. Journal of Clinical Lipidology. 2016;10(5):1230-1239. doi:10.1016/j.jacl.2016.08.001

299. Safran C, Rind DM, Davis RB, et al. A clinical trial of a knowledge-based medical record. MEDINFO. 1995;95:1076-1080.

300. Safran C, Rind DM, Davis RM, et al. An electronic medical record that helps care for patients with HIV infection. AMIA. 1994:224-228.

301. Salama MA, Soliman OS, Maglogiannis I, Hassanien AE, Fahmy AA. Rough set-based identification of heart valve diseases using heart sounds. In: Skowron A, Suraj Z, eds. Rough Sets and Intelligent Systems - Professor Zdzisław Pawlak in Memoriam. Vol 43. Berlin, Heidelberg: Springer Berlin Heidelberg; 2013:475-491. doi:10.1007/978-3-642-30341-8_25

302. Sali R, Roohafza H, Sadeghi M, Andalib E, Shavandi H, Sarrafzadegan N. Validation of the revised stressful life event questionnaire using a hybrid model of genetic algorithm and artificial neural networks. Computational and Mathematical Methods in Medicine. 2013;2013:1-7. doi:10.1155/2013/601640

303. Salmeron JL, Rahimi SA, Navali AM, Sadeghpour A. Medical diagnosis of Rheumatoid Arthritis using data driven PSO–FCM with scarce datasets. Neurocomputing. 2017;232:104-112. doi:10.1016/j.neucom.2016.09.113

304. Saraiva R, Perkusich M, Silva L, Almeida H, Siebra C, Perkusich A. Early diagnosis of gastrointestinal cancer by using case-based and rule-based reasoning. Expert Systems with Applications. 2016;61:192-202. doi:10.1016/j.eswa.2016.05.026

305. Scala PL, Pasquale D, Tresoldi D, Lafortuna CL, Rizzo G, Padula M. Ontology-supported clinical profiling for the evaluation of obesity and related comorbidities. Studies in Health Technology and Informatics. 2012:1025–1029. doi:10.3233/978-1-61499-101-4-1025

306. Schilithz AOC, Kale PL, Gama SGN, Nobre FF. Risk groups in children under six months of age using self-organizing maps. Computer Methods and Programs in Biomedicine. 2014;115(1):1-10. doi:10.1016/j.cmpb.2014.02.011

307. Schilling C, Mortimer D, Dalziel K, Heeley E, Chalmers J, Clarke P. Using classification and regression trees (CART) to identify prescribing thresholds for cardiovascular disease. PharmacoEconomics. 2016;34(2):195-205. doi:10.1007/s40273-015-0342-3

308. Sen A, Sinha AP. An ontological model of the practice transformation process. Journal of Biomedical Informatics. 2016;61:298-318. doi:10.1016/j.jbi.2016.05.001

309. Seroussi B, Bouaud J, Dreau H, et al. ASTI: a guideline-based drug-ordering system for primary care. MEDINFO. 2001;2001:528-532.

310. Shalev V, Kinar Y, Kalkstein N, et al. Computational analysis of blood counts significantly increases detection rate of gastric and colorectal cancers. In: Esophageal, Gastric and Duodenal Disorders. Vol 28. Journal of Gastroenterology and Hepatology; 2013:761.

311. Shankar RD, Martins SB, Tu SW, Goldstein MK, Musen MA. Building an explanation function for a hypertension decision-support system. MEDINFO. 2001;2001:538-542.

312. Shavandi H, Mahlooji H. Fuzzy hierarchical queueing models for the location set covering problem in congested systems. Scientia Iranica. 2008;15(3):378-388.

313. Shen Y, Yuan K, Chen D, et al. An ontology-driven clinical decision support system (IDDAP) for infectious disease diagnosis and antibiotic prescription. Artificial Intelligence in Medicine. 2018;86:20-32. doi:10.1016/j.artmed.2018.01.003

314. Siepmann JP, Bachman JW. HTN-APT: computer aid in hypertension management. The Journal of Family Practice. 1987;24(3):313-316.

315. Simon A. A multivariate portable noninvasive diagnostic approach to actively assess cognitive function in Alzheimer's disease (AD) and mild cognitive impairment (MCI). Alzheimer's & Dementia. 2012;8(4):115-116. doi:10.1016/j.jalz.2012.05.299

316. Simon A. A multi-modal biosensor-based approach to aid in the diagnosis of concussion and Alzheimer's Disease. In: Vol 10. Alzheimer's & Dementia; 2014:662. doi:10.1016/j.jalz.2014.05.1181

317. Simon A, Ashrafiuon H, Ghorbanian P, Devilbiss D. A noninvasive device-based approach to aid in the diagnosis of mTBI and Alzheimer's disease: Preliminary findings from Clinical Pilot Studies. Alzheimer's & Dementia. 2013;9(4):P222. doi:10.1016/j.jalz.2013.05.417

318. Simon A, Devilbiss D. Multivariate models of biosensor data to actively assess sports concussion and mild traumatic brain injury. In: Vol 86. Neurology; 2016.

319. Sinnott MM, Carr B, Markey J, et al. Knowledge based lipid management system for general practitioners. Clinica Chimica Acta. 1993;222(1-2):71-77. doi:10.1016/0009-8981(93)90093-J

320. Sola D, Borioli GS, Quaglia R. Predicting GPs' engagement with artificial intelligence. British Journal of Healthcare Management. 2018;24(3):134-140. doi:10.12968/bjhc.2018.24.3.134

321. Soler JK, Corrigan D, Kazienko P, et al. Evidence-based rules from family practice to inform family practice; the learning healthcare system case study on urinary tract infections. BMC Family Practice. 2015;16(63). doi:10.1186/s12875-015-0271-4

322. Solms F, Smit E, Nel ZJ. A neural network diagnostic tool for the chronic fatigue syndrome. In: Proceedings of International Conference on Neural Networks (ICNN'96). Vol 2. Washington, DC, USA: IEEE; 1996:778-781. doi:10.1109/ICNN.1996.548995

323. Song MH, Lee YH, Kang UG. Comparison of machine learning algorithms for classification of the sentences in three clinical practice guidelines. Healthcare Informatics Research. 2013;19(1):16. doi:10.4258/hir.2013.19.1.16

324. Sourla E, Syrimpeis V, Stamatopoulou K-M, Merekoulias G, Tsakalidis A, Tzimas G. Exploiting fuzzy expert systems in cardiology. In: Iliadis L, Papadopoulos H, Jayne C, eds. Engineering Applications of Neural Networks. Vol 384. Berlin, Heidelberg: Springer Berlin Heidelberg; 2013:80-89. doi:10.1007/978-3-642-41016-1_9

325. South BR, Shen S, Jones M, et al. Developing a manually annotated clinical document corpus to identify phenotypic information for inflammatory bowel disease. BMC Bioinformatics. 2009;10(S9). doi:10.1186/1471-2105-10-S9-S12

326. Spyropoulos J, Boutsalis G, Chatterjee P, LaCouture M. Improving clinical decisions in dyslipidemia: simulation-based medical education in action. In: Innovations in Practice Management and Social Media. Vol 69. Journal of the American College of Cardiology; 2017:2502. doi:10.1016/S0735-1097(17)35891-6

327. Srinivas S, Ravindran AR. Optimizing outpatient appointment system using machine learning algorithms and scheduling rules: A prescriptive analytics framework. Expert Systems with Applications. 2018;102:245-261. doi:10.1016/j.eswa.2018.02.022

328. St-Maurice J, Kuo M-H, Gooch P. A proof of concept for assessing emergency room use with primary care data and natural language processing. Methods of Information in Medicine. 2013;52(01):33-42. doi:10.3414/ME12-01-0012

329. St-Maurice J, Kuo MH. Analyzing primary care data to characterize inappropriate emergency room use. Studies in Health Technology and Informatics. 2012:990–994. doi:10.3233/978-1-61499-101-4-990

330. Stan O, Avram C, Stefan I, Astilean A. Integrated innovative solutions to improve healthcare scheduling. In: 2016 IEEE International Conference on Automation, Quality and Testing, Robotics (AQTR). Cluj-Napoca, Romania: IEEE; 2016:1-6. doi:10.1109/AQTR.2016.7501301

331. Stanganelli I, Brucale A, Calori L, et al. Computer-aided diagnosis of melanocytic lesions. Anticancer Res. 2005;25:4577-4582.

332. Šter M, Švab I, Šter B. Prediction of intended career choice in family medicine using artificial neural networks. European Journal of General Practice. 2015;21(1):63-69. doi:10.3109/13814788.2014.933314

333. Stiglic G, Kokol P. Intelligent patient and nurse scheduling in ambulatory health care centers. In: 2005 IEEE Engineering in Medicine and Biology 27th Annual Conference. Shanghai, China: IEEE; 2005:5475-5478. doi:10.1109/IEMBS.2005.1615722

334. Suhasini A, Palanivel S, Ramalingam V. Multimodel decision support system for psychiatry problem. Expert Systems with Applications. 2011;38(5):4990-4997. doi:10.1016/j.eswa.2010.09.152

335. Sumner 2nd W, Hagen MD, Rovinelli R. The item generation methodology of an empiric simulation project. Adv Health Sci Educ. 1999;4:49-66.

336. Sumner 2nd W, Truszczynski M, Marek VM. Simulating patients with parallel health state networks. AMIA. 1998:438-442.

337. Sumner 2nd. W, Xu JZ. Modeling fatigue. In: AMIA Annual Symposium Proceedings. 2002:747-751.

338. Sumner 2nd. W, Xu JZ, Roussel G, Hagen MD. Modeling relief. In: AMIA Annual Symposium Proceedings. 2007:706-710.

339. Syed Z, Leeds D, Curtis D, Nesta F, Levine RA, Guttag J. A framework for the analysis of acoustical cardiac signals. IEEE Transactions on Biomedical Engineering. 2007;54(4):651-662. doi:10.1109/TBME.2006.889189

340. Syiam MM. A neural network expert system for diagnosing eye diseases. In: Proceedings of the Tenth Conference on Artificial Intelligence for Applications. San Antonia, TX, USA: IEEE; 1994:491-492. doi:10.1109/CAIA.1994.323624

341. Szecsenyi J, Buschhorn A, Kochen M. General practitioners' attitudes towards future developments in practice computing—a representative survey in the north of Germany. Family Practice. 1992;9(3):357-361. doi:10.1093/fampra/9.3.357

342. Tan A, Liang MW, Chio MT, et al. Designing a mobile imaging system for early melanoma detection. In: 6th Singapore Health & Biomedical Congress. Vol 44. Singapore: Annals of the Academy of Medicine; 2015:S187. doi:10.1016/j.jaad.2015.02.366

343. Tape TG, Rn RAS, Campbell JR. Implementing guidelines in ambulatory practice. AMIA. 1993:806-807.

344. Teipel SJ, Kilimann I, Dyrba M, Brüggen K, Thyrian JR, Hoffmann W. Structural connectivity as a biomarker for Alzheimer's disease: Evaluation in a multicenter trial and a primary care cohort. Alzheimer's & Dementia. 2015;11(7):P696. doi:10.1016/j.jalz.2015.06.1545

345. Tekin C, Atan O, Van Der Schaar M. Discover the expert: context-adaptive expert selection for medical diagnosis. IEEE Transactions on Emerging Topics in Computing. 2015;3(2):220-234. doi:10.1109/TETC.2014.2386133

346. Teles G, Oliveira C, Braga R, et al. Using Bayesian networks to improve the decision-making process in public health systems. In: 2014 IEEE 16th International Conference on E-Health Networking, Applications and Services (Healthcom). Natal: IEEE; 2014:565-570. doi:10.1109/HealthCom.2014.7001904

347. Thurin N, Castilloux AM, Reich C, Hermann R, Frise S, Moride Y. Assessing the feasibility and performance of the HAWK electronic medical records for drug safety surveillance. In: Drug Safety. Vol 38. Prague, Czech Republic: Drug Safety; 2015:1033. doi:10.1007/s40264-015-0346-0

348. Timpka T. Knowledge-based decision support for general practitioners: an integrated design. Computer Methods and Programs in Biomedicine. 1987;25:49-60.

349. Ting SL, Kwok SK, Tsang AHC, Lee WB. A hybrid knowledge-based approach to supporting the medical prescription for general practitioners: Real case in a Hong Kong medical center. Knowledge-Based Systems. 2011;24(3):444-456. doi:10.1016/j.knosys.2010.12.011

350. Tirunagari S, Bull SC, Vehtari A, Farmer C, de Lusignan S, Poh N. Automatic detection of acute kidney injury episodes from primary care data. In: 2016 IEEE Symposium Series on Computational Intelligence (SSCI). Athens, Greece: IEEE; 2016:1-6. doi:10.1109/SSCI.2016.7849885

351. Trevisan J, Angelov PP, Carmichael PL, Scott AD, Martin FL. Extracting biological information with computational analysis of Fourier-transform infrared (FTIR) biospectroscopy datasets: current practices to future perspectives. The Analyst. 2012;137(14):3202. doi:10.1039/c2an16300d

352. Trtica-Majnaric L, Zekic-Susac M, Sarlija N, Vitale B. Prediction of influenza vaccination outcome by neural networks and logistic regression. Journal of Biomedical Informatics. 2010;43(5):774-781. doi:10.1016/j.jbi.2010.04.011

353. Van der Niepen P, Woestenburg A, Brié H, et al. Effectiveness of valsartan for treatment of hypertension: patient profiling and hierarchical modeling of determinants and outcomes (the Preview Study). Annals of Pharmacotherapy. 2009;43(5):849-861. doi:10.1345/aph.1L576

354. Van Voorhees BW, Paunesku D, Gollan J, Kuwabara S, Reinecke M, Basu A. Predicting future risk of depressive episode in adolescents: The Chicago Adolescent Depression Risk Assessment (CADRA). The Annals of Family Medicine. 2008;6(6):503-511. doi:10.1370/afm.887

355. VanQuekelberghe P, Jakob T, Hoffmann D, Wetter T, Finkeissen E. Minimalist knowledge representation of primary care diseases in the medrapid.info knowledge base. Journal of Innovation in Health Informatics. 2005;13(4):239-248. doi:10.14236/jhi.v13i4.603

356. Vega R, Sanchez-Ante G, Falcon-Morales LE, Sossa H, Guevara E. Retinal vessel extraction using Lattice Neural Networks with dendritic processing. Computers in Biology and Medicine. 2015;58:20-30. doi:10.1016/j.compbiomed.2014.12.016

357. Velickovski F, Ceccaroni L, Marti R, et al. Automated spirometry quality assurance: supervised learning from multiple experts. IEEE Journal of Biomedical and Health Informatics. 2018;22(1):276-284. doi:10.1109/JBHI.2017.2713988

358. Vijayakrishnan R, Steinhubl SR, Ng K, et al. Prevalence of heart failure signs and symptoms in a large primary care population identified through the use of text and data mining of the electronic health record. Journal of Cardiac Failure. 2014;20(7):459-464. doi:10.1016/j.cardfail.2014.03.008

359. Viveros MS, Nearhos JP, Rothman MJ. Applying data mining techniques to a health insurance information system. In: Proceedings of the 22nd VLDB Conference. Mumbai (Bombay), India: VLDB; 1996:286-294.

360. Vos R, Aarts S, van Mulligen E, et al. Finding potentially new multimorbidity patterns of psychiatric and somatic diseases: exploring the use of literature-based discovery in primary care research. JAMIA. 2014;21(1):139-145. doi:10.1136/amiajnl-2012-001448

361. Walton R, Ilic Z. Knowledge engineering for drug prescribing guidelines. Health Telematics for Clinical Guidelines and Protocols. 1995:75-85.

362. Wang W, Richards G, Rea S. Hybrid data mining ensemble for predicting osteoporosis risk. In: 2005 IEEE Engineering in Medicine and Biology 27th Annual Conference. Shanghai, China: IEEE; 2005:886-889. doi:10.1109/IEMBS.2005.1616557

363. Wang W, Rea S. Intelligent ensemble system aids osteoporosis early detection. In: Proceedings of the 6th WSEAS Int. Conf. on Evolutionary Computing. Lisbon, Portugal; 2005:123-128.

364. Wang X, Valdez TA, Bi J. Detecting tympanostomy tubes from otoscopic images via offline and online training. Computers in Biology and Medicine. 2015;61:107-118. doi:10.1016/j.compbiomed.2015.03.025

365. Wang Y, Patrick J, Miller G, O'Hallaran J. A computational linguistics motivated mapping of ICPC-2 PLUS to SNOMED CT. BMC Medical Informatics and Decision Making. 2008;8(S1). doi:10.1186/1472-6947-8-S1-S5

366. Wang Y, Chen H, Li R, Duan N, Lewis-Fernández R. Prediction-based structured variable selection through the receiver operating characteristic curves. Biometrics. 2011;67(3):896-905. doi:10.1111/j.1541-0420.2010.01533.x

367. Wang Z, Shah AD, Tate AR, Denaxas S, Shawe-Taylor J, Hemingway H. Extracting diagnoses and investigation results from unstructured text in electronic health records by semi-supervised machine learning. Brusic V, ed. PLoS ONE. 2012;7(1):e30412. doi:10.1371/journal.pone.0030412

368. Warren, Warren, Freedman. Interviewing expertise in primary care medicine: a knowledge-based support system. In: Proceedings of the Twenty-Seventh Hawaii International Conference on System Sciences HICSS-94. Wailea, HI, USA: IEEE Comput. Soc. Press; 1994:173-182. doi:10.1109/HICSS.1994.323354

369. Warren J, Beliakov G, van der Zwaag B. Fuzzy logic in clinical practice decision support systems. In: Proceedings of the 33rd Annual Hawaii International Conference on System Sciences. Vol vol.1. Maui, HI, USA: IEEE Comput. Soc; 2000:10. doi:10.1109/HICSS.2000.926789

370. Warren JR, Davidovic A, Spenceley S, Bolton P. Mediface: anticipative data entry interface for general practitioners. In: Proceedings 1998 Australasian Computer Human Interaction Conference. OzCHI'98 (Cat. No.98EX234). Adelaide, SA, Australia: IEEE Comput. Soc; 1998:192-199. doi:10.1109/OZCHI.1998.732214

371. Warren JR, Posey B, Thornton T, ParangA P. Can computer autoacquisition of medical information meet the needs of the future? A feasibility study in direct computation of the fine grained electronic medical record. AMIA. 1999:445-449.

372. Warren JR, Tyerman SP. Webifying a patient interview support application. Medical Informatics. 1998;23(1):63-74. doi:10.3109/14639239809001392

373. Watrous RL, Thompson WR, Ackerman SJ. The impact of computer-assisted auscultation on physician referrals of asymptomatic patients with heart murmurs. Clinical Cardiology. 2008;31(2):79-83. doi:10.1002/clc.20185

374. Weng SF, Reps J, Kai J, Garibaldi JM, Qureshi N. Can machine-learning improve cardiovascular risk prediction using routine clinical data? Liu B, ed. PLOS ONE. 2017;12(4):e0174944. doi:10.1371/journal.pone.0174944

375. Wenner AR, Ferrante M, Belser D. Instant medical history. AMIA. 1994:1036.

376. Wickramasinghe K, Alahakoon D, Schattner P, Georgeff M. Self-organizing maps for translating health care knowledge: a case study in diabetes management. In: Wang D, Reynolds M, eds. AI 2011: Advances in Artificial Intelligence. Vol 7106. Berlin, Heidelberg: Springer Berlin Heidelberg; 2011:162-171. doi:10.1007/978-3-642-25832-9_17

377. Widmer G, Horn W, Nagele B. Automatic knowledge base refinement: Learning from examples and deep knowledge in rheumatology. Artif Intell Med. 1993;5(3):225-243. doi:10.1016/0933-3657(93)90026-Y

378. Witteman CL, Renooij S, Koele P. Medicine in words and numbers: a cross-sectional survey comparing probability assessment scales. BMC Medical Informatics and Decision Making. 2007;7(1). doi:10.1186/1472-6947-7-13

379. Wright A, Ricciardi TN, Zwick M. Application of information-theoretic data mining techniques in a national ambulatory practice outcomes research network. AMIA. 2005:829-833.

380. Wu J, Roy J, Stewart WF. Prediction modeling using EHR data. Medical Care. 2010;48(6):8.

381. Yamauchi K, Fukatsu T. A decision support system for diagnostic consultation in laboratory tests. MEDINFO. 1995;95:1034.

382. Yin Z, Zhao Y, Lu X, Duan H. A hybrid intelligent diagnosis approach for quick screening of Alzheimer's Disease based on multiple neuropsychological rating scales. Computational and Mathematical Methods in Medicine. 2015;2015:1-13. doi:10.1155/2015/258761

383. Zahlmann G, Kochner B, Ugi I, et al. Hybrid fuzzy image processing for situation assessment. A Knowledge-based system for early detection of diabetic retinopathy. IEEE Engineering in Medicine and Biology. 2000:76-83.

384. Zahlmann G, Scherf M, Wegner A, Obermaier M, Mertz M. Situation assessment of glaucoma using a hybrid fuzzy neural network. IEEE Engineering in Medicine and Biology Magazine. 2000;19(1):84-91. doi:10.1109/51.816247

385. Zamora A, Carrion C, Vazquez G, et al. A clinical decision support system can improves the quality of lipid-lowering therapy in coronary patients. In: Vol 36. European Heart Journal; 2015:285-286.

386. Zamora M, Baradad M, Amado E, et al. Characterizing chronic disease and polymedication prescription patterns from electronic health records. In: 2015 IEEE International Conference on Data Science and Advanced Analytics (DSAA). Campus des Cordeliers, Paris, France: IEEE; 2015:1-9. doi:10.1109/DSAA.2015.7344870

387. Zavrsnik J, Kancler K, Kokol P, Mernik M, Malcic I. The computer tool of the mitral valve prolapse determination based on automatic learning. In: Proceedings of IEEE Symposium on Computer-Based Medical Systems (CBMS). Winston-Salem, NC, USA: IEEE Comput. Soc. Press; 1994:170-175. doi:10.1109/CBMS.1994.316006

388. Zavrsnik J, Kokol P, Maleiae I, Kancler K, Mernik M, Bigec M. ROSE: Decision trees, automatic learning, and their applications in cardiac medicine. MEDINFO. 1995;95:1688.

389. Zeuner R, Gsell U, Hubenthal M, et al. Popgen-ossa: development of an organ specific self assessment (OSSA) for interdisciplinary documentation of patient reported clinical outcomes. In: Validation of Outcome Measures and Biomarkers. BMJ Annals of the Rheumatic Diseases; 2017:465. doi:10.1136/annrheumdis-2017-eular.4722

390. Zhang H, Plutzky J, Shubina M, Turchin A. Continued statin prescriptions after adverse reactions and patient outcomes: a cohort study. Annals of Internal Medicine. 2017;167(4):221-228. doi:10.7326/M16-0838

391. Zhang J, Gajjala S, Agrawal P, et al. A computer vision pipeline for automated determination of cardiac structure and function and detection of disease by two-dimensional echocardiography. 2017. http://arxiv.org/abs/1706.07342. Accessed July 3, 2019.

392. Zhang J, Xiong H, Huang Y, Wu H, Leach K, Barnes LE. M-SEQ: Early detection of anxiety and depression via temporal orders of diagnoses in electronic health data. In: 2015 IEEE International Conference on Big Data (Big Data). Santa Clara, CA, USA: IEEE; 2015:2569-2577. doi:10.1109/BigData.2015.7364054

393. Zhang Q, Zhang G, Lu J, Wu D. A framework of hybrid recommender system for personalized clinical prescription. In: 2015 10th International Conference on Intelligent Systems and Knowledge Engineering (ISKE). Taipei, Taiwan: IEEE; 2015:189-195. doi:10.1109/ISKE.2015.98

394. Zheng K, Mei Q, Yang L, Manion FJ, Balis UJ, Hanauer DA. Voice-dictated versus typed-in clinician notes: linguistic properties and the potential implications on natural language processing. AMIA Annual Symposium Proceedings Archive. 2011;2011:1630-1638.

395. Zhou S-M, Fernandez-Gutierrez F, Kennedy J, et al. Defining disease phenotypes in primary care electronic health records by a machine learning approach: a case study in identifying rheumatoid arthritis. Pappalardo F, ed. PLOS ONE. 2016;11(5):e0154515. doi:10.1371/journal.pone.0154515

396. Zhou S-M, Muhammad RA, Sheppard S, Howe R, Lyons RA, Brophy S. Mining electronic health records to identify predictive factors associated with hospital admission for Campylobacter infections. The Lancet. 2017;390:S99. doi:10.1016/S0140-6736(17)33034-9

397. Zhou S-M, Rahman MA, Atkinson M, Brophy S. Mining textual data from primary healthcare records: Automatic identification of patient phenotype cohorts. In: 2014 International Joint Conference on Neural Networks (IJCNN). Beijing, China: IEEE; 2014:3621-3627. doi:10.1109/IJCNN.2014.6889494

398. Zhu J, Min C, Wang F. An ontology based intelligent telehealth system for long-term management of hypertension. In: Zhang Y-T, ed. The International Conference on Health Informatics. Vol 42. Cham: Springer International Publishing; 2014:80-83. doi:10.1007/978-3-319-03005-0_21

399. Zhuang ZY, Churilov L, Sikaris K. Uncovering the patterns in pathology ordering by Australian general practitioners: a data mining perspective. In: Proceedings of the 39th Annual Hawaii International Conference on System Sciences (HICSS'06). Kauia, HI, USA: IEEE; 2006:92c-92c. doi:10.1109/HICSS.2006.513

400. Zhuang ZY, Churilov L, Burstein F, Sikaris K. Combining data mining and case-based reasoning for intelligent decision support for pathology ordering by general practitioners. Eur J Oper Res. 2009;195(3):662-675. doi:10.1016/j.ejor.2007.11.003

401. Zhuang ZY, Wilkin CL, Ceglowski A. A framework for an intelligent decision support system: A case in pathology test ordering. Decis Support Sys. 2013;55(2):476-487. doi:10.1016/j.dss.2012.10.006

402. Zhuang ZY, Amarasiri R, Churilov L, Alahakoon D, Sikaris K. Exploring the clinical notes of pathology ordering by Australian general practitioners: a text mining perspective. In: 2007 40th Annual Hawaii International Conference on System Sciences (HICSS'07). Waikoloa, HI: IEEE; 2007:136-136. doi:10.1109/HICSS.2007.220

403. Zortea M, Schopf TR, Thon K, et al. Performance of a dermoscopy-based computer vision system for the diagnosis of pigmented skin lesions compared with visual evaluation by experienced dermatologists. Artificial Intelligence in Medicine. 2014;60(1):13-26. doi:10.1016/j.artmed.2013.11.006

404. Zouridakis G, Wadhawan T, Situ N, et al. Melanoma and other skin lesion detection using smart handheld devices. In: Rasooly A, Herold KE, eds. Mobile Health Technologies. Vol 1256. New York, NY: Springer New York; 2015:459-496. doi:10.1007/978-1-4939-2172-0_30

405. Zvarova J, Peleska J, Hanzlicek P, Zvara K. Enhanced care of hypertensive patients using the Internet. Medical Informatics and the Internet in Medicine. 2002;27(3):161-168. doi:10.1080/1463923021000034651

Appendix D

Objective 2 Extended Information

D.1 RECORD Checklist

	Item No.	STROBE items	Location in manuscript where items are reported	RECORD items	Location in manuscript where items are reported
Title and abstract					
	1	(a) Indicate the study's design with a commonly used term in the title or the abstract (b) Provide in the abstract an informative and balanced summary of what was done and what was found	Abstract	RECORD 1.1: The type of data used should be specified in the title or abstract. When possible, the name of the databases used should be included. RECORD 1.2: If applicable, the geographic region and timeframe within which the study took place should be reported in the title or abstract. RECORD 1.3: If linkage between databases was conducted for the study, this should be clearly stated in the title or abstract.	Abstract
Introduction					
Background rationale	2	Explain the scientific background and rationale for the investigation being reported	Introduction		Introduction
Objectives	3	State specific objectives, including any prespecified hypotheses	Introduction		Introduction
Methods					
Study Design	4	Present key elements of study design early in the paper	Introduction, Methods – Study population and data source		Introduction, Methods – Study

					population and data source
Setting	5	Describe the setting, locations, and relevant dates, including periods of recruitment, exposure, follow-up, and data collection	Methods – Study population and data source; additional details specific to analyses are presented under the appropriate sub-heading		Methods – Study population and data source; additional details specific to analyses are presented under the appropriate sub-heading
Participants	6	*(a) Cohort study* - Give the eligibility criteria, and the sources and methods of selection of participants. Describe methods of follow-up *Case-control study* - Give the eligibility criteria, and the sources and methods of case ascertainment and control selection. Give the rationale for the choice of cases and controls *Cross-sectional study* - Give the eligibility criteria, and the sources and methods of selection of participants *(b) Cohort study* - For matched studies, give matching criteria and number of exposed and unexposed *Case-control study* - For matched studies, give matching	Methods	RECORD 6.1: The methods of study population selection (such as codes or algorithms used to identify subjects) should be listed in detail. If this is not possible, an explanation should be provided. RECORD 6.2: Any validation studies of the codes or algorithms used to select the population should be referenced. If validation was conducted for this study and not published elsewhere, detailed methods and results should be provided. RECORD 6.3: If the study involved linkage of databases, consider use of a flow diagram or other graphical display to demonstrate the data linkage process, including the number of individuals with linked data at each stage.	Methods

		criteria and the number of controls per case			
Variables	7	Clearly define all outcomes, exposures, predictors, potential confounders, and effect modifiers. Give diagnostic criteria, if applicable.	Methods, Supplementary Table S1	RECORD 7.1: A complete list of codes and algorithms used to classify exposures, outcomes, confounders, and effect modifiers should be provided. If these cannot be reported, an explanation should be provided.	Supplementary Table S1
Data sources/ measurement	8	For each variable of interest, give sources of data and details of methods of assessment (measurement). Describe comparability of assessment methods if there is more than one group	Methods, Supplementary Table S1		Methods, Supplementary Table S1
Bias	9	Describe any efforts to address potential sources of bias	N/A		N/A
Study size	10	Explain how the study size was arrived at	Methods		Methods
Quantitative variables	11	Explain how quantitative variables were handled in the analyses. If applicable, describe which groupings were chosen, and why	Methods, Supplementary Table S1		Methods, Supplementary Table S1
Statistical methods	12	(a) Describe all statistical methods, including those used to control for confounding (b) Describe any methods used to examine subgroups and interactions (c) Explain how missing data were addressed	Methods		Methods

		(d) *Cohort study* - If applicable, explain how loss to follow-up was addressed *Case-control study* - If applicable, explain how matching of cases and controls was addressed *Cross-sectional study* - If applicable, describe analytical methods taking account of sampling strategy (e) Describe any sensitivity analyses			
Data access and cleaning methods		..		RECORD 12.1: Authors should describe the extent to which the investigators had access to the database population used to create the study population. RECORD 12.2: Authors should provide information on the data cleaning methods used in the study.	Methods, Supplementary Table S1
Linkage		..		RECORD 12.3: State whether the study included person-level, institutional-level, or other data linkage across two or more databases. The methods of linkage and methods of linkage quality evaluation should be provided.	No linkage
Results					
Participants	13	(a) Report the numbers of individuals at each stage of the study (*e.g.*, numbers potentially	Supplementary Appendix 3	RECORD 13.1: Describe in detail the selection of the persons included in the study (*i.e.,* study population selection)	Methods

		eligible, examined for eligibility, confirmed eligible, included in the study, completing follow-up, and analysed) (b) Give reasons for non-participation at each stage. (c) Consider use of a flow diagram		including filtering based on data quality, data availability and linkage. The selection of included persons can be described in the text and/or by means of the study flow diagram.	
Descriptive data	14	(a) Give characteristics of study participants (*e.g.*, demographic, clinical, social) and information on exposures and potential confounders (b) Indicate the number of participants with missing data for each variable of interest (c) *Cohort study* - summarise follow-up time (*e.g.*, average and total amount)	Results, Supplementary Appendix 3		
Outcome data	15	*Cohort study* - Report numbers of outcome events or summary measures over time *Case-control study* - Report numbers in each exposure category, or summary measures of exposure *Cross-sectional study* - Report numbers of outcome events or summary measures	Results, Supplementary Appendix 3		
Main results	16	(a) Give unadjusted estimates and, if applicable, confounder-adjusted estimates and their precision (e.g., 95% confidence	Results, Supplementary Appendix 3		

		interval). Make clear which confounders were adjusted for and why they were included (b) Report category boundaries when continuous variables were categorized (c) If relevant, consider translating estimates of relative risk into absolute risk for a meaningful time period			
Other analyses	17	Report other analyses done—e.g., analyses of subgroups and interactions, and sensitivity analyses	Results, Supplementary Appendix 1,3		
Discussion					
Key results	18	Summarise key results with reference to study objectives	Discussion		Discussion
Limitations	19	Discuss limitations of the study, taking into account sources of potential bias or imprecision. Discuss both direction and magnitude of any potential bias	Discussion	RECORD 19.1: Discuss the implications of using data that were not created or collected to answer the specific research question(s). Include discussion of misclassification bias, unmeasured confounding, missing data, and changing eligibility over time, as they pertain to the study being reported.	Discussion
Interpretation	20	Give a cautious overall interpretation of results considering objectives, limitations, multiplicity of analyses, results from similar studies, and other relevant evidence	Discussion		Discussion

Generalisability	21	Discuss the generalisability (external validity) of the study results	N/A		N/A
Other Information					
Funding	22	Give the source of funding and the role of the funders for the present study and, if applicable, for the original study on which the present article is based	Declarations		Declarations
Accessibility of protocol, raw data, and programming code		..		RECORD 22.1: Authors should provide information on how to access any supplemental information such as the study protocol, raw data, or programming code.	Given the sensitive nature of the data, this information is not shared.

D.2 Extra Figures

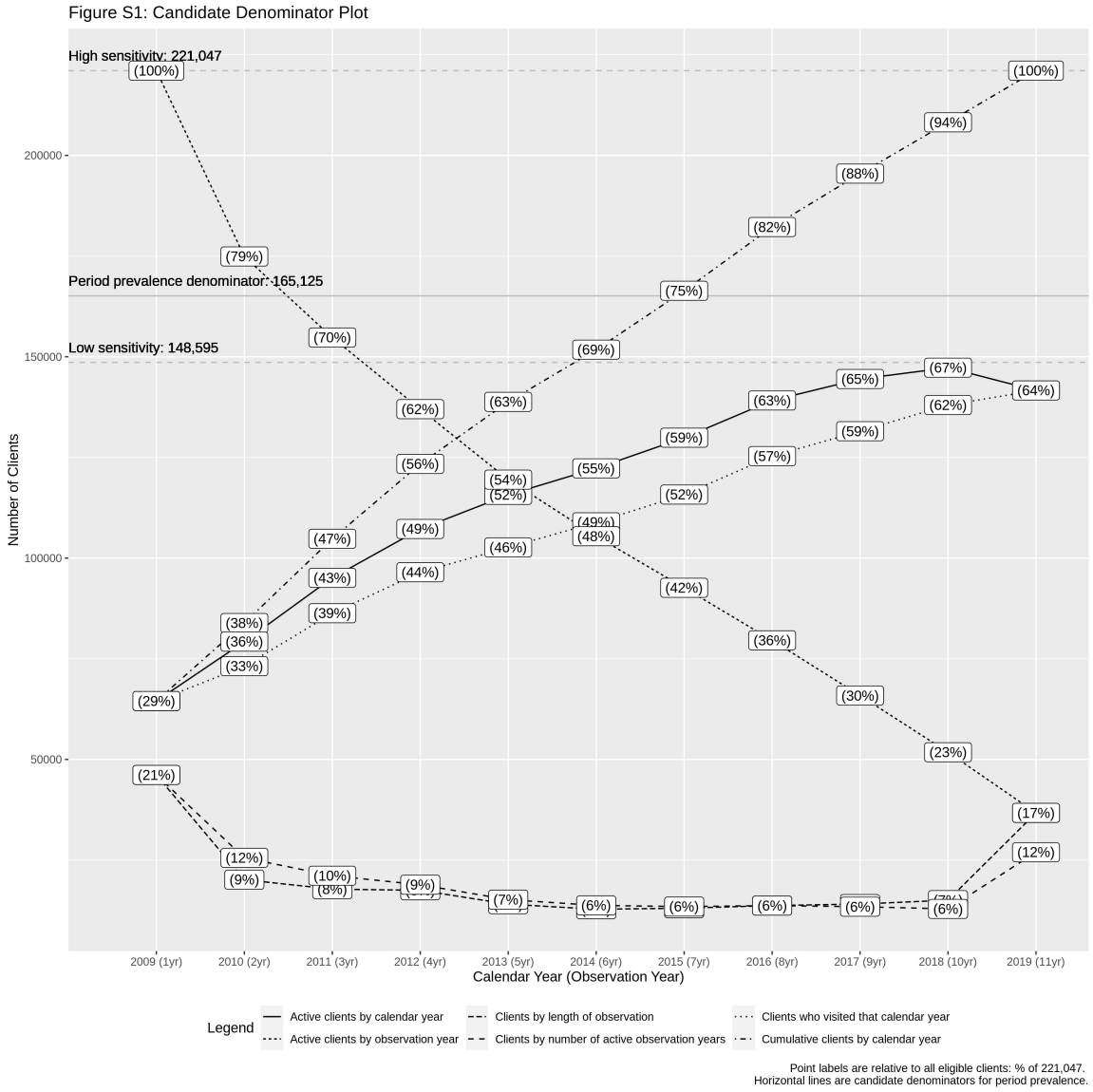
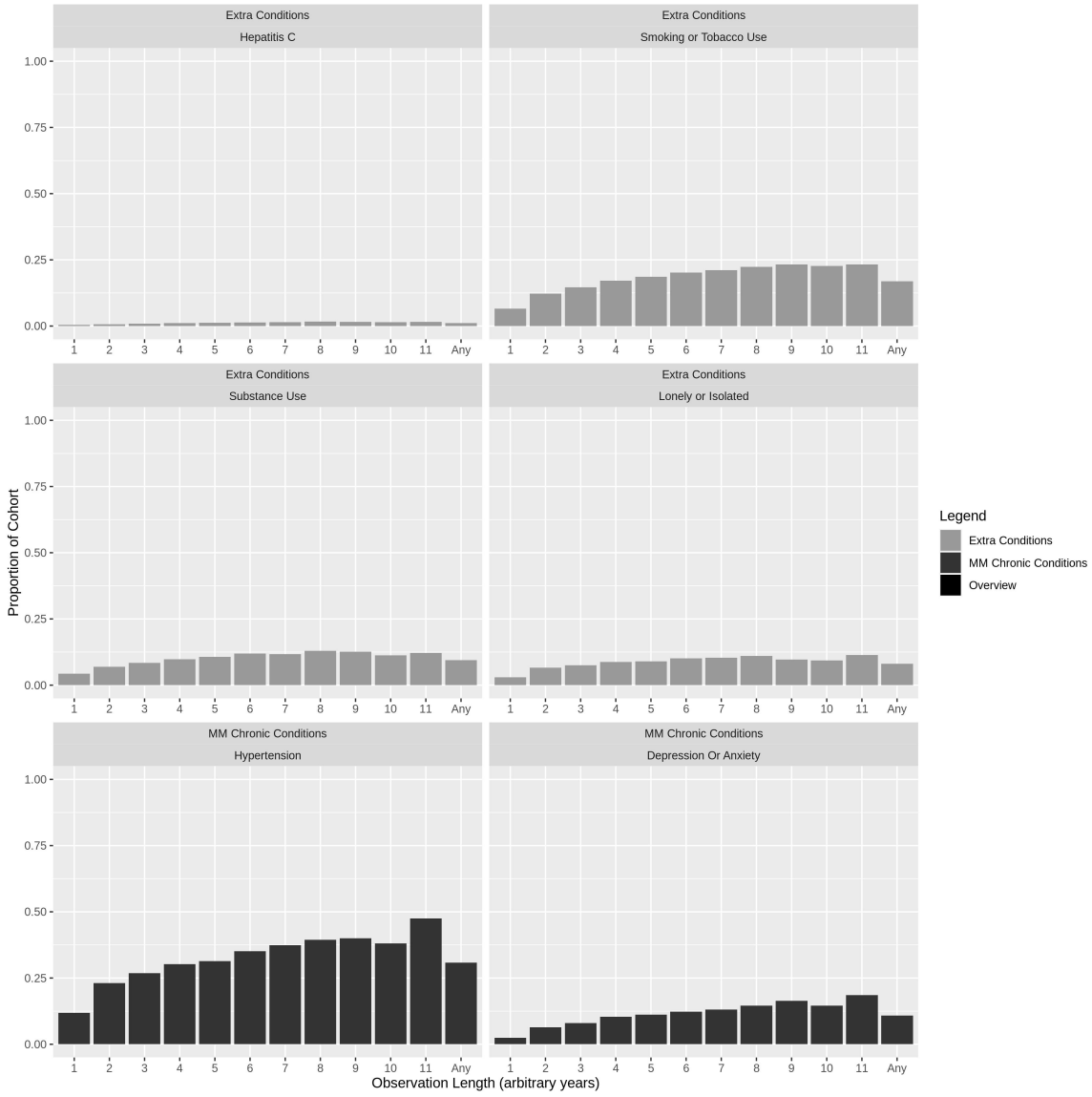
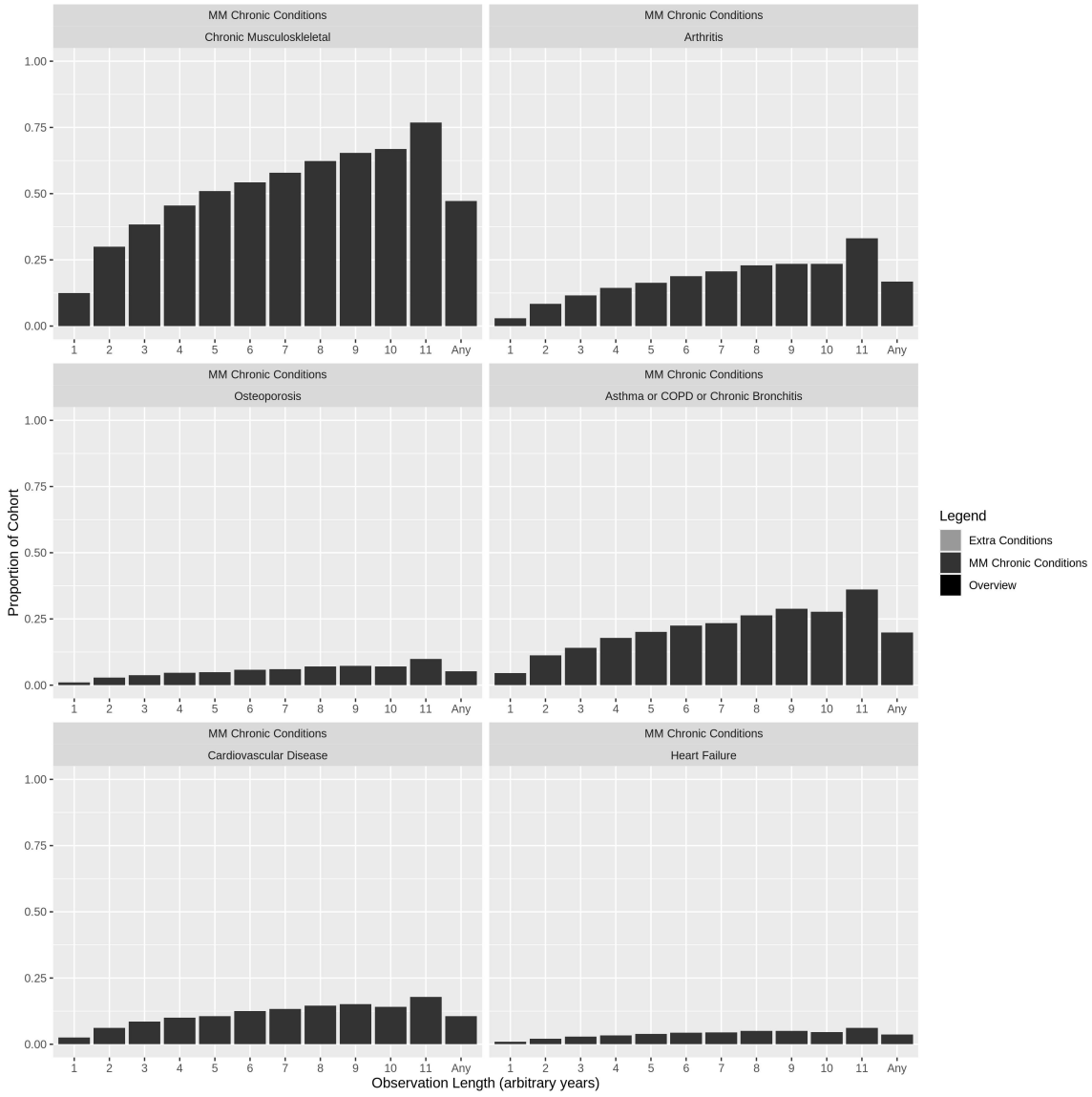
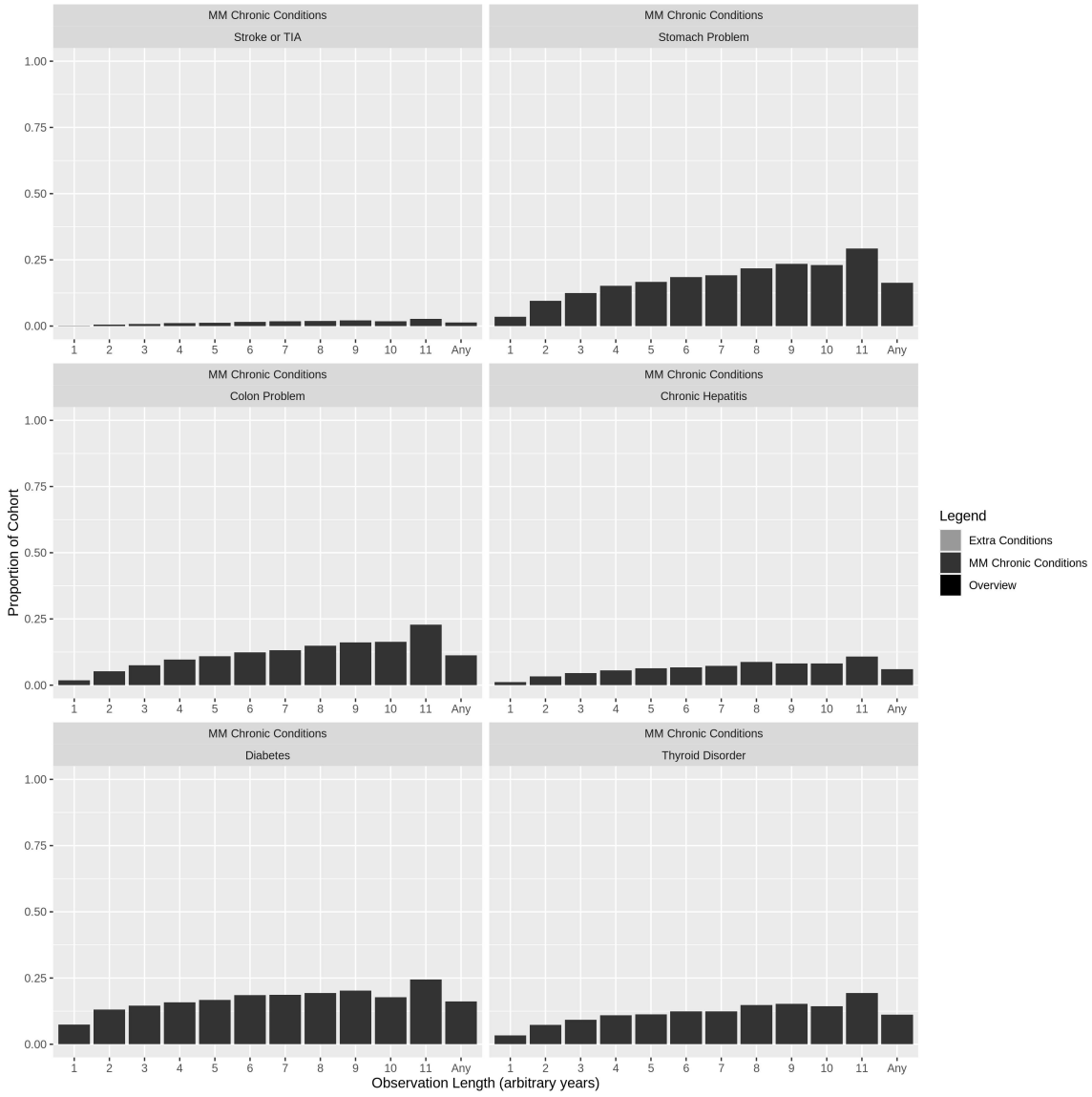


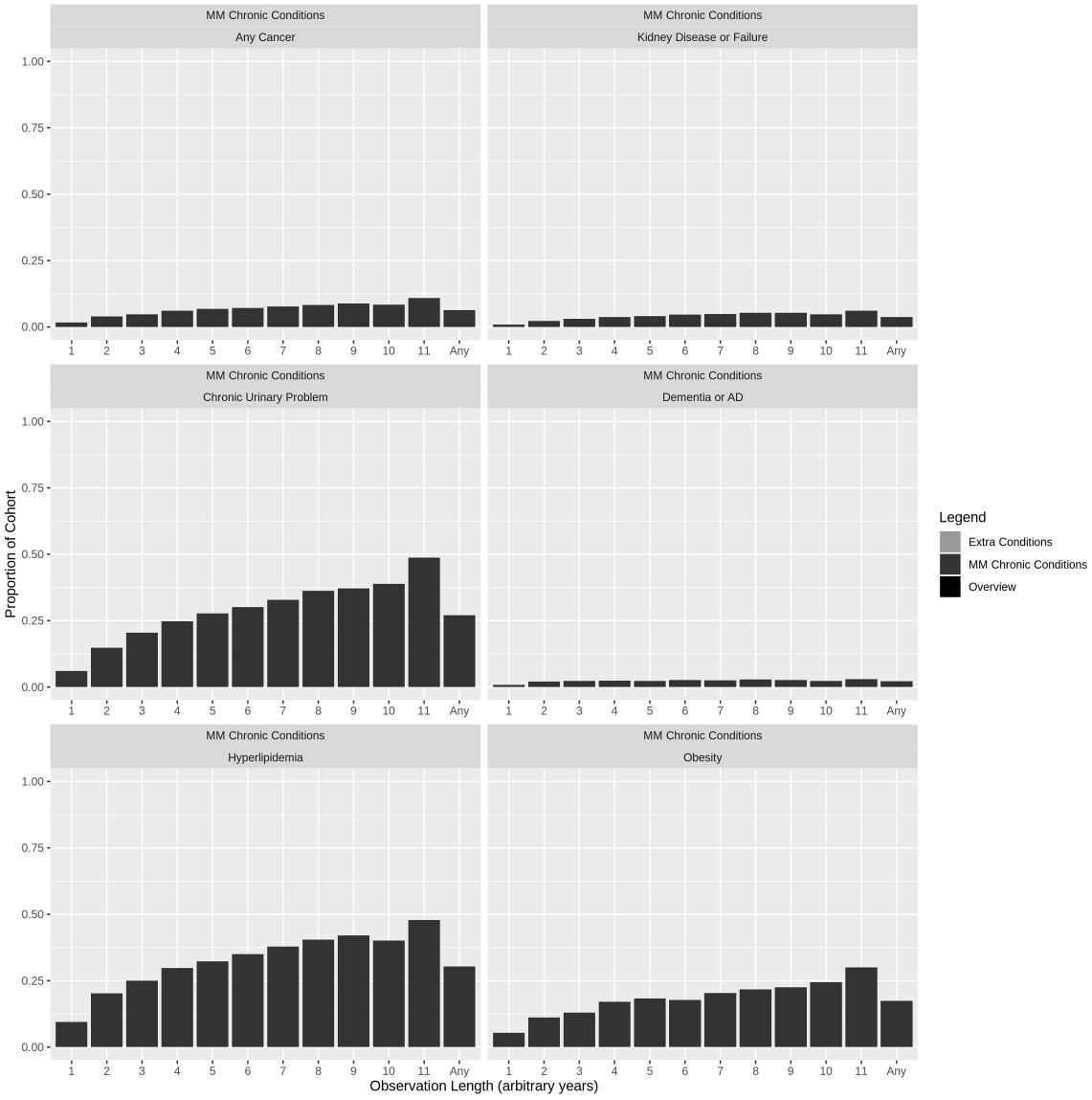
Figure D.1: Cohort size by calendar- and observation-based time.

Notes: Active clients had at least one event during or after the year (calendar- or observation-based) of interest (gap years counted). The number of active observation years refers to the number of 365.25 day periods, counted from the first calendar date that an event was recorded for that client, that clients had at least one event recorded (gap years not counted). Length of observation refers to the number of years from the first to the last year that at least one event was recorded during (gap years counted). Cumulative clients refers to the number of clients who had at least one event during or before the year of interest. *Legend*: COPD = Chronic Obstructive Pulmonary Disease; TIA = Transient Ischemic Attack; AD = Alzheimer's Disease.


Extra Conditions
Hepatitis C
Extra Conditions
Smoking or Tobacco Use
Extra Conditions
Substance Use
Extra Conditions
Lonely or Isolated
MM Chronic Conditions
Hypertension
MM Chronic Conditions
Depression Or Anxiety
Proportion of Cohort
1.00
0.75
0.50
0.25
0.00
1
2
3
4
5
6
7
8
9
10
11
Any
Observation Length (arbitrary years)
Legend
Extra Conditions
MM Chronic Conditions
Overview


MM Chronic Conditions
Chronic Musculoskleletal
MM Chronic Conditions
Arthritis
MM Chronic Conditions
Osteoporosis
MM Chronic Conditions
Asthma or COPD or Chronic Bronchitis
MM Chronic Conditions
Cardiovascular Disease
MM Chronic Conditions
Heart Failure
Proportion of Cohort
1.00
0.75
0.50
0.25
0.00
1 2 3 4 5 6 7 8 9 10 11 Any
Observation Length (arbitrary years)
Legend
Extra Conditions
MM Chronic Conditions
Overview


MM Chronic Conditions
Stroke or TIA
MM Chronic Conditions
Stomach Problem
MM Chronic Conditions
Colon Problem
MM Chronic Conditions
Chronic Hepatitis
MM Chronic Conditions
Diabetes
MM Chronic Conditions
Thyroid Disorder
Proportion of Cohort
1.00
0.75
0.50
0.25
0.00
1 2 3 4 5 6 7 8 9 10 11 Any
Observation Length (arbitrary years)
Legend
Extra Conditions
MM Chronic Conditions
Overview


MM Chronic Conditions
Any Cancer
MM Chronic Conditions
Kidney Disease or Failure
MM Chronic Conditions
Chronic Urinary Problem
MM Chronic Conditions
Dementia or AD
MM Chronic Conditions
Hyperlipidemia
MM Chronic Conditions
Obesity
Proportion of Cohort
1.00
0.75
0.50
0.25
0.00
1 2 3 4 5 6 7 8 9 10 11 Any
Observation Length (arbitrary years)
Legend
Extra Conditions
MM Chronic Conditions
Overview

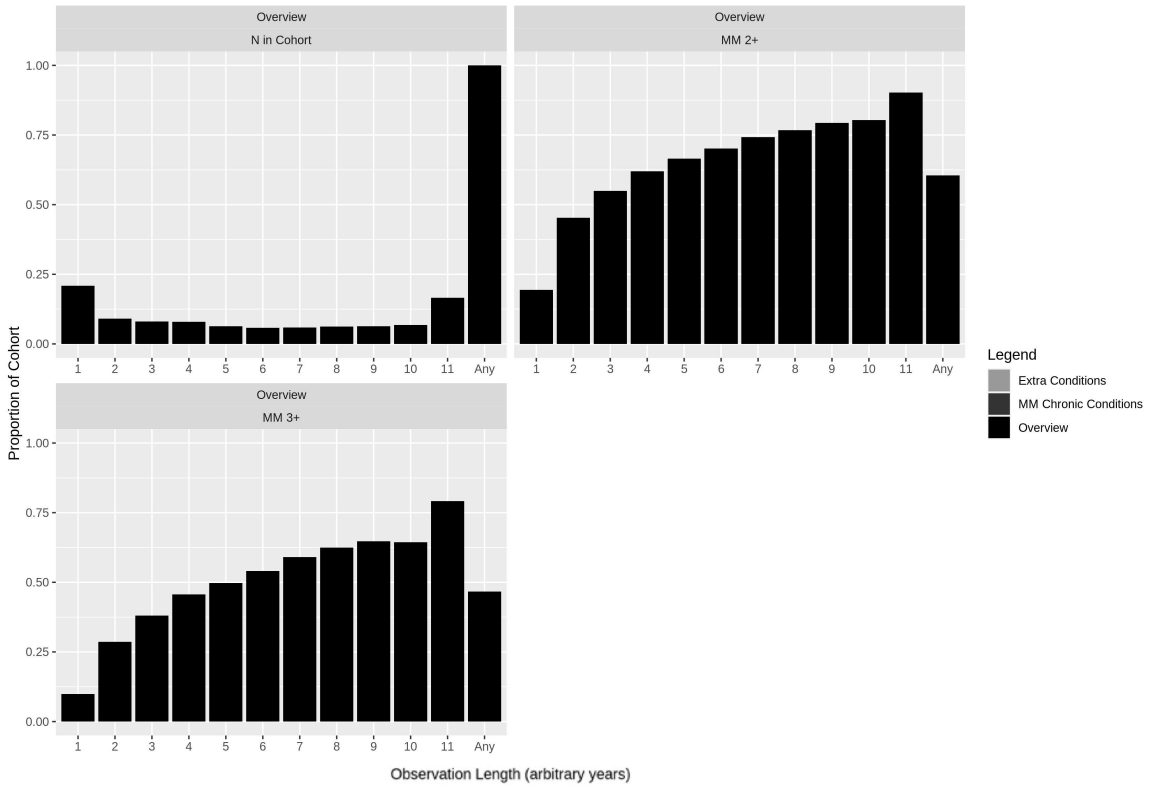
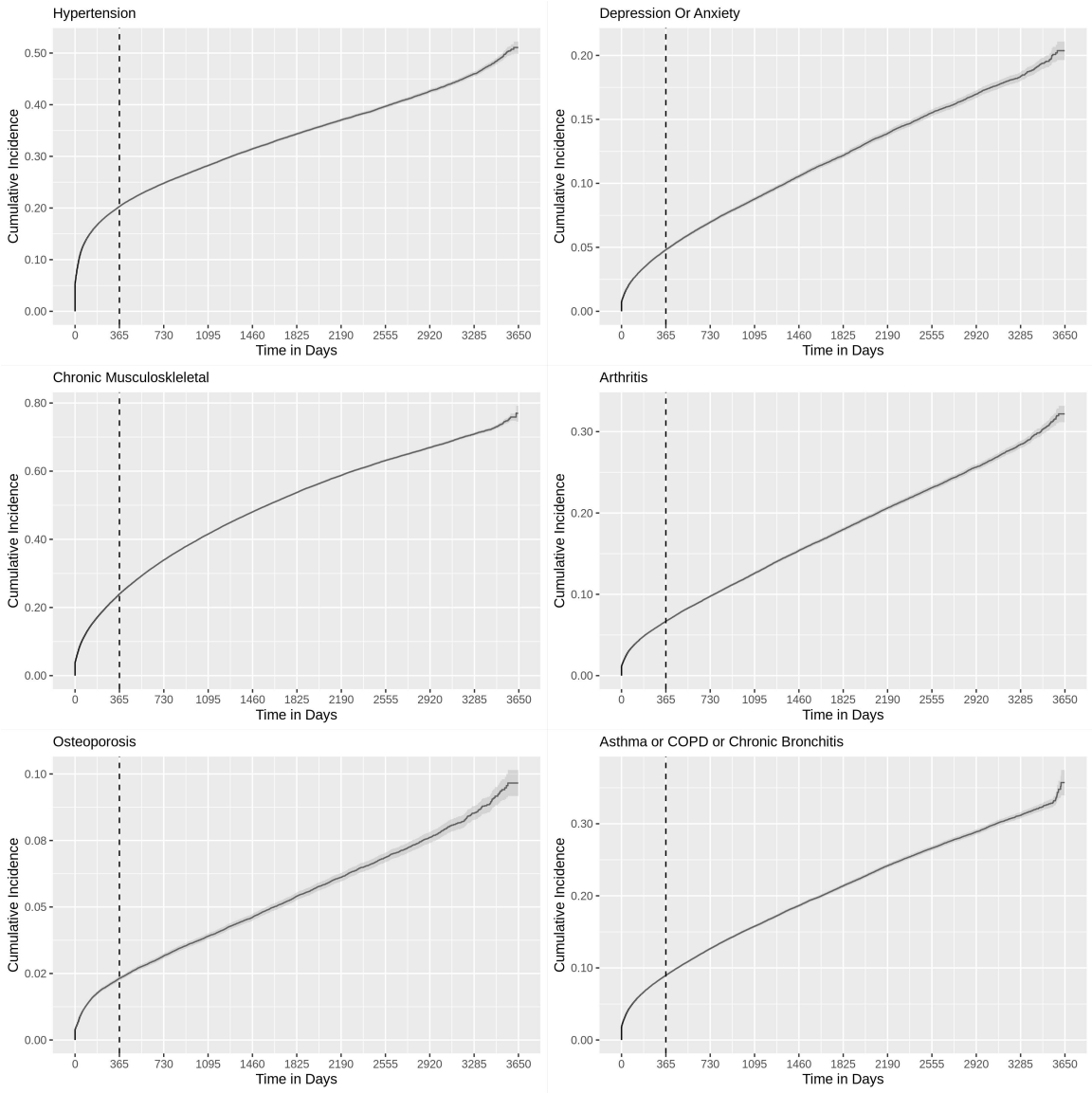
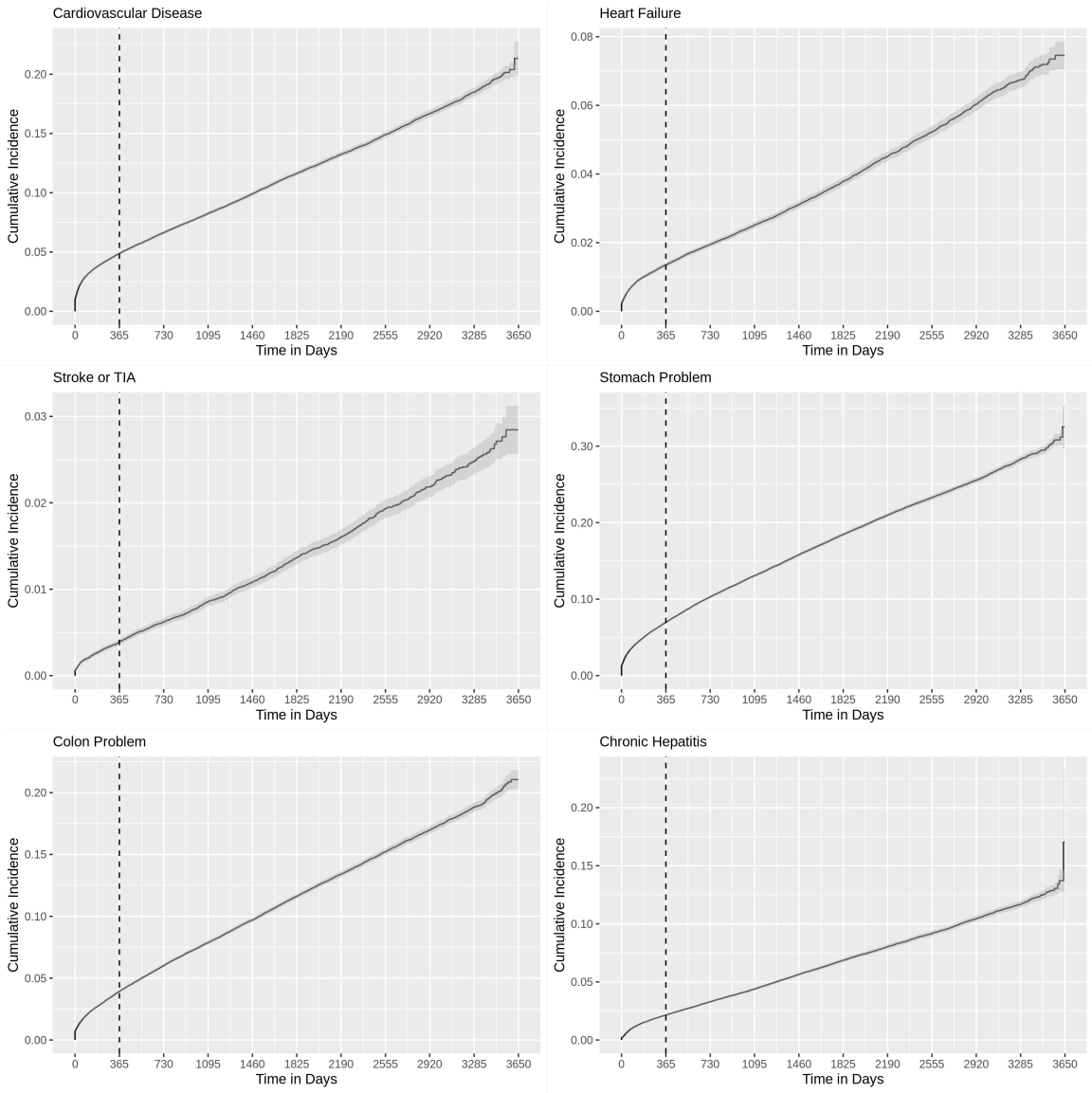
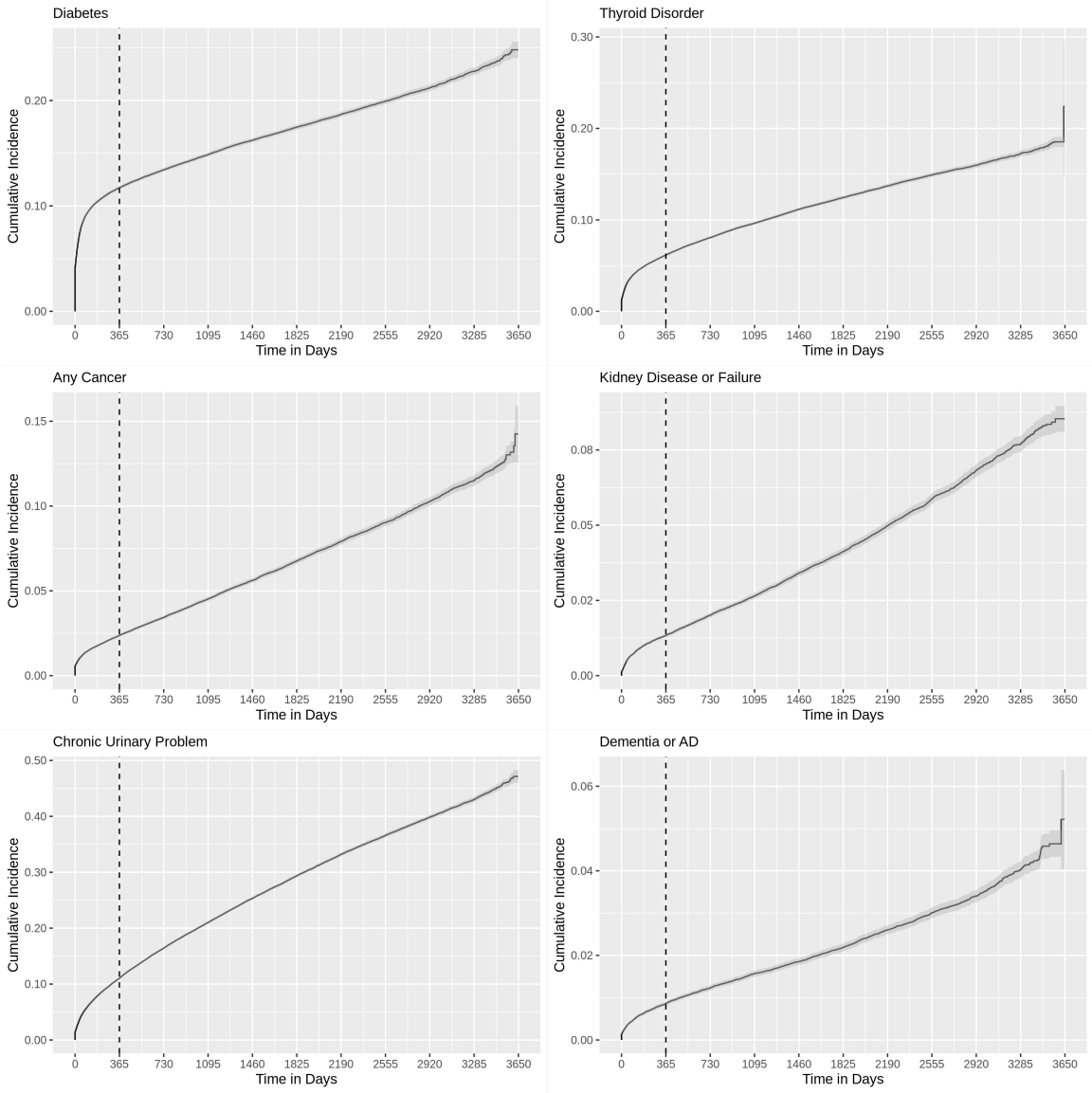


Figure D.2: Observation-based period prevalence.

Notes: Each bar represents the proportion of clients within that observation-based cohort (years are arbitrary 365.25 day consecutive periods between the first and last recorded events) that had at least one indication of the condition of interest across their entire observation history. Conditions were grouped to represent 1) Extra conditions of interest to Alliance stakeholders, 2) 20 chronic conditions, which make up multimorbidity (MM) status, and 3) Overview indicators for the cohorts. *Legend*: COPD = Chronic Obstructive Pulmonary Disease; TIA = Transient Ischemic Attack; AD = Alzheimer's Disease.


Hypertension
Cumulative Incidence
0.00
0.10
0.20
0.30
0.40
0.50
0 365 730 1095 1460 1825 2190 2555 2920 3285 3650
Time in Days
Depression Or Anxiety
Cumulative Incidence
0.00
0.05
0.10
0.15
0.20
0 365 730 1095 1460 1825 2190 2555 2920 3285 3650
Time in Days
Chronic Musculoskleletal
Cumulative Incidence
0.00
0.20
0.40
0.60
0.80
0 365 730 1095 1460 1825 2190 2555 2920 3285 3650
Time in Days
Arthritis
Cumulative Incidence
0.00
0.10
0.20
0.30
0 365 730 1095 1460 1825 2190 2555 2920 3285 3650
Time in Days
Osteoporosis
Cumulative Incidence
0.00
0.02
0.05
0.08
0.10
0 365 730 1095 1460 1825 2190 2555 2920 3285 3650
Time in Days
Asthma or COPD or Chronic Bronchitis
Cumulative Incidence
0.00
0.10
0.20
0.30
0 365 730 1095 1460 1825 2190 2555 2920 3285 3650
Time in Days


Cardiovascular Disease
Cumulative Incidence
0.00
0.05
0.10
0.15
0.20
0
365
730
1095
1460
1825
2190
2555
2920
3285
3650
Time in Days
Heart Failure
Cumulative Incidence
0.00
0.02
0.04
0.06
0.08
Time in Days
Stroke or TIA
Cumulative Incidence
0.00
0.01
0.02
0.03
Time in Days
Stomach Problem
Cumulative Incidence
0.00
0.10
0.20
0.30
Time in Days
Colon Problem
Cumulative Incidence
0.00
0.05
0.10
0.15
0.20
Time in Days
Chronic Hepatitis
Cumulative Incidence
0.00
0.05
0.10
0.15
0.20
Time in Days


Diabetes
Cumulative Incidence
0.00
0.10
0.20
0 365 730 1095 1460 1825 2190 2555 2920 3285 3650
Time in Days
Thyroid Disorder
Cumulative Incidence
0.00
0.10
0.20
0.30
0 365 730 1095 1460 1825 2190 2555 2920 3285 3650
Time in Days
Any Cancer
Cumulative Incidence
0.00
0.05
0.10
0.15
0 365 730 1095 1460 1825 2190 2555 2920 3285 3650
Time in Days
Kidney Disease or Failure
Cumulative Incidence
0.00
0.02
0.05
0.08
0 365 730 1095 1460 1825 2190 2555 2920 3285 3650
Time in Days
Chronic Urinary Problem
Cumulative Incidence
0.00
0.10
0.20
0.30
0.40
0.50
0 365 730 1095 1460 1825 2190 2555 2920 3285 3650
Time in Days
Dementia or AD
Cumulative Incidence
0.00
0.02
0.04
0.06
0 365 730 1095 1460 1825 2190 2555 2920 3285 3650
Time in Days

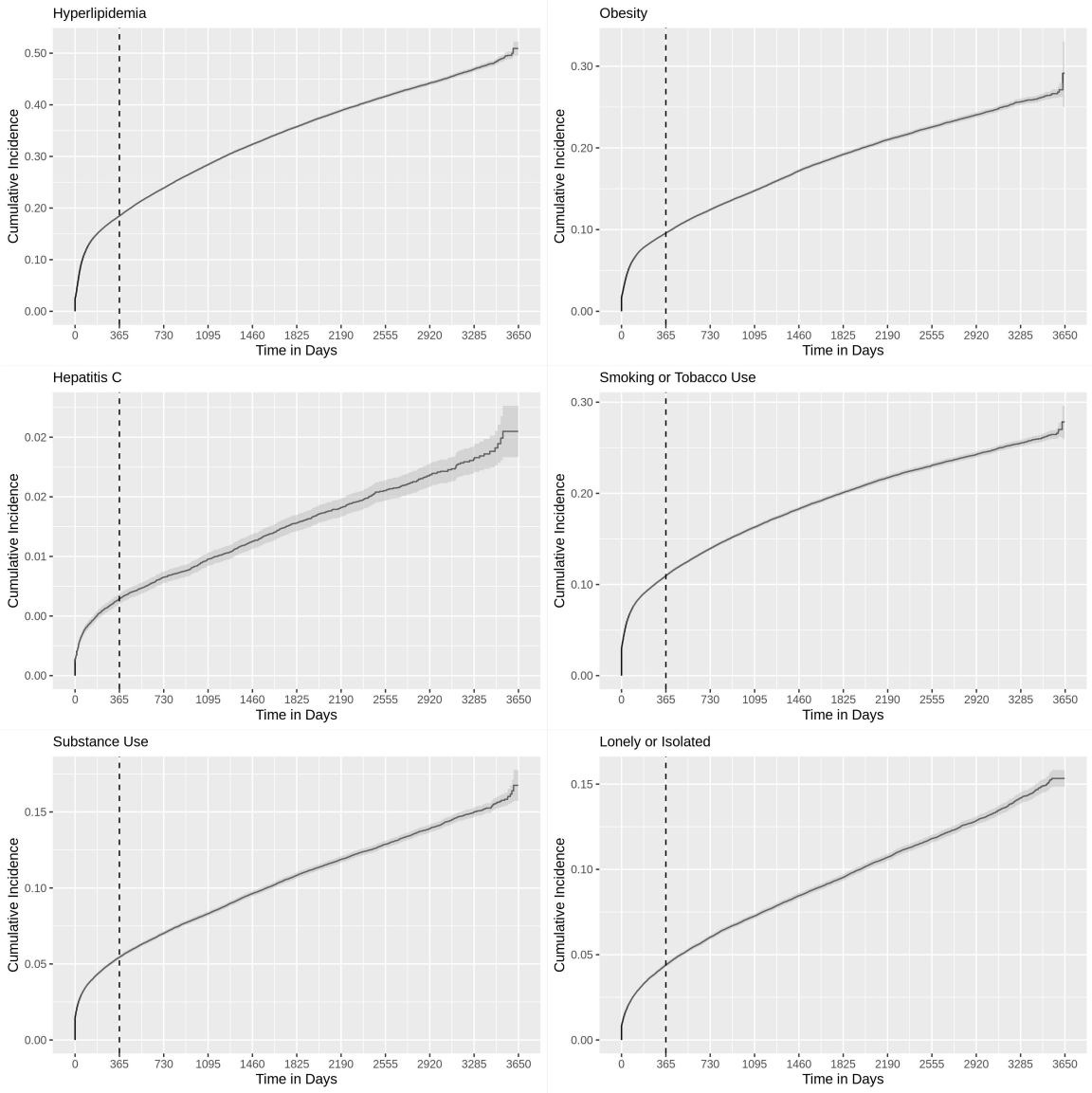


Figure D.3: Cumulative incidence plots by days of observation since the first recorded event.

Notes: Clients eligible for this analysis must not have had any care recorded in the first calendar-year of available data (2009). *Legend*: COPD = Chronic Obstructive Pulmonary Disease; TIA = Transient Ischemic Attack; AD = Alzheimer's Disease.

D.3 Extra Tables

Table D.1: Characteristic variable definitions.

Variable Name	**Definition**	**Source used to guide variable operationalization, if any**
Sociodemographic		
Age in 2015	2015 minus Year of Birth	
Geography	Geography of place of residence based on Forward Sortation Area: Rural if second digit is 0; Urban if any other valid digit; NA otherwise.	[1]
Sex	Categories as recorded in client characteristic table	
Gender	Collapsed client characteristic table categories	[2]
Sexual Orientation	Categories as recorded in client characteristic table	
Highest Level of Education Completed	Collapsed client characteristic table categories	[2] followed to the extent possible
Primary Spoken Language	Collapsed client characteristic table categories into the two official languages of Canada with remaining languages categorized as Other.	
Race and Ethnicity	Collapsed client characteristic table categories	[3,4]
Year Since Arrival in Canada	Cleaned free text entries from client characteristic table and collapsed into 5 years, 6 or more year, and None Recorded categories. None Recorded cannot differentiate between never-immigrated and never-asked.	
Household Income	Collapsed client characteristic table categories	Note: could not reliably follow guidelines in [2]
Household Composition	Categories as recorded in client characteristic table	
Stable Residence	Collapsed client characteristic table categories into Stable or Unstable (homeless, shelter, other temporary). Additional unstable residence situations were identified as the presence of at least one ENCODE-FM code: 8990, 9433, 9434, 9435, 9436, 9437, 9438, 9439, 9440, 9441, 9442, 9443, 9432, 8982, 8986, 9419, 9424, 8985, 9431, 9415, 9425, 9412, 9414, which were given priority.	List of codes are from an Alliance for Healthier Communities stakeholder.
Food Insecurity	At least one ENCODE-FM code: 8972, 9782, 9802, 8971, 9568, 9805	List of codes are from an Alliance for Healthier Communities stakeholder
Clinical		
Hypertension	At least one ICD-10 code: i10, i11, i12, i13, i14, i15	[5]
Depression or Anxiety	At least one ICD-10 code: f33, f40, f41	[5]

Chronic Musculoskeletal Conditions causing pain or limitation	At least one ICD-10 code: m40, m41, m42, m43, m44, m45, m46, m47, m48, m49, m50, m51, m52, m53, m54, m60, m61, m62, m63, m65, m66, m67, m68, m70, m71, m72, m73, m74, m75, m76, m77, m78, m79	[5]	
Arthritis and/or Rheumatoid Arthritis	At least one ICD-10 code: m05.9, m13.0, m13.9, m15, m16, m17, m18, m19	[5]	
Osteoporosis	At least one ICD-10 code: m81	[5]	
Asthma, Chronic Obstructive Pulmonary Disease, or Chronic Bronchitis	At least one ICD-10 code: j40, j41, j42, j43, j44, j45, j46	[5]	
Cardiovascular Disease (angina, myocardial infarction, atrial fibrillation, poor circulation in the lower limbs)	At least one ICD-10 code: i20, i25, i48, i70, i71, i72, i73, i74, i75, i76, i77, i78, i79	[5]	
Heart Failure (including valve problems or replacement)	At least one ICD-10 code: i05, i06, i07, i08, i09, i34, i35, i36, i37, i38, i39, i42, i43, i50	[5]	
Stroke and Transient Ischemic Attack	At least one ICD-10 code: g45, i62	[5]	
Stomach Problem (irritable bowel, Chron's disease, ulcerative colitis, diverticulosis)	At least one ICD-10 code: k21, k25.7, k29.5	[5]	
Colon Problem	At least one ICD-10 code: k50, k51, k52, k57, k58	[5]	
Chronic Hepatitis	At least one ICD-10 code: k70, k71, k72, k73, k74, k75, k76, k77	[5]	
Diabetes	At least one ICD-10 code: e10, e11, e12, e13, e14	[5]	
Thyroid Disorder	At least one ICD-10 code: e00, e01, e02, e03, e04, e05, e06, e07	[5]	
Any Cancer (including melanoma, but excluding other skin cancers)	At least one ICD-10 code: c00, c01, c02, c03, c04, c05, c06, c07, c08, c09, c10, c11, c12, c13, c14, c15, c16, c17, c18, c19, c20, c21, c22, c23, c24, c25, c26, c27, c28, c29, c30, c31, c32, c33, c34, c35, c36, c37, c38, c39, c40, c41, c42, c43, c44, c45, c46, c47, c48, c49, c50, c51, c52, c53, c54, c55, c56, c57, c58, c59, c60, c61, c62, c63, c64, c65, c66, c67, c68, c69, c70, c71, c72, c73, c74, c75, c76, c77, c78, c79, c80, c81, c82, c83, c84, c85, c86, c87, c88, c89, c90, c91, c92, c93, c94, c95, c96, c97	[5] modified by removing the 5 year restriction; taking any cancer indication within the 10 year period	
Kidney Disease or Failure	At least one ICD-10 code: n18, n19	[5]	
Chronic Urinary Problem	At least one ICD-10 code: n03, n11, n18, n20, n21, n22, n23, n25, n26, 27, n28, n29, n30, n31, n32, n33, n34, n35, n36, n37, n38, n39, n40, n41, n42, n43, n44, n45, n46, n47, n48, n49, n50, n51+B38	[5]	

Dementia or Alzheimer's Disease	At least one ICD-10 code: f00, f01, f02, f03	[5]
Hyperlipidemia (high cholesterol)	At least one ICD-10 code: e78	[5]
Obesity	At least one ICD-10 code: e66	[5] ICD-10 only; no BMI
Hepatitis C	At least one ICD-10 code: b18.2, b19.20, b19.21	[6]
Smoking or Tobacco Use	At least one ENCODE-FM code: 10072, 5520, 679, 9910, 5339, 5340, 5341, 5342, 5343, 5344, 5345, 5346, 5347, 5348, 5349	JKK selected relevant ENCODE-FM codes based on manual review
Substance Use	At least one ENCODE-FM code: 5304, 10004, 10005, 5305, 5306, 5307, 9754, 5308, 5309, 5310, 5311, 5312, 5313, 5314, 5315, 5316, 5317, 5318, 5319, 5320, 5321, 5322, 5323, 5324, 5325, 5326, 5327, 5328, 5329, 5330, 5331, 5332, 5333, 5334, 5335, 5336, 5337, 5338, 5350, 5351, 5352, 5353, 5354, 5355, 5356, 5357, 5358, 5359, 5360, 5361, 5362, 5363, 5364, 5365, 5366, 5367, 5368, 5369, 5370, 5371, 10007, 5372, 5373, 5374, 5375, 5376, 5377, 5378, 5379, 5380, 5381, 5382, 5383, 5384, 5385, 5386, 5387, 5388, 5389, 5390, 5391, 5392, 5393, 5394, 5395, 5396, 5397, 5398, 5399, 5400, 9845, 5401, 9844, 5401, 5402, 5403, 5404, 5405, 5406, 5407, 5408, 5409, 5410, 5411, 5412, 5413, 5414, 5415, 5416, 5417, 5418, 5419, 5420, 5421, 5422, 5423, 5424, 5425, 5426, 5427, 5428, 5429, 5430, 5431, 5432, 5433, 5434, 5435, 5436, 5437, 5438, 5439, 5440, 5441, 5442, 5443, 5444, 5445, 5446, 5447, 5448, 5449, 9277, 9278, 5450, 5451, 5452, 5453, 5454, 5455, 5456, 5457, 5458, 5459, 5460, 5461, 5462, 5463, 5464, 5465, 5466, 5467, 5468, 5469, 5470, 5471, 5472, 5473, 5474, 5475, 5476, 5477, 5478, 5479, 5480, 5481 or recorded in Disabilities Table	JKK selected relevant ENCODE-FM codes based on manual review
Lonely or Isolated	At least one ENCODE-FM code: 5138, 5139, 9265, 9267, 9268, 9512	List of codes are from an Alliance for Healthier Communities stakeholder
Health Care Use		
# Years of Observation	Based on records in the service event table: Ceiling of number of days from first to last recorded event divided by 365.25	
# Provider Types Seen	Number of unique provider types recorded in providers involved table. Provider types were maintained as entered except Other, Unknown, and Undefined were collapsed	
# Internal Referrals	Number of records in the internal referrals table	
# External Referrals	Number of records in the external referrals table	

Avg. # Days per Year	Based on records in the service event table: Sum of unique calendar days with at least one event recorded divided by Number of Years of Observation
Max # Days per Year	Based on records in the service event table: Maximum of number of unique calendar days care is accessed in a single calendar year
Avg. # Events per Day	Based on records in the service event table: Sum of events divided by number of calendar days care is accessed at least once
Max # Events per Day	Based on records in the service event table: Maximum number of events recorded in a single calendar day

Legend: # = Number; Avg. = Average; CHC = Community Health Centre; ENCODE-FM = Electronic Nomenclature and Classification Of Disorders and Encounters for Family Medicine; ICD = International Classification of Disease; SD = Standard Deviation; UAR = Urban at Risk. *References:* [1] Canada Post. Addressing guidelines - Forward Sortation Area (FSA). Canada Post. 2022 [cited 2022 Feb 8]. Available from: https://www.canadapost-postescanada.ca/cpc/en/support/articles/addressing-guidelines/postal-codes.page [2] CIHI. In Pursuit of Health Equity: Defining Stratifiers for Measuring Health Inequality - A Focus on Age, Sex, Gender, Income, Education and Geographic Location. Ottawa, ON: Canadian Institute for Health Information; 2018 Apr. Available from: https://www.cihi.ca/sites/default/files/document/defining-stratifiers-measuring-health-inequalities-2018-en-web.pdf [3] CIHI. Proposed Standards for Race-Based and Indigenous Identity Data Collection and Health Reporting in Canada. Ottawa, ON: Canadian Institute for Health Information; 2020. Available from: https://www.cihi.ca/en/proposed-standards-for-race-based-and-indigenous-identity-data [4] Flanagin A, Frey T, Christiansen SL, AMA Manual of Style Committee. Updated guidance on the reporting of race and ethnicity in medical and science journals. *JAMA*. 2021;326(7):621–7. [5] Fortin M, Almirall J, Nicholson K. Development of a research tool to document self-reported chronic conditions in primary care. *J Comorb.* 2017;7(1):117–23. [6] Support Path. Hepatitis C ICD-10 Codes. Gilead Sciences; 2015 [cited 2020 Sep 25]. Available from: https://www.cvph.org/data/files/mysupportpath.pdf

Table D.2: Sociodemographic characteristics sub-strata.

Characteristic	Values	All n (%)	UAR & MM n (%)	UAR & Non-MM n (%)	Non-UAR & MM n(%)	Non-UAR & Non-MM n(%)
n		221047	19237	16761	83935	101114
Age in 2015	25-34	55505 (25.11)	1864 (9.69)	6112 (36.47)	7482 (8.91)	40047 (39.61)
	35-44	45646 (20.65)	3154 (16.40)	4386 (26.17)	12388 (14.76)	25718 (25.43)
	45-54	44653 (20.20)	4784 (24.87)	3402 (20.30)	19198 (22.87)	17269 (17.08)
	55-64	37848 (17.12)	4935 (25.65)	1855 (11.07)	20643 (24.59)	10415 (10.3)
	65-74	23162 (10.48)	2952 (15.35)	692 (4.13)	14828 (17.67)	4690 (4.64)
	75+	14233 (6.44)	1548 (8.05)	314 (1.87)	9396 (11.19)	2975 (2.94)
Rural Geography	Rural	49275 (22.29)	3479 (18.08)	2652 (15.82)	23339 (27.81)	19805 (19.59)
	Urban	167728 (75.88)	15291 (79.49)	13247 (79.03)	59720 (71.15)	79470 (78.59)
	Missing	4044 (1.83)	467 (2.43)	862 (5.14)	876 (1.04)	1839 (1.82)
Sex	Female	127070 (57.49)	10647 (55.35)	8052 (48.04)	49299 (58.73)	59072 (58.42)
	Male	93294 (42.21)	8561 (44.5)	8590 (51.25)	34563 (41.18)	41580 (41.12)
	Other	331 (0.15)	3 (0.02)	40 (0.24)	16 (0.02)	272 (0.27)
	Missing	352 (0.16)	26 (0.14)	79 (0.47)	57 (0.07)	190 (0.19)
Gender	Female	41352 (18.71)	3352 (17.42)	2157 (12.87)	18479 (22.02)	17364 (17.17)
	Gender Diverse	340 (0.15)	52 (0.27)	60 (0.36)	92 (0.11)	136 (0.13)
	Male	29366 (13.28)	2425 (12.61)	2160 (12.89)	12308 (14.66)	12473 (12.34)
	Prefer not to answer	1001 (0.45)	37 (0.19)	14 (0.08)	339 (0.4)	611 (0.60)
	Missing	148988 (67.40)	13371 (69.51)	12370 (73.8)	52717 (62.81)	70530 (69.75)
Sexual Orientation	Bisexual	1578 (0.71)	141 (0.73)	144 (0.86)	549 (0.65)	744 (0.74)
	Gay	708 (0.32)	94 (0.49)	98 (0.58)	212 (0.25)	304 (0.30)
	Heterosexual	57065 (25.82)	4703 (24.45)	3744 (22.34)	24402 (29.07)	24216 (23.95)
	Lesbian	485 (0.22)	45 (0.23)	25 (0.15)	199 (0.24)	216 (0.21)
	Queer	323 (0.15)	14 (0.07)	20 (0.12)	77 (0.09)	212 (0.21)
	Two-Spirit	128 (0.06)	40 (0.21)	40 (0.24)	21 (0.03)	27 (0.03)
	Other	246 (0.11)	21 (0.11)	13 (0.08)	122 (0.15)	90 (0.09)
	Do not know	924 (0.42)	113 (0.59)	88 (0.53)	372 (0.44)	351 (0.35)
	Prefer not to answer	7561 (3.42)	565 (2.94)	312 (1.86)	3513 (4.19)	3171 (3.14)

	Missing	152029 (68.78)	13501 (70.18)	12277 (73.25)	54468 (64.89)	71783 (70.99)
Highest Level of Education	Post-secondary or equivalent	84888 (38.40)	6463 (33.60)	5593 (33.37)	29300 (34.91)	43532 (43.05)
	Secondary or equivalent	61831 (27.97)	6656 (34.60)	5127 (30.59)	25961 (30.93)	24087 (23.82)
	Less than high school	18941 (8.57)	1886 (9.80)	1380 (8.23)	8732 (10.40)	6943 (6.87)
	Other	8507 (3.85)	384 (2.00)	335 (2.00)	3694 (4.4)	4094 (4.05)
	Do not know	4860 (2.20)	734 (3.82)	584 (3.48)	1616 (1.93)	1926 (1.90)
	Prefer not to answer	2950 (1.33)	273 (1.42)	149 (0.89)	1312 (1.56)	1216 (1.20)
	Missing	39070 (17.67)	2841 (14.77)	3593 (21.44)	13320 (15.87)	19316 (19.10)
Primary Language	English	167163 (75.62)	17036 (88.56)	14622 (87.24)	62563 (74.54)	72942 (72.14)
	French	22547 (10.20)	554 (2.88)	390 (2.33)	10537 (12.55)	11066 (10.94)
	Other	26847 (12.15)	1473 (7.66)	1475 (8.80)	9237 (11.00)	14662 (14.50)
	Missing	4490 (2.03)	174 (0.90)	274 (1.63)	1598 (1.90)	2444 (2.42)
Race and Ethnicity	Black	8861 (4.01)	337 (1.75)	388 (2.31)	3420 (4.07)	4716 (4.66)
	East/SouthEast Asian	3739 (1.69)	248 (1.29)	236 (1.41)	1297 (1.55)	1958 (1.94)
	Indigenous	2944 (1.33)	838 (4.36)	739 (4.41)	803 (0.96)	564 (0.56)
	Latino	4350 (1.97)	102 (0.53)	104 (0.62)	1606 (1.91)	2538 (2.51)
	Middle Eastern	2046 (0.93)	149 (0.77)	195 (1.16)	689 (0.82)	1013 (1.00)
	Other	567 (0.26)	79 (0.41)	69 (0.41)	227 (0.27)	192 (0.19)
	South Asian	3597 (1.63)	232 (1.21)	91 (0.54)	1620 (1.93)	1654 (1.64)
	White	38464 (17.40)	2661 (13.83)	1870 (11.16)	18843 (22.45)	15090 (14.92)
	Do not know	838 (0.38)	91 (0.47)	60 (0.36)	396 (0.47)	291 (0.29)
	Prefer not to answer	2649 (1.20)	165 (0.86)	96 (0.57)	1348 (1.61)	1040 (1.03)
	Missing	152992 (69.21)	14335 (74.52)	12913 (77.04)	53686 (63.96)	72058 (71.26)
Years since Arrival in Canada	0-5 years	13654 (6.18)	315 (1.64)	876 (5.23)	2732 (3.25)	9731 (9.62)
	6+ years	51815 (23.44)	2863 (14.88)	2077 (12.39)	19859 (23.66)	27016 (26.72)
	None recorded	155578 (70.38)	16059 (83.48)	13808 (82.38)	61344 (73.09)	64367 (63.66)
Household Income	$0 to $14,999	40519 (18.33)	4476 (23.27)	4253 (25.37)	13281 (15.82)	18509 (18.31)
	$15,000 to $24,999	21102 (9.55)	2095 (10.89)	1460 (8.71)	8986 (10.71)	8561 (8.47)
	$25,000 to $39,999	20877 (9.44)	1772 (9.21)	1216 (7.25)	8964 (10.68)	8925 (8.83)

	$40,000 to $59,999	17245 (7.80)	1455 (7.56)	966 (5.76)	7216 (8.60)	7608 (7.52)
	$60,000 or more	28494 (12.89)	2092 (10.87)	1770 (10.56)	10776 (12.84)	13856 (13.70)
	Do not know	15408 (6.97)	1301 (6.76)	1357 (8.10)	4963 (5.91)	7787 (7.70)
	Prefer not to answer	27621 (12.50)	2437 (12.67)	1693 (10.10)	12453 (14.84)	11038 (10.92)
	Missing	49781 (22.52)	3609 (18.76)	4046 (24.14)	17296 (20.61)	24830 (24.56)
Household Composition	Couple with children	53398 (24.16)	3280 (17.05)	3479 (20.76)	17433 (20.77)	29206 (28.88)
	Couple without child	39664 (17.94)	3907 (20.31)	2038 (12.16)	19043 (22.69)	14676 (14.51)
	Extended Family	7632 (3.45)	578 (3.00)	545 (3.25)	3003 (3.58)	3506 (3.47)
	Grandparents with Grandchild(ren)	1746 (0.79)	187 (0.97)	60 (0.36)	996 (1.19)	503 (0.50)
	Siblings	1622 (0.73)	140 (0.73)	110 (0.66)	529 (0.63)	843 (0.83)
	Single Parent	14445 (6.53)	1344 (6.99)	1183 (7.06)	5004 (5.96)	6914 (6.84)
	Sole Member	32782 (14.83)	4503 (23.41)	2942 (17.55)	14094 (16.79)	11243 (11.12)
	Unrelated housemates	8622 (3.90)	669 (3.48)	898 (5.36)	2180 (2.60)	4875 (4.82)
	Other	8913 (4.03)	788 (4.10)	688 (4.10)	3414 (4.07)	4023 (3.98)
	Do not know	2475 (1.12)	301 (1.56)	342 (2.04)	978 (1.17)	854 (0.84)
	Prefer not to answer	3727 (1.69)	262 (1.36)	229 (1.37)	1665 (1.98)	1571 (1.55)
	Missing	46021 (20.82)	3278 (17.04)	4247 (25.34)	15596 (18.58)	22900 (22.65)
Stable Residence	True	199349 (90.18)	14813 (77.00)	13414 (80.03)	75666 (90.15)	95456 (94.4)
Food Insecurity	True	10985 (4.97)	2066 (10.74)	881 (5.26)	5257 (6.26)	2781 (2.75)

Legend: CHC = Community Health Centre; MM = Multimorbidity; UAR = Urban at Risk

Table D.3: Health care use characteristics.

Measure	Value	All Clients	UAR CHC	Rural CHC	Multimorbidity
# Years of Observation	Min, Median, Max	(1, 5, 11)	(1, 6, 11)	(1, 7, 11)	(1, 8, 11)
	Mean (SD)	5.6 (3.7)	6.1 (3.8)	6.7 (3.6)	7.4 (3.3)
11 Years of Observation	n (%)	36 724 (16.6)	7976 (22.2)	5374 (25)	29 062 (28.2)
# Provider Types Seen	Min, Median, Max	(0, 4, 19)	(0, 5, 19)	(0, 5, 14)	(0, 6, 19)
	Mean (SD)	4.5 (2.3)	5.1 (2.6)	4.8 (2.1)	5.8 (2.2)
# Internal Referrals	Min, Median, Max	(0, 0, 300)	(0, 1, 300)	(0, 0, 51)	(0, 1, 300)
	Mean (SD)	1.7 (4.3)	2.8 (7.0)	1.4 (2.8)	2.8 (5.7)
# External Referrals	Min, Median, Max	(0, 1, 309)	(0, 2, 309)	(0, 1, 46)	(0, 3, 309)
	Mean (SD)	2.9 (4.5)	3.8 (5.9)	2.5 (3.2)	4.8 (5.5)
Avg. # Days/Year	Min, Median, Max	(0.2, 6, 176.9)	(0.2, 6.9, 129.7)	(0.2, 6.2, 120.3)	(0.3, 9.2, 176.9)
	Mean (SD)	8 (7.4)	9.4 (8.9)	8 (6.7)	11.4 (8.4)
Max # Days/Year	Min, Median, Max	(1, 10, 349)	(1, 12, 245)	(1, 11, 349)	(1, 17, 349)
	Mean (SD)	13.7 (13)	16.8 (16.6)	14.2 (12.1)	20.3 (14.4)
Avg. # Events/Day	Min, Median, Max	(1, 1.2, 66)	(1, 1.2, 29)	(1, 1.2, 31)	(1, 1.2, 31)
	Mean (SD)	1.3 (0.5)	1.3 (0.3)	1.3 (0.3)	1.3 (0.3)
Max # Events/Day	Min, Median, Max	(1, 3, 635)	(1, 3, 635)	(1, 3, 224)	(1, 3, 635)
	Mean (SD)	3.9 (7.8)	4.1 (8.2)	3.8 (6.1)	5.5 (10.7)

Legend: # = Number; Avg. = Average; CHC = Community Health Centre; SD = Standard Deviation; UAR = Urban at Risk.

Table D.4: Provider type counts.

Provider Type	Number of Events	% of Events
Provider Involved in Care		
Physician	3693760	30.13
Nurse Practitioner (RN-EC)	2608238	21.28
Nurse	2475621	20.19
Registered Practical Nurse (RPN)	990144	8.08
Social worker	452641	3.69
OtherUnknownUndefined	448761	3.66
Dietitian/Nutritionist	268395	2.19
Chiropodist	259101	2.11
Counselor	212799	1.74
Physiotherapist	171291	1.40
Internal Referral		
OtherUnknownUndefined	100649	26.71
Physician	73070	19.39
Nurse Practitioner (RN-EC)	37333	9.91
Dietitian/Nutritionist	30670	8.14
Nurse	29326	7.78
Social worker	28357	7.52
Physiotherapist	11210	2.97
Chiropractor	9881	2.62
Chiropodist	9741	2.58
Counselor	6068	1.61
External Referral		
OtherUnknownUndefined	183804	28.54
Dermatologist	41388	6.43
Surgeon - general	40736	6.32
Gastroenterologist	33737	5.24
Surgeon - speciality (eye, heart, brain, etc.)	29370	4.56
Physiotherapist	27639	4.29
E.N.T. specialist	25791	4.00
Urologist	22546	3.50
Gynecologist	21701	3.37
Cardiologist	20592	3.20

Table D.5: Time series clustering of care access frequency.

Cluster ID	# Clients	% Clients	Medoid
Short Term by Year			
K = 2 (SS = 0.502)			
1	11552	30.5	20, 8
2	26368	69.5	6, 2
K = 3 (SS = 0.301)			
1	15067	39.7	12, 3
2	16791	44.3	4, 1
3	6062	16.0	24, 9
K = 4 (SS = 0.142)			
1	12931	34.1	5, 1
2	13063	34.4	12, 3
3	5602	14.8	1, 2
4	6324	16.7	25, 8
K = 5 (SS = 0.211)			
1	3639	9.6	31, 8
2	11533	30.4	8, 1
3	11155	29.4	3, 2
4	7722	20.4	12, 5
5	3871	10.2	17, 11
Short Term by Quarter			
K = 2 (SS = 0.541)			
1	6068	16.0	10, 5, 5, 7, 3
2	31852	84.0	3, 1, 0, 1, 2
K = 3 (SS = 0.249)			
1	10780	28.4	6, 3, 1, 2, 1
2	20431	53.9	2, 0, 0, 0, 1
3	6709	17.7	6, 1, 3, 4, 3
K = 4 (SS = 0.044)			
1	14389	37.9	3, 0, 1, 1, 1
2	8939	23.6	6, 1, 0, 1, 2
3	9072	23.9	2, 1, 0, 1, 2
4	5520	14.6	9, 4, 2, 5, 3
K = 5 (SS = 0.121)			
1	4163	11.0	11, 8, 4, 5, 2
2	7084	18.7	3, 1, 0, 4, 1
3	17282	45.6	3, 1, 0, 1, 2
4	6111	16.1	5, 2, 1, 0, 1
5	3280	8.6	6, 0, 1, 6, 1
Long Term by Year			
K = 2 (SS = 0.553)			
1	34265	80.0	8, 3, 3, 2, 0, 1, 2, 6
2	8590	20.0	15, 24, 20, 19, 20, 27, 23, 11
K = 3 (SS = 0.149)			
1	15831	36.9	9, 4, 8, 3, 3, 2, 5, 2
2	10557	24.6	24, 9, 13, 19, 12, 12, 16, 6
3	16467	38.4	4, 0, 0, 1, 0, 0, 1, 4
K = 4 (SS = 0.155)			

1	2402	5.6	18, 35, 34, 46, 34, 39, 27, 9
2	23637	55.2	8, 3, 2, 2, 4, 1, 2, 3
3	8440	19.7	5, 0, 0, 0, 0, 1, 13, 3
4	8376	19.5	20, 8, 10, 12, 16, 11, 19, 9
K = 5 (SS = 0.136)			
1	6206	14.5	17, 7, 1, 1, 2, 2, 4, 2
2	5166	12.1	9, 13, 11, 11, 15, 21, 22, 9
3	4716	11.0	27, 16, 11, 7, 10, 10, 13, 5
4	11254	26.3	6, 2, 6, 6, 7, 7, 12, 3
5	15513	36.2	3, 0, 0, 0, 0, 1, 5, 2
Long Term by Quarter			
K = 2 (SS = 0.536)			
1	36775	85.8	4, 1, 0, 2, 0, 0, 0, 0, 0, 1, 4, 0, 0, 0, 0, 1, 2, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 1, 0, 2
2	6080	14.2	7, 1, 6, 2, 3, 10, 6, 4, 5, 6, 5, 5, 2, 4, 4, 7, 6, 4, 5, 3, 5, 4, 7, 6, 9, 11, 6, 8, 3, 4
K = 3 (SS = 0.007)			
1	16528	38.6	1, 0, 5, 0, 1, 2, 0, 0, 0, 3, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 1, 0, 2, 1, 0, 1
2	9729	22.7	5, 3, 1, 1, 1, 3, 2, 1, 3, 3, 8, 3, 1, 1, 2, 1, 1, 1, 0, 0, 1, 1, 1, 0, 1, 3, 4, 7, 2
3	16598	38.7	3, 0, 1, 0, 0, 0, 0, 0, 0, 2, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 1, 3, 0, 1, 1, 4
K = 4 (SS = 0.236)			
1	998	2.3	7, 0, 3, 4, 1, 2, 2, 1, 1, 1, 1, 1, 1, 0, 1, 2, 1, 1, 0, 2, 2, 1, 6, 2, 3, 9, 20, 10, 13, 9, 12, 13, 2
2	26775	62.5	3, 0, 1, 0, 2, 0, 0, 0, 1, 2, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 2, 1, 1, 0, 1
3	10169	23.7	3, 2, 0, 0, 4, 0, 0, 0, 0, 1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 1, 6, 2, 4, 2, 5, 1, 2
4	4913	11.5	8, 5, 4, 3, 3, 2, 3, 4, 2, 4, 6, 4, 5, 4, 4, 5, 3, 7, 5, 4, 5, 4, 3, 3, 2, 4, 2, 7, 3
K = 5 (SS = -0.031)			
1	11624	27.1	3, 2, 0, 1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 1, 0, 0, 0, 0, 2, 1, 0, 1, 0, 0, 0, 3, 1
2	6981	16.3	2, 0, 0, 0, 0, 1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 5, 3, 0, 1, 0, 1, 0, 0, 0, 2, 1
3	6448	15.0	6, 8, 3, 1, 1, 2, 3, 3, 3, 2, 6, 4, 1, 1, 3, 1, 2, 2, 0, 2, 2, 1, 2, 2, 3, 2, 2, 1, 2
4	5840	13.6	4, 3, 0, 0, 3, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 3, 0, 0, 2, 4, 9, 5, 2
5	11962	27.9	2, 0, 2, 0, 0, 0, 0, 0, 0, 0, 0, 1, 0, 0, 0, 0, 3, 0, 0, 0, 0, 0, 0, 1, 0, 2, 0, 0, 1, 0, 0, 0, 1

Legend: K = Number of clusters.

D.4 Extra Technical Details

Eleven-year period prevalence technical details:

Since not all clients received care from CHCs 2009-2019, they were not all at-risk of condition indications in their electronic health record (EHR) for the entire calendar-based period of observation. Thus, the denominator required estimation of the average or mid-point size of the population. This was challenging given that primary care EHRs represent an open cohort with no standard expectation for frequency of care, and the overall number of clients receiving care increased across calendar time (see Supplementary Figure 1). We used the following process to calculate 11-year period prevalence: *Numerator*: number of clients with at least one relevant code at any point from 2009 through 2019. *Denominator*: First, we calculated the median number of calendar-based years of observation across all eligible clients (i.e., median number of "at-risk" years): 5 years. Second, we calculated the number of clients who received any type of care at least once in each of the seven possible five-year intervals (2009-13; 2010-14; 2011-15; 2012-16; 2013-17; 2014-18; 2015-19), representing the size of the population within each of those five-year intervals. Finally, the median size of those seven cohorts was used as the denominator, representing the overall average size of the population across 11 years. The same process was followed to get estimates for the entire eligible population and for the subset of clients who receive care from urban at risk community health centres.

Appendix E

Objective 3 Extended Information

E.1 Illustrative Example: Hybrid Feature- and Similarity- Based Model

The feature- and similarity-based parts of the model (HFSM) can represent primal and dual forms, respectively.[153] Thus, a model where some features are combined using an unpenalized linear kernel will be equivalent to a model where all features are entered in logistic regression. To demonstrate this, we generated 10,000 observations according to the data generating mechanism $P(Y) = \sigma(0.25 - 1X_1 + 2X_2)$ where $X_{1,2} \sim N(0, 1)$. Logistic regression was fit with an intercept, X_1, and X_2; *HFSM* was fit with the intercept and X_1 maintained as features and X_2 included with a linear kernel.

Table E.1: Illustrative example feature coefficients

	LR	HFSM
β_0	0.24	0.24
β_1	-1.04	-1.04
β_2	2.04	NA

Table E.1 shows the learned coefficients whereby the intercept and X_1 coefficients were the same for the two models. The unpenalized α's from *HFSM* ranged from -0.001 to 0.001. As expected, predictions based on the two model forms were also equivalent (not shown).

E.2 Hybrid Model Code

The following Python code can be used to fit the four main models we used: logistic regression (M1), kernel logistic regression (M2), hybrid model sequential fit (M3), and hybrid model simultaneous fit (M4).

```
import numpy as np
import cvxpy as cp
from sklearn.metrics import roc_auc_score
from scipy.special import expit

# Function to fit feature only model
# @param Xtrain the training feature data
# @param yTrain training binary outcome
# @return betas and auc on training data

# Note that cp.logistic(x) is log(1 + exp(x)), not sigmoid

def FIT_M1(Xtrain, yTrain, save=False, fnHead=None):

    beta = cp.Variable((Xtrain.shape[1], 1))

    problemM1 = cp.Problem(cp.Maximize(cp.sum(
                    cp.multiply(yTrain, (Xtrain @ beta))
                    - cp.logistic((Xtrain @ beta)))/Xtrain.shape[0]))

    problemM1.solve(verbose=False, solver=cp.ECOS)

    # Get training AUC value
    aucTrain = roc_auc_score(yTrain, expit(Xtrain @ beta.value))

    print(f"\n␣***␣DONE␣M1␣FIT␣***"
          f"\nStatus␣of␣M1␣problem:␣{problemM1.status}"
          f"and␣Optimal␣value:␣{problemM1.value}"
          f"and␣solve␣time:␣{problemM1._solve_time}"
          f"\n**Training␣AUC:␣{aucTrain}"
          )

    if (save):
        np.save(fnHead + "_Betas.npy", beta.value)
        np.save(fnHead + "_SolveTime.npy", problemM1._solve_time)
        np.save(fnHead + "_OptValue.npy", problemM1.value)
        np.save(fnHead + "_Status.npy", problemM1.status)
        np.save(fnHead + "_aucTrain.npy", aucTrain)

    return beta.value, aucTrain



# Function to fit kernel only model with L1 penalty
# @param Ktrain precomputed training kernel
# @param yTrain training outcome
# @param l1 strength of L1 penalty for alphas
# @param fnHead start path to save object
#  including directory and foldO
# @return alphas, auc on training data
def FIT_M2(Ktrain, yTrain, l1, save=False, fnHead=None):

    alpha = cp.Variable((Ktrain.shape[1], 1))

    lam = cp.Parameter(nonneg=True, value=l1)

    problemM2 = cp.Problem(cp.Maximize(cp.sum(
                    cp.multiply(yTrain, (Ktrain @ alpha))
```

```
                    - cp.logistic(Ktrain @ alpha))/Ktrain.shape[0]
                    - lam * cp.norm(alpha, 1)))

    problemM2.solve(verbose=False, solver=cp.ECOS)

    # Get training AUC value
    aucTrain = roc_auc_score(yTrain, expit(Ktrain @ alpha.value))

    print(f"\n␣***␣DONE␣M2␣FIT␣WITH␣LAM␣{l1}␣***"
          f"\nStatus␣of␣M2␣problem:␣{problemM2.status}"
          f"and␣Optimal␣value:␣{problemM2.value}"
          f"and␣solve␣time:␣{problemM2._solve_time}"
          f"\n**Training␣AUC:␣{aucTrain}"
          )

    if (save):
        np.save(fnHead + "_Alphas.npy", alpha.value)
        np.save(fnHead + "_SolveTime.npy", problemM2._solve_time)
        np.save(fnHead + "_OptimalValue.npy", problemM2.value)
        np.save(fnHead + "_Status.npy", problemM2.status)
        np.save(fnHead + "_aucTrain.npy", aucTrain)

    return alpha.value, aucTrain


# Function to fit HFSM-Seq with L1 penalty
# Betas are fit first and fixed while learning alphas
# @param Xtrain precomputed training kernel
# @param Ktrain precomputed training kernel
# @param yTrain training outcome
# @param l1 strength of L1 penalty for alphas
# @param fnHead start path to save object
#  including directory and foldO
# @return betas, alphas, auc on training data
def FIT_M3(Xtrain, Ktrain, yTrain, l1, fixedBeta=None, save=False, fnHead=None):

    if (fixedBeta==None):
        # learn the betas ignoring alphas
        fixedBeta, aucNotUsed = FIT_M1(Xtrain, yTrain,
                                       save=True, fnHead=fnHead + "_m1Part")

    # betas are set up as fixed parameter for learning alphas
    betaM1 = cp.Parameter(fixedBeta.shape, value=fixedBeta)
    # Alphas are learned
    alpha = cp.Variable((Ktrain.shape[0], 1))
    # L1 penalty strength is fixed parameter
    lam = cp.Parameter(nonneg=True, value=l1)

    # problem to solve
    problemM3 = cp.Problem(cp.Maximize(cp.sum(
                        cp.multiply(yTrain, (Ktrain @ alpha + Xtrain @ betaM1))
                        - cp.logistic(Ktrain @ alpha + Xtrain @ betaM1))
                        / Ktrain.shape[0]
                        - lam * cp.norm(alpha, 1)))

    # call the solver; default max iters is 10,000
    problemM3.solve(verbose=False, warm_start=True, solver=cp.ECOS)

    aucTrain = roc_auc_score(yTrain,
                             expit(Xtrain @ betaM1.value + Ktrain @ alpha.value))

    print(f"\n␣***␣DONE␣M3␣FIT␣WITH␣LAM␣{l1}␣***"
          f"\nStatus␣of␣problem:␣{problemM3.status}"
          f"and␣Optimal␣value:␣{problemM3.value}"
          f"and␣solve␣time:␣{problemM3._solve_time}"
          f"\n**Training␣AUC:␣{aucTrain}"
          )
```

```
    if (save):
        np.save(fnHead + "_BetasM1.npy", betaM1.value)
        np.save(fnHead + "_Alphas.npy", alpha.value)
        np.save(fnHead + "_SolveTime.npy", problemM3._solve_time)
        np.save(fnHead + "_OptimalValue.npy", problemM3.value)
        np.save(fnHead + "_Status.npy", problemM3.status)
        np.save(fnHead + "_aucTrain.npy", aucTrain)

    return betaM1.value, alpha.value, aucTrain


# Function to fit HFSM-Sim with L1 penalty
# @param Xtrain precomputed training kernel
# @param Ktrain precomputed training kernel
# @param yTrain training outcome
# @param l1 strength of L1 penalty for alphas
# @param fnHead start path to save object
#  including directory and foldO
# @return betas, alphas, auc on training data data
def FIT_M4(Xtrain, Ktrain, yTrain, l1, save=False, fnHead=None):

    # Variables can be scalars, vectors, or matrices
    beta = cp.Variable((Xtrain.shape[1], 1))
    # vector of values (n,1) to fit
    alpha = cp.Variable((Ktrain.shape[0], 1))
    # Parameter - this one is positive scalar for lam
    lam = cp.Parameter(nonneg=True, value=l1)

    # problem to solve
    problemM4 = cp.Problem(cp.Maximize(cp.sum(
                        cp.multiply(yTrain, (Ktrain @ alpha + Xtrain @ beta))
                        - cp.logistic(Ktrain @ alpha + Xtrain @ beta))
                        / Ktrain.shape[0]
                        - lam * cp.norm(alpha, 1)))

    # call the solver; default max iters is 10,000
    problemM4.solve(verbose=False, warm_start=True, solver=cp.ECOS)

    aucTrain = roc_auc_score(yTrain,
                             expit(Xtrain @ beta.value + Ktrain @ alpha.value))

    print(f"\n *** DONE M4 FIT WITH LAM {l1} ***"
          f"\nStatus of problem: {problemM4.status}"
          f"and Optimal value: {problemM4.value}"
          f"and solve time: {problemM4._solve_time}"
          f"\n**Training AUC: {aucTrain}"
          )

    if (save):
        np.save(fnHead + "_Betas.npy", beta.value)
        np.save(fnHead + "_Alphas.npy", alpha.value)
        np.save(fnHead + "_SolveTime.npy", problemM4._solve_time)
        np.save(fnHead + "_OptimalValue.npy", problemM4.value)
        np.save(fnHead + "_Status.npy", problemM4.status)
        np.save(fnHead + "_aucTrain.npy", aucTrain)

    return beta.value, alpha.value, aucTrain
```

E.3 Simulation Study Details

E.3.1 RBF kernel sigma selection

Three candidate hyperparameter values for the RBF kernel were selected to provide a range of diagonal dominance as assessed by the following equation:

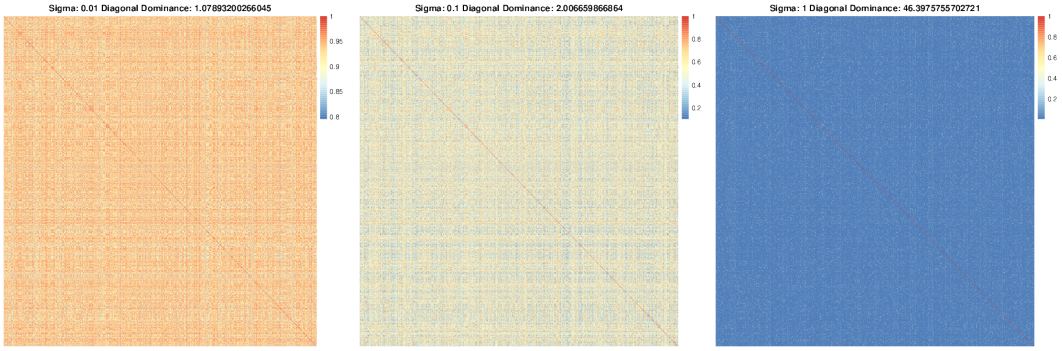
$$DD = \sum \frac{|diagonals|}{|off - diagonals|}$$

For a matrix with $i, j = n$ observations:

$$DD = \sum_{i,j} \frac{|a_{i,i}|}{|a_{i,j}| - |a_{i,i}|}$$

For a kernel matrix $a_{i,i} = 1$ and the range will be $\left[\frac{n}{(n-1)}, \infty\right)$.

In addition to looking at the above scalar measure, we generated heat plots for RBF kernels with a range of σ values on a random sample of 1000 observations of the six variables in MONK's data problems. Example plots in Figure E.1 provide another view at how varying the σ values alters the similarity captured by the RBF kernel.



$\sigma = 0.01$ $\sigma = 0.1$ $\sigma = 1$

Figure E.1: Heatmap demonstrating similarity of RBF kernel with various σ on a random sample.

E.3.2 Selected hyperparameters and model coefficients

Table E.2: Synthetic data scenario 1: selected hyperparameters

	Fold 1	Fold 2	Fold 3	Fold 4	Fold 5
Sigma for RBF Kernel					
LR	NA	NA	NA	NA	NA
KLR	1.000	1.000	1.000	1.000	1.000
HFSM-Seq	0.100	1.000	0.100	0.100	1.000
HFSM-Sim	0.100	1.000	0.100	0.100	0.100
L1 Penalty Strength					
LR	NA	NA	NA	NA	NA
KLR	0.032	0.001	0.001	1.000	1.000
HFSM-Seq	0.001	0.001	0.001	0.001	0.001
HFSM-Sim	0.001	0.001	0.001	0.001	0.001

Table E.3: Synthetic data scenario 1: model interpretation

(a) Average Feature Coefficients

	LR	HFSM-Sim	True
β_0	-1.263	-1.021	-1.500
β_1	0.320	0.319	0.300
β_2	0.440	0.444	0.400
β_3	0.568	0.567	0.600
β_4	0.695	0.695	0.700

(b) Average Kernel Coefficients

	KLR	HFSM-Seq	HFSM-Sim
Non-0	262.600	197.400	180.400
Max	0.000	0.014	0.005
Min	-0.019	-0.014	-0.019
Mean	-0.004	0.000	-0.005
Median	-0.003	-0.002	-0.006

Table E.4: Synthetic data scenario 2: selected hyperparameters

	Fold 1	Fold 2	Fold 3	Fold 4	Fold 5
Sigma for RBF Kernel					
LR	NA	NA	NA	NA	NA
KLR	1.000	1.000	1.000	1.000	1.000
HFSM-Seq	1.000	1.000	1.000	1.000	1.000
HFSM-Sim	1.000	1.000	1.000	1.000	1.000
L1 Penalty Strength					
LR	NA	NA	NA	NA	NA
KLR	0.001	0.001	0.001	0.001	0.001
HFSM-Seq	0.001	0.001	0.001	0.001	0.001
HFSM-Sim	0.001	0.001	0.001	0.001	0.001

Table E.5: Synthetic data scenario 2: model interpretation

(a) Average Feature Coefficients

	LR	HFSM-Sim	True
β_0	-0.844	-0.557	-2.100
β_1	0.185	0.198	0.300
β_2	0.377	0.408	0.400
β_3	0.443	0.457	0.600
β_4	0.566	0.605	0.700

(b) Average Kernel Coefficients

	KLR	HFSM-Seq	HFSM-Sim
Non-0	1696.200	1732.800	1631.000
Max	0.082	0.085	0.082
Min	-0.098	-0.096	-0.109
Mean	-0.002	-0.001	-0.010
Median	0.001	0.001	-0.006

Table E.6: Synthetic data scenario 3: selected hyperparameters

	Fold 1	Fold 2	Fold 3	Fold 4	Fold 5
Sigma for RBF Kernel					
LR	NA	NA	NA	NA	NA
KLR	1.000	1.000	1.000	1.000	1.000
HFSM-Seq	1.000	1.000	1.000	1.000	1.000
HFSM-Sim	1.000	1.000	1.000	1.000	1.000
L1 Penalty Strength					
LR	NA	NA	NA	NA	NA
KLR	0.001	0.001	0.001	0.001	0.001
HFSM-Seq	0.001	0.001	0.001	0.001	0.001
HFSM-Sim	0.001	0.001	0.001	0.001	0.001

Table E.7: Synthetic data scenario 3: model interpretation

(a) Average Feature Coefficients

	LR	HFSM-Sim	True
β_0	-0.459	0.414	-3.200
β_1	0.150	0.160	0.300
β_2	0.163	0.203	0.400
β_3	0.235	0.325	0.600
β_4	0.280	0.385	0.700

(b) Average Kernel Coefficients

	KLR	HFSM-Seq	HFSM-Sim
Non-0	2575.800	2599.600	2546.400
Max	0.094	0.102	0.092
Min	-0.264	-0.269	-0.297
Mean	-0.004	-0.003	-0.021
Median	0.006	0.006	-0.013

E.4 Clinical Case Study Details

E.4.1 Cohort overview

Loss to Follow Up Of the 10,687 people excluded for having less than three years between their first and last care records, 6,276 (58.7%) had their first event in 2017 or later so there was not enough calendar time for sufficient observation; bias due to their exclusions is expected to be minimal. The remaining 4,411 (41.3%) were "true" loss to follow-up under a more traditional research study paradigm; we do not know if they stopped receiving care altogether of if they switched to another health care organization. After applying additional eligibility criteria there were 1,430 clients and among them there were 108 cases of the outcome of which 22 (16.9%) occurred at least one year from the first recorded event. If our study was more application than methods testing focused we would perform sensitivity analyses to assess whether there is bias due to these lost to follow up as in a real world setting the future length of care when applying a predictive model is unknown.

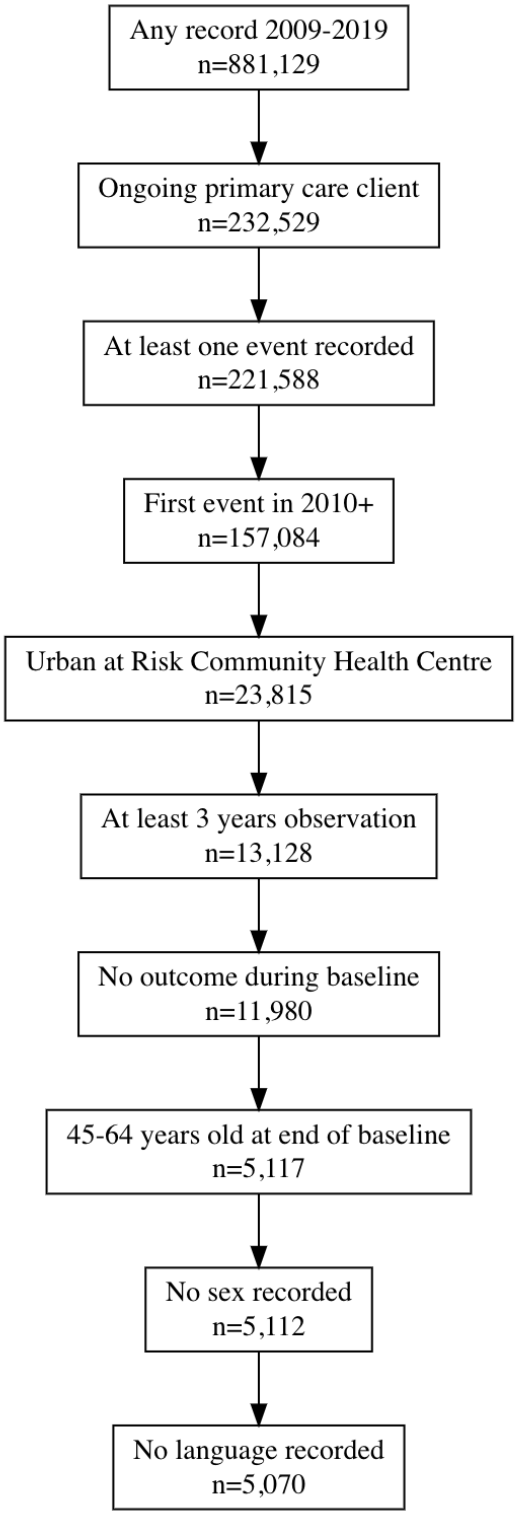


Figure E.2: Clinical case study cohort flow diagram

Table E.8: Clinical case study baseline features.

Feature	Values	n (%)
Sex	Female	2379 (46.92)
	Male	2691 (53.08)
Rural Residence	Rural	1011 (19.94)
	Urban	3942 (77.75)
	Missing	117 (2.31)
Household Income	$0 to $14,999	1254 (24.73)
	$15,000 to $24,999	454 (8.95)
	$25,000 to $34,999	274 (5.40)
	$35,000 to $59,000	535 (10.55)
	$60,000 or more	600 (11.83)
	Do not know	274 (5.40)
	Prefer not to answer	587 (11.58)
	Missing	1092 (21.54)
Household Composition	Couple	1897 (37.42)
	Other Family	519 (10.24)
	Unrelated housemates	217 (4.28)
	Sole Member	1205 (23.77)
	Do Not Know or Other	255 (5.03)
	Prefer not to answer	57 (1.12)
	Missing	920 (18.15)
Education Completed	Post-secondary or equivalent	1717 (33.87)
	Secondary or equivalent	1849 (36.47)
	Less than high school	395 (7.79)
	Do Not Know or Other	269 (5.31)
	Prefer not to answer	54 (1.07)
	Missing	786 (15.50)
Language	English	4691 (92.52)
	French	82 (1.62)
	Other	297 (5.86)
LGBTQ	Lgbtq	67 (1.32)
	Non-Lgbtq	1084 (21.38)
	Missing	3919 (77.30)
Years in Canada	True	627 (12.37)
Physical Disability	True	240 (4.73)
Depression or Anxiety	True	410 (8.09)
Chronic Urinary Problem	True	852 (16.80)
Obesity	True	737 (14.54)
Personality Disorder	True	145 (2.86)
Stable Housing	True	556 (10.97)
Substance Use	True	753 (14.85)
Smoking or Tobacco Use	True	1454 (28.68)
Food Insecurity	True	200 (3.94)

E.4.2 Application 1: prediction

Table E.9: Clinical case study selected hyperparameters.

	Fold 1	Fold 2	Fold 3	Fold 4	Fold 5
L1 Penalty Strength					
KLR	1e-04	1e-03	1e-04	1e-04	1e-04
HFSM-Seq	1e-04	1e-03	1e-03	1e-03	1e-04
HFSM-Sim	1e-03	1e-03	1e-03	1e-03	1e-04
Kernel Function & Data					
KLR	SCR_PIST2	J_ST	J_PIST2	SCR_PIST2	SCR_PIST2
HFSM-Seq	SCR_ST	SCR_PIST2	J_ST	SCR_PIST2	J_PIST2
HFSM-Sim	J_ST	SCR_PIST2	J_PIST2	SCR_PIST2	J_PIST2

Legend: J = Jaccard similarity; SCR = Common and rare code similarity;
PI = Provider type data; ST = Service type data; PIST2 = both.

E.4.3 Application 2: inference/interpretation

Table E.10: Feature coefficients for models re-trained on all data

Variable	Values	HFSM-Seq	HFSM-Sim
Intercept		-4.275	-5.268
Sex	Male	-0.261	-0.202
Rural Residence	Urban	0.462	0.227
	Missing	0.774	0.636
Household Income	$15,000 to $24,999	0.137	0.036
	$25,000 to $34,999	-0.492	-0.444
	$35,000 to $59,000	-0.948	-0.781
	$60,000 or more	-1.640	-1.508
	Do not know	0.517	0.446
	Prefer not to answer	-0.627	-0.521
	Missing	-0.310	0.061
Household Composition	Other Family	0.659	0.608
	Unrelated housemates	0.743	0.677
	Sole Member	0.909	0.833
	Do not know or other	0.069	-0.056
	Prefer not to answer	0.443	0.455
	Missing	0.742	0.462
Education Level	Secondary or equivalent	0.011	-0.075
	Less than high school	0.127	0.056
	Do not know or other	0.292	0.294
	Prefer not to answer	0.081	0.334
	Missing	-0.268	-0.156
Primary Language	French	0.114	0.671
	Other	0.448	0.509
LGBTQ	Non-Lgbtq	-0.031	0.009
	Missing	0.593	0.534
Years in Canada	True	0.277	0.102
Physical Disability	True	-0.249	-0.188
Depression or Anxiety	True	0.629	0.379
Chronic Urinary Problem	True	0.098	0.079
Obesity	True	0.145	0.019
Personality Disorder	True	0.179	0.114
Stable Housing	True	0.934	0.626
Substance Use	True	0.225	0.114
Smoking or Tobacco Use	True	0.178	-0.063
Food Insecurity	True	0.116	-0.106

Table E.11: Client characteristics stratified by kernel coefficient

Variable	Values	Zero Alpha	Positive Alpha	Negative Alpha
# of Clients		5038 (100%)	13 (100%)	19 (100%)
Loneliness/Social Isolation	Present	270 (5.36%)	6 (46.15%)	0 (0.00%)
Sex	Female	2362 (46.88%)	7 (53.85%)	10 (52.63%)
	Male	2676 (53.12%)	6 (46.15%)	9 (47.37%)
Rural Residence	Rural	1010 (20.05%)	1 (7.69%)	0 (0.00%)
	Urban	3912 (77.65%)	11 (84.62%)	19 (100.00%)
	Missing	116 (2.30%)	1 (7.69%)	0 (0.00%)
Household Income	$0 to $14,999	1240 (24.61%)	7 (53.85%)	7 (36.84%)
	$15,000 to $24,999	451 (8.95%)	3 (23.08%)	0 (0.00%)
	$25,000 to $34,999	272 (5.40%)	1 (7.69%)	1 (5.26%)
	$35,000 to $59,000	533 (10.58%)	1 (7.69%)	1 (5.26%)
	$60,000 or more	596 (11.83%)	0 (0.00%)	4 (21.05%)
	Do not know	273 (5.42%)	0 (0.00%)	1 (5.26%)
	Prefer not to answer	584 (11.59%)	1 (7.69%)	2 (10.53%)
	Missing	1089 (21.62%)	0 (0.00%)	3 (15.79%)
Household Composition	Couple	1886 (37.44%)	5 (38.46%)	6 (31.58%)
	OtherFamily	516 (10.24%)	0 (0.00%)	3 (15.79%)
	Unrelated housemates	214 (4.25%)	2 (15.38%)	1 (5.26%)
	Sole Member	1197 (23.76%)	5 (38.46%)	3 (15.79%)
	Do not know/Other	254 (5.04%)	0 (0.00%)	1 (5.26%)
	Prefer not to answer	55 (1.09%)	1 (7.69%)	1 (5.26%)
	Missing	916 (18.18%)	0 (0.00%)	4 (21.05%)
Education Level	Post-secondary or equiv	1705 (33.84%)	4 (30.77%)	8 (42.11%)
	Secondary or equivalent	1837 (36.46%)	5 (38.46%)	7 (36.84%)
	Less than high school	392 (7.78%)	1 (7.69%)	2 (10.53%)
	Do not know/Other	266 (5.28%)	3 (23.08%)	0 (0.00%)
	Prefer not to answer	54 (1.07%)	0 (0.00%)	0 (0.00%)
	Missing	784 (15.56%)	0 (0.00%)	2 (10.53%)
Primary Language	English	4660 (92.50%)	13 (100.00%)	18 (94.74%)
	French	81 (1.61%)	0 (0.00%)	1 (5.26%)
	Other	297 (5.90%)	0 (0.00%)	0 (0.00%)
LGBTQ	Lgbtq	66 (1.31%)	1 (7.69%)	0 (0.00%)
	Non-Lgbtq	1077 (21.38%)	2 (15.38%)	5 (26.32%)
	Missing	3895 (77.31%)	10 (76.92%)	14 (73.68%)
Years in Canada	True	624 (12.39%)	2 (15.38%)	1 (5.26%)
Physical Disability	True	239 (4.74%)	1 (7.69%)	0 (0.00%)
Depression or Anxiety	True	408 (8.10%)	2 (15.38%)	0 (0.00%)
Chronic Urinary Problem	True	848 (16.83%)	2 (15.38%)	2 (10.53%)
Obesity	True	732 (14.53%)	1 (7.69%)	4 (21.05%)
Personality Disorder	True	144 (2.86%)	1 (7.69%)	0 (0.00%)
Stable Housing	True	549 (10.90%)	4 (30.77%)	3 (15.79%)
Substance Use	True	745 (14.79%)	4 (30.77%)	4 (21.05%)

Smoking or Tobacco Use	True	1443 (28.64%)	8 (61.54%)	3 (15.79%)
Food Insecurity	True	195 (3.87%)	4 (30.77%)	1 (5.26%)

Following are the top 10 codes for each topic from non-negative matrix factorization on provider type and service type (PIST2) data for sub-cohorts of clients with positive, negative, and zero α coefficients.

Negative α

Topic 1 with top 10 weights [('Diagnostic test request', 1.46), ('Intermediate assessment', 1.46), ('Physician', 1.42), ('Nurse', 1.30), ('Discussion regarding the treatment plan', 1.30), ('Health advice/instructions', 1.23), ('Case management/coordination', 1.17), ('Minor assessment', 1.03), ('Discussion regarding the diagnostic findings', 0.96), ('General assessment', 0.91)]

Topic 2 with top 10 weights [('discussion', 1.41), ('Recommendation/assistance', 1.29), ('Basic support', 0.88), ('Forms completion', 0.87), ('internal referral', 0.78), ('Information provision about community resources', 0.75), ('counselling', 0.72), ('Internal consultation', 0.54), ('Case management/coordination', 0.44), ('Counselor', 0.42)]

Topic 3 with top 10 weights [('Counselor', 0.75), ('Individual counselling', 0.74), ('Forms completion', 0.73), ('Foot care', 0.66), ('Chiropodist', 0.66), ('Client intake/interview', 0.59), ('Service access coordinator', 0.49), ('Blank Services (grandfathered)', 0.48), ('Preventive care', 0.47), ('medication prescription', 0.47)]

Topic 4 with top 10 weights [('Periodic health examination', 1.10), ('Client intake/interview', 0.88), ('medication prescription', 0.75), ('Nurse Practitioner (RN-EC)', 0.65), ('discussion', 0.64), ('Discussion regarding the diagnostic findings', 0.46), ('Discussion regarding the treatment plan', 0.43), ('Diagnostic test request', 0.37), ('Intermediate assessment', 0.37), ('Preventive care', 0.37)]

Topic 5 with top 10 weights [('Individual counselling', 1.47), ('Nurse Practitioner (RN-EC)', 1.05), ('internal referral', 0.60), ('Minor assessment', 0.58), ('External referral', 0.57), ('Dietitian/Nutritionist', 0.55), ('assessment', 0.55), ('Health advice/instructions', 0.52), ('Discussion regarding the treatment plan', 0.41), ('Medication renewal', 0.41)]

Positive α

Topic 1 with top 10 weights [('Consultation (grandfathered)', 1.25), ('Health advice/instructions', 1.12), ('referral', 1.05), ('discussion', 1.05), ('Advocacy', 1.05), ('Internal consultation', 1.05), ('Physician', 1.02), ('Nurse', 1.02), ('assessment', 0.97), ('Basic support', 0.97)]

Topic 2 with top 10 weights [('External consultation', 0.74), ('External referral', 0.61), ('Minor assessment', 0.60), ('Social worker', 0.57), ('Transportation assistance', 0.49), ('Individual counselling', 0.49), ('Intermediate assessment', 0.45), ('Information provision about community resources', 0.41), ('medication prescription', 0.39), ('Community Health Worker', 0.35)]

Topic 3 with top 10 weights [('Preventive care', 0.93), ('Client intake/interview', 0.68), ('Discussion regarding the treatment plan', 0.61), ('Chronic illness monitoring', 0.60), ('Discussion regarding the diagnostic findings', 0.60), ('assessment', 0.58), ('Basic support', 0.58), ('Community Health Worker', 0.57), ('care', 0.55), ('Health advice/instructions', 0.54)]

Topic 4 with top 10 weights [('Social worker', 0.80), ('Registered Practical Nurse (RPN)', 0.80), ('Individual counselling', 0.79), ('Case management/coordination', 0.66), ('health examination', 0.59), ('Minor assessment', 0.55), ('External referral', 0.5), ('Physician', 0.46), ('Nurse', 0.46), ('Outreach Worker', 0.45)]

Topic 5 with top 10 weights [('Minor assessment', 0.85), ('Outreach Worker', 0.62), ('Case management/coordination', 0.60), ('Social worker', 0.59), ('General assessment', 0.56), ('Dietitian/Nutritionist', 0.46), ('Foot care', 0.46), ('Diagnostic test request', 0.46), ('Discussion regarding the diagnostic findings', 0.46), ('Registered Practical Nurse (RPN)', 0.37)]

Zero α

Topic 1 with top 10 weights [('Health advice/instructions', 5.89), ('Nurse Practitioner (RN-EC)', 5.34), ('Discussion regarding the treatment plan', 4.96), ('Intermediate assessment', 4.85), ('Minor assessment', 4.46), ('Nurse', 4.29), ('Physician', 3.85), ('Diagnostic test request', 3.79), ('medication prescription', 3.75), ('Discussion regarding the diagnostic findings', 3.61)]

Topic 2 with top 10 weights [('Basic support', 2.98), ('Advocacy', 2.90), ('Recommendation/assistance', 2.74), ('discussion', 2.58), ('counselling', 2.32), ('Consultation (grandfathered)', 2.26), ('assessment', 2.04), ('Triage', 1.85), ('referral', 1.82), ('Internal consultation', 1.63)]

Topic 3 with top 10 weights [('General assessment', 2.99), ('internal referral', 2.77), ('Individual counselling', 2.56), ('Physician', 2.29), ('Internal consultation', 2.21), ('Dietitian/Nutritionist', 2.16), ('Nurse', 2.13), ('External referral', 1.93), ('Diagnostic test request', 1.81), ('Consultation (grandfathered)', 1.73)]

Topic 4 with top 10 weights [('care', 2.42), ('Mental health care', 2.08), ('Preventive care', 1.93), ('Chronic illness monitoring', 1.88), ('Individual counselling', 1.85), ('Dietitian/Nutritionist', 1.40), ('assessment', 1.36), ('Blank Services (grandfathered)', 1.36), ('Dispensing medication', 1.33), ('counselling', 1.33)]

Topic 5 with top 10 weights [('Information provision about community resources', 3.65), ('Community Health Worker', 2.09), ('Client intake/interview', 2.01), ('Case management/coordination', 1.99), ('Forms completion', 1.96), ('Recommendation/assistance', 1.86), ('Social worker', 1.50), ('Individual counselling', 1.24), ('internal referral', 1.23), ('Health advice/instructions', 1.21)]

Appendix F

Frequent Visitor Analyses

Objective: To describe the characteristics of frequent visitors to primary care providers (PCPs: nurse practitioner or medical doctor), and risk factors for a period of frequent visits.

Rationale: Clients on the highest end of the visit frequency spectrum may be accessing care more than necessary and benefit from interventions such as social prescribing that address upstream drivers of care access patterns [1]. Community Health Centres (CHCs) within the Alliance for Healthier Communities provide individual and group programs from a variety of disciplines to support clients in this way, so understanding the needs of frequent visitors is a key area of interest.

F.1 Methods

Two analyses were performed: one to compare all-time frequent visitors with remaining clients and one to explore risk factors for a quarter-year period of frequent visits. Note that in contrast to the care frequency analyses in Chapter 4, which included care visits to any type of care provider, the following two analyses calculated client observation time and care frequency based solely on PCP-associated visits. Cut-offs to classify "frequent visitors" were CHC-specific such that the same definition was used across CHCs, but the resulting values were allowed to differ by CHC.

F.1.1 All-time frequent visitor characteristic comparison

For each CHC, *frequent visitors* included clients in the top 10% of average days of PCP care per observation year. Table-based comparisons of frequent visitors with remaining clients, grouped across all CHCs, were performed for sociodemographic, clinical, and healthcare use characteristics that have over 50% completeness and perceived importance. Characteristics were defined as in Chapter 4 except categories with less than 1% were treated as missing. Eligible clients must have had at least one PCP-associated visit in their care history.

Categorical variables were compared using chi-squared test of independence. The exact P-values were reported and the chi-square value converted to Cohen's w effect size. A one-sided 95% confidence interval was provided; the upper-bound was fixed. Continuous variables were compared using Cohen's d with P-value and two-sided 95% confidence interval reported. Guidelines were used to interpret the Cohen's d and w effect sizes [2,3].

F.1.2 Risk factors for a period of frequent visits

The risk factor analysis considered each quarter-year, not including the first year that care is recorded, as a period where a client may have been a frequent visitor. The *frequent visitor* cut-off for a given CHC was the top 10% of the maximum quarter-year care access frequencies (not including the first year due to it being distinctly different in terms of care needs) across all clients from that CHC. Any given client may have

had zero or more "frequent visitor" quarter-years of observation based on their CHC-specific cut-off. Note the maximum year was used instead of the average as was done for the first analysis because of the sparsity of data; over 60% of clients had at least one quarter-year in the top decile of their CHC average client days/observation-quarter, which suggested a lot of clients had some higher quarter(s) but this analysis was intended to focus on the extremes. Eligible clients required at least one visit with a PCP that was over one year from their first recorded PCP visit.

To estimate the risk of frequent visitor status in the next quarter-year (after year one), we performed multilevel modelling using logistic regression and blocked variable entry. Quarter-years were nested within clients, the intercept was random, and the estimated effects of risk factors were fixed. Characteristics from above with expected importance and under 50% missingness were included as candidate risk factors, entered in three sequential blocks based on expected ordering of effect:

1. **Sociodemographic baseline factors**: Age, Rural residence, Sex, Education, Language, Household income, Number of people supported, Length of time in Canada.
2. **Clinical conditions or issues addressed**: Stable residence, Food insecurity, Substance use, Lonely or isolated, Smoking or tobacco use, Number of chronic conditions, Hep C.
3. **Past healthcare use**: Average number of events/day, Number of External referrals, Number of provider types seen.

To operationalize each risk factor, data up to and including the "baseline quarter", or quarter prior to the outcome quarter, were used. Block 1 factors were time-invariant except for age. For block 2, the count of chronic conditions (20 possible) was cumulative while the remaining conditions (e.g., substance use) were considered present or absent depending on care received only in the baseline quarter. All block 3 measures were cumulative. We used the bam method from R package `mgcv` to fit the models using a Maximum Likelihood estimator to allow use of likelihood ratio tests to compare the nested models [4,5]. An intercept-only model wqs also compared.

F.2 Results

F.2.1 All-time frequent visitor characteristic comparison

Of the 210,488 eligible clients there were 2,608,238 (41.4%) nurse practitioner- and 3,693,760 (58.6%) medical doctor-associated events. CHC-specific "frequent visitor" cut-offs ranged from 4.5 to 26.8 days per observation year (Figure F.1).

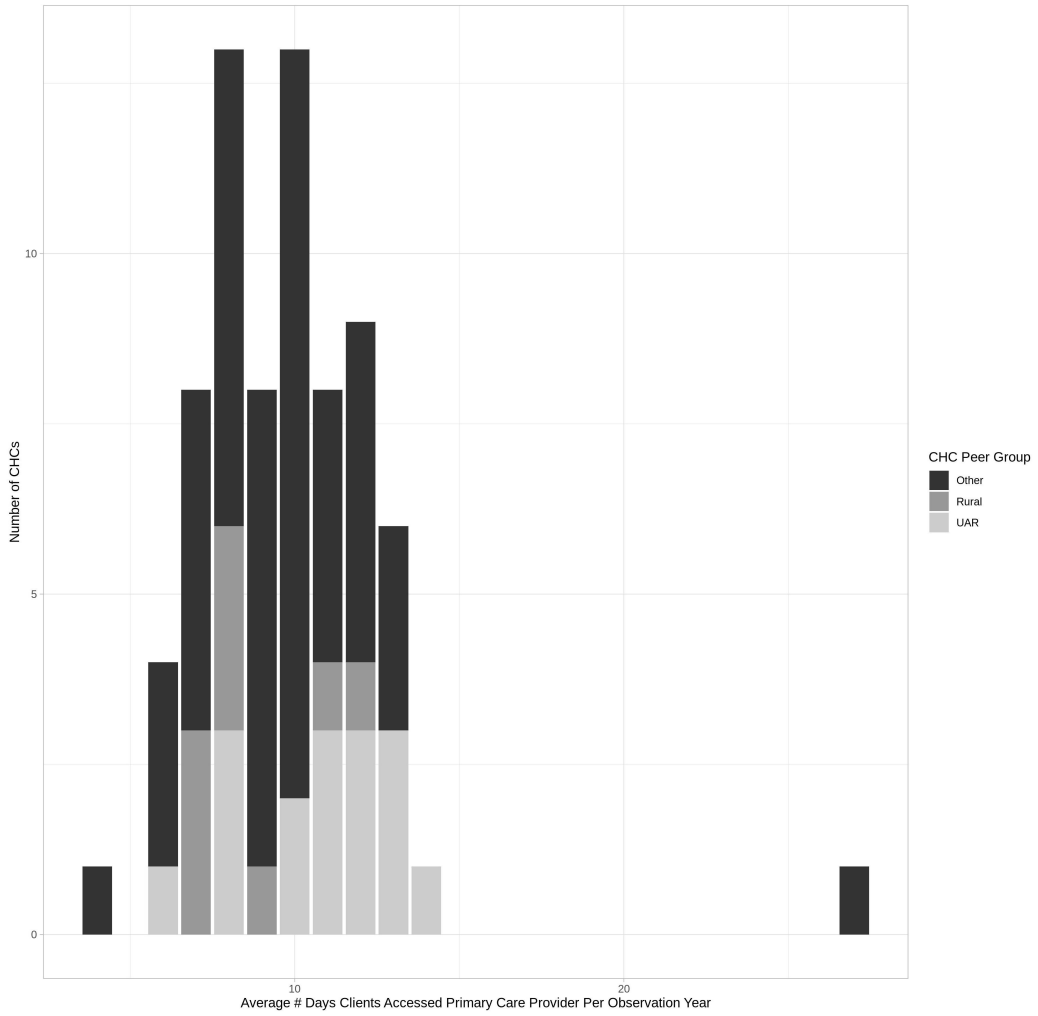


Figure F.1: Frequent visitor cut-off for all-time access.

Table F.1 shows the crude characteristic comparisons; effect estimates were in Table F.2. Number of chronic conditions, number external referrals by PCP, and number of provider types seen had large effect sizes; age had a moderate effect size; and remaining characteristics had small effect sizes.

Table F.1: Frequent visitor univariate characteristic comparisons.

Characteristic	Values	Frequent Visitor	Non-Frequent Visitor
n		21447 (100%)	189041 (100%)
Age in 2019	min, median, max	(30, 61, 113)	(30, 51, 113)
	mean (sd)	60.6 (17.3)	52.3 (15.4)
Rural Residence	Urban	16719 (77.95%)	142144 (75.19%)
	Rural	4349 (20.28%)	43535 (23.03%)
	Missing	379 (1.77%)	3362 (1.78%)
Sex	Male	7038 (32.82%)	81855 (43.3%)
	Female	14376 (67.03%)	106621 (56.4%)
	Missing	33 (0.15%)	565 (0.3%)
Education Level	Post-secondary or equivalent	6589 (30.72%)	74720 (39.53%)
	Secondary or equivalent	6798 (31.7%)	52436 (27.74%)
	Less than high school	2738 (12.77%)	15183 (8.03%)
	Other	787 (3.67%)	7360 (3.89%)
	Do not know	634 (2.96%)	3844 (2.03%)
	Prefer not to answer	350 (1.63%)	2489 (1.32%)
	Missing	3551 (16.56%)	33009 (17.46%)
Household Income	$0 to $14,999	5203 (24.26%)	32705 (17.3%)
	$15,000 to $24,999	2795 (13.03%)	17184 (9.09%)
	$25,000 to $39,999	2070 (9.65%)	17957 (9.5%)
	$40,000 to $59,999	1233 (5.75%)	15454 (8.17%)
	$60,000 or more	1334 (6.22%)	26459 (14%)
	Do not know	1627 (7.59%)	12934 (6.84%)
	Prefer not to answer	2743 (12.79%)	23974 (12.68%)
	Missing	4442 (20.71%)	42374 (22.42%)
Primary Language	English	16432 (76.62%)	142883 (75.58%)
	French	2273 (10.6%)	19510 (10.32%)
	Other	2406 (11.22%)	22817 (12.07%)
	Missing	336 (1.57%)	3831 (2.03%)
Years in Canada	0-5 years	1136 (5.3%)	11545 (6.11%)
	6+ years	4543 (21.18%)	44121 (23.34%)
	None recorded	15768 (73.52%)	133375 (70.55%)
# Chronic Conditions	min, median, max	(0, 6, 18)	(0, 2, 16)
	mean (sd)	5.8 (3.4)	2.8 (2.5)
Hepatitis C	True	602 (2.81%)	1817 (0.96%)
Smoking or Tobacco Use	True	5475 (25.53%)	31628 (16.73%)
Substance Use	True	4032 (18.8%)	16457 (8.71%)
Food Insecurity	True	2798 (13.05%)	7878 (4.17%)
Lonely or Isolated	True	4580 (21.35%)	12989 (6.87%)
Stable Residence	True	16542 (77.13%)	173121 (91.58%)
# Provider Types	min, median, max	(1, 6, 19)	(1, 4, 19)
	mean (sd)	6.4 (2.6)	4.4 (2.2)
# Ext Refs by PCP	min, median, max	(0, 3, 182)	(0, 1, 54)

	mean (sd)	4.9 (6)	1.8 (2.8)
Avg. # PCP Events/Day	min, median, max mean (sd)	(1, 1.1, 3.1) 1.1 (0.1)	(1, 1, 36) 1 (0.1)
# Years Obsv	min, median, max mean (sd)	(1, 5, 11) 5.8 (3.7)	(1, 5, 11) 5.4 (3.7)

Table F.2: All-time frequent visitor univariate statistical comparisons.

Characteristic	Chi-Square	P-value	Effect Estimate	95% CI	Interpretation
Age in 2019	NA	NA	-0.535	(-0.55, -0.52)	moderate
Rural residence	83.6	0	0.020	(0.02, 1)	very small
Sex	894.2	0	0.065	(0.06, 1)	very small
Education Level	1103.7	0	0.072	(0.07, 1)	very small
Household Income	1899.7	0	0.095	(0.09, 1)	very small
Primary Language	36.4	0	0.013	(0.01, 1)	very small
Years Since Arrival in Canada	83.7	0	0.020	(0.02, 1)	very small
# Chronic Conditions	NA	NA	-1.172	(-1.19, -1.16)	large
Hepatitis C	576.0	0	0.052	(0.05, 1)	very small
Smoking or Tobacco Use	1026.1	0	0.070	(0.07, 1)	very small
Substance Use	2232.6	0	0.103	(0.1, 1)	small
Food Insecurity	3151.9	0	0.122	(0.12, 1)	small
Lonely or Isolated	5280.2	0	0.158	(0.15, 1)	small
Stable Residence	4509.1	0	0.146	(0.14, 1)	small
# Provider Types	NA	NA	-0.896	(-0.91, -0.8)	large
# Ext Refs by PCP	NA	NA	-0.930	(-0.94, -0.92)	large
Avg. # PCP Events/Day	NA	NA	-0.202	(-0.22, -0.19)	small
# Years Obsv	NA	NA	-0.101	(-0.11, -0.09)	very small

F.2.2 Risk factors for a period of frequent visits

Of the 163,230 eligible clients and 4,132,848 client-quarters of observation, 21,339 (13.1%) had at least frequent visit quarter-year and 8,398 (39.4%) had more than one frequent visitor quarter-year. Cut-offs for a frequent visitor quarter ranged from three to 17 (Figure F.2).

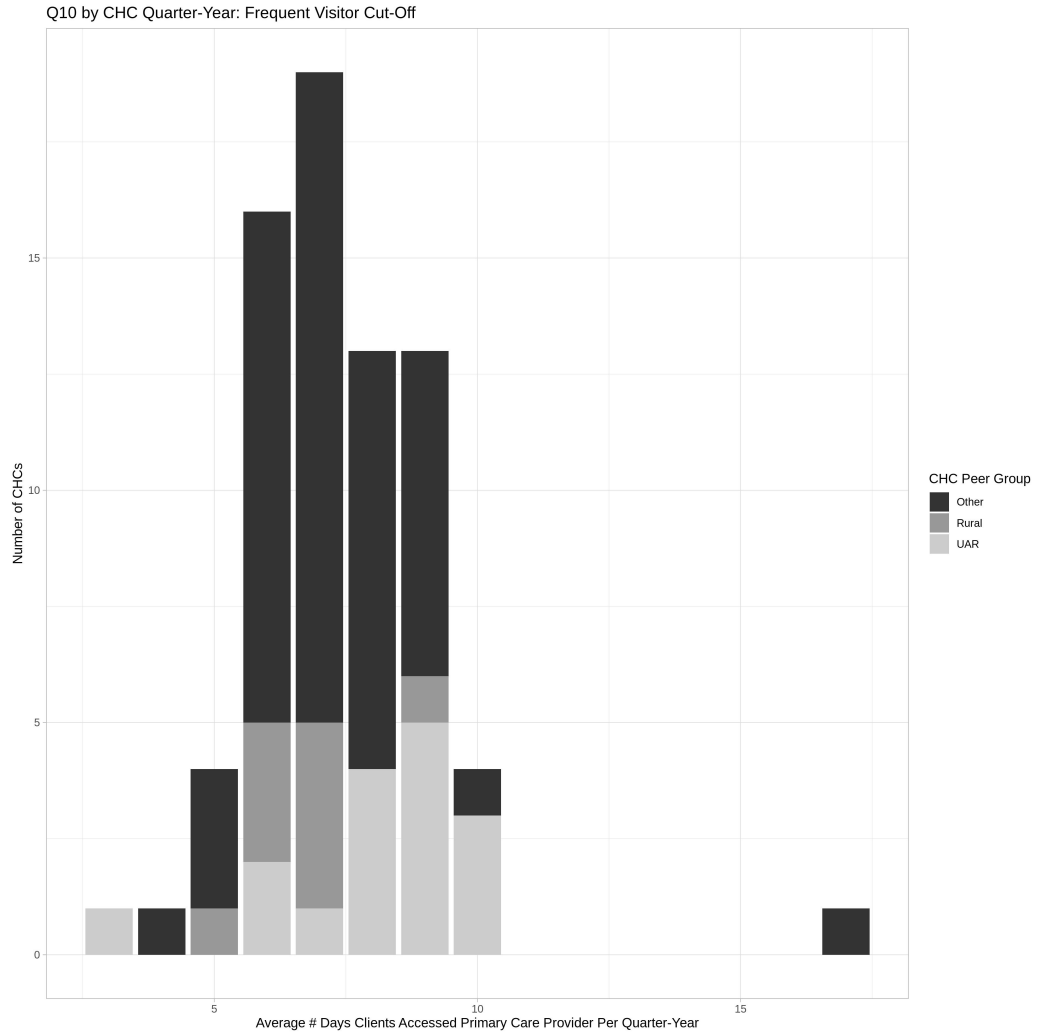


Figure F.2: Frequent visitor cut-off for quarter-year.

Odds ratios for all three models are in Table F.3. In general, odds ratio magnitudes were attenuated as more variables entered the model. Likelihood ratio tests between sequential models were all statistically significant (P-value $<$ 2.2e-16).

Table F.3: Frequent visitor risk factor analyses.

Variable	Value	Block 1	Block 2	Block 3
Quarter Age	Years	1.02	1.00	1.00
Rural	Rural	0.78	0.89	0.92
	Missing	1.41	1.06	1.07
Sex	Female	1.50	1.48	1.45
	Missing	1.42	1.27	1.33
Education	Secondary or equivalent	1.17	1.05	1.06
	Less than high school	1.41	1.27	1.27
	Other	1.29	1.17	1.20

	Do not know	1.40	1.20	1.22
	Prefer not to answer	1.53	1.31	1.33
	Missing	1.10	1.11	1.14
Household Income	$40,000 to $59,999	0.75	0.87	0.88
	$25,000 to $39,999	0.56	0.79	0.80
	$15,000 to $24,999	0.41	0.62	0.64
	$0 to $14,999	0.32	0.52	0.54
	Do not know	0.77	0.96	0.94
	Prefer not to answer	0.58	0.80	0.82
	Missing	0.62	0.86	0.90
Language	French	0.75	0.94	0.96
	Other	0.77	0.81	0.79
	Missing	0.77	0.79	0.78
Year in Canada	6+ years	1.00	0.86	0.79
	0 to 5 years	1.26	0.93	0.87
# Chronic Conditions	Count	-	1.27	1.20
Hepatitis C	Present	-	2.84	2.65
Smoking or Tobacco Use	Present	-	1.95	1.91
Substance Use	Present	-	3.95	3.83
Food Insecurity	Present	-	1.77	1.66
Lonely or Isolated	Present	-	2.15	2.00
Stable Housing	Present	-	2.50	2.24
Cumulative # Provider Types	Count	-	-	1.07
Cumulative # External Referrals	Count	-	-	1.03
Average # Events/Day	Count	-	-	1.40

F.3 Discussion

Univariate comparisons between frequent visitors based on all-time healthcare use found large effect sizes for frequent visitors having more provider types involved in care, external referrals, and chronic conditions. There were no large effect sizes for sociodemographic conditions in these analyses; however, some associations did emerge in the risk factor analysis based on discrete periods of healthcare use. In general findings further support the limited existing research on care frequency of primary care in Canada showing the importance of both social and medical complexity in general [6,7]. Specific findings in contrast with other populations include a positive association of female sex [6] and housing stability [7] with the outcome. Future work should include causal analyses to identify the potential impact on intervenable risk factors, or the development of a predictive model to identify individual clients who are at high risk of frequent healthcare use in the future. Future work could explore how to identify specific clients that are expected to experience higher than necessary care frequency in the future and whether interventions such as social prescribing may be beneficial.

F.4 References

1. Alliance for Healthier Communities. Towards a Learning Health System: Better Care Tomorrow When We Learn from Today. Alliance for Healthier Communities; 2020:15. Accessed November 23, 2020. https://www.allianceon.org/sites/default/files/documents/Learning%20Health%20System%20report%202020-10-20%20-%20FINAL_JR.pdf
2. Gignac GE, & Szodorai ET. Effect size guidelines for individual differences researchers. *Pers Individ Differ.* 2016;102:74–78. https://doi.org/10.1016/j.paid.2016.06.069
3. CohenJ. (1988). Statistical power analysis for the behavioral sciences (2nd ed). L. Erlbaum Associates.
4. Wood SN, Goude Y, & ShawS. Generalized additive models for large data sets. *J R Stat Soc Ser C Appl Stat.* 2015;64(1):139–155. https://doi.org/10.1111/rssc.12068
5. Li Z, & Wood SN. Faster model matrix crossproducts for large generalized linear models with discretized covariates. *Stat Comput.* 2020;30(1):19–25. https://doi.org/10.1007/s11222-019-09864-2
6. Singer A, Kosowan L, Katz A, Ronksley P, McBrien K, Halas G, & Williamson T. Characterizing patients with high use of the primary and tertiary care systems: A retrospective cohort study. *Health Policy.* 2020;124(3):291–297. https://doi.org/10.1016/j.healthpol.2020.01.011
7. Currie LB, Patterson ML, Moniruzzaman A, McCandless LC, & Somers JM. Examining the relationship between health-related need and the receipt of care by participants experiencing homelessness and mental illness. *BMC Health Serv Res.* 2014;14(1):404. https://doi.org/10.1186/1472-6963-14-404

Appendix G

Curriculum Vitae

Jacqueline K. Kueper

Education

The University of Western Ontario 2017-Present
Ph.D. Candidate in Epidemiology & Computer Science

The University of Western Ontario 2015-2017
M.Sc. Epidemiology & Biostatistics

McGill University 2010-2015
B.Sc. Biological, Biomedical, & Life Sciences

Honours and Awards

Canadian Institutes for Health Research Frederick Banting & Charles Best Canada Graduate Scholarship–Doctoral 2018-2021

Queen Elizabeth II Graduate Scholarship in Science & Technology (declined) 2018-2019

Ontario TUTOR-PHC Fellowship funded by INSPIRE-PHC 2018

Alzheimer Society of London and Middlesex Master's Scholarship 2016-2017

Dr. Carol Buck Graduate Scholarship in Epidemiology 2015-2016

McGill University J.W. McConnell Scholarship 2010-2011

Research Grants

Canadian Institutes for Health Research, The Institution of Population and Public Health Planning and Dissemination Grant 2022-2028
CAD $2,400,000 for *Artificial Intelligence for Public Health (AI4PH) Training Platform.* Rosella LC, Anderson M, Buckeridge D, Fan L, Lee J, Lix L, Osgood N, Banack H, Brook J, *Carabali M, Chaiton M, DiRuggiero E, Dolatabadi E, Donnelly C,*Fisher S, Fuller D, Gibson J, Green M, Guttman A, Hardcastle L, *Harish V, Janjua N, King N, *Kueper J, Lizotte D, Manuel D, McGrail K, McLaughlin J, Mitani A, *Morgenstern J, Ondrusek N, Pagalan L, Quesnel-Vallee A, Risling T, Saarela O, Sanmartin C, Shaw J, Siddiqi A, Stoner B, Sultan R, Tranmer J, Wang B, Weichenthal S, Williamson T, Damestoy N, Leggett C, Linke C, Michel P, Moore B, Pisano V, Schull M, Strome E. (*Denotes trainees)

Canadian Institutes for Health Research, Operating Grant: Addressing the Wider Health Impacts of COVID-19 02/2022-01/2024
CAD $455,244 for *Exploring the Untold Story of COVID-19: Understanding the Wider and Future Impacts of the Pandemic and Finding Solutions to Improve Population Health, Resiliency, and Preparedness.* Ryan BL, (Nominated PI), Terry AL (Co-PI): Co-investigators: Ali S, Bayliss L, Black J, Brown JB, Cejic S, Cipriano L, Freeman T, Jan S, Kueper J, Lizotte D, Mathews M, Meyer M, Nicholson K, Ranade S, Rayner J, Sedig K, Silverman M, Speechley M, Stranges S, Summers A, Thind A, Vingilis E, Wetmore S.

INSPIRE-PHC Applied Health Research Question (AHRQ) 2020-2021
CAD $36,166 for *Creating an Action Plan for the Use of Artificial Intelligence: COVID-19 Pandemic and Recovery in Primary Health Care in Ontario.* Co-PIs: Kueper JK, Lizotte DJ, Terry AL. Co-Investigators: Beleno R, Brown JB, Cejic S, McKay S, Léger D, Pinto A, Ryan BL, Stewart M, Zwarenstein M.

Related Work Experience

CFPC-AMS Tech*Forward* Fellow, The College of Family Physicians of Canada 04/2021-Present

Graduate Teaching Assistant, The University of Western Ontario 2017-2021

Weekend Residential Support & Relief Worker, Native Women's Shelter of Montreal 02/2014-08/2015

Selected Publications

Kueper JK, Terry A, Bahniwal R, Meredith L, Beleno R, Brown JB, Dang J, Leger D, McKay S, Pinto A, Ryan BL, Zwarenstein M, Lizotte DJ. Connecting Artificial Intelligence and Primary Care Challenges: Findings from a Multi Stakeholder Col-

laborative Consultation. *BMJ Health & Care Informatics*. 2022;29:e100493. doi: 10.1136/bmjhci-2021-100493. *Selected as BMJ HCI Editors' Choice in May 2022.*

Chisholm A, Wang J, Bonnell LN, Duwe E, Gilfoyle M, Kueper JK, Locher I, Gebauer S. From NAPCRG: Primary Care Research Through the lens of NAPCRG's Trainee Committee: A Year of Reflection in a Pandemic and a Call to Action. *The Annals of Family Medicine*. 2022;20(1):98-99. doi: 10.1370/afm.2778. [Special Report]

Kueper JK. Primer for Artificial Intelligence in Primary Care. *Canadian Family Physician*. 2021;67:889-893. [Commentary]

Brogly C, Shoemaker JK, Lizotte DJ, Kueper JK, Bauer MA. Smart Healthy Campus: A Mobile Application to Identify Lifestyle Indicators Related to Undergraduate Mental Health. *JMIR Formative Research*. 2021;5(10):e29160. doi:10.2196/29160

Black J, Kueper JK, Terry AL, Lizotte DJ. Development of a Prognostic Prediction Model to Estimate the Risk of Multiple Chronic Diseases: Constructing a Copula-Based Model Using Canadian Primary Care Electronic Medical Record Data. *International Journal of Population Data Science*. 2021;6(1). doi: 10.23889/ijpds.v6i1.1395

Kueper JK, Terry AL, Zwarenstein M, Lizotte DJ. Artificial Intelligence and Primary Care Research: A Scoping Review. *Annals of Family Medicine*. 2020;18(3):250-258. doi:10.1370/afm.2518

Alsabbagh W, Kueper JK, Johnston S, Burge F, Glazier R, McGrail K, Blackman S, Wong S. Development and Evaluation of the Comparability of Primary Health Care Indicators from Administrative Data Across Three Canadian Provinces. *International Journal of Population Data Science*. 2020;5(1). doi:10.23889/ijpds.v5i1.1340

Kueper JK, Alsabbagh W, Peterson S, Wong ST, for the TRANSFORMATION Study. A Concept Dictionary for Achieving Cross Provincial Comparisons of Osteoporosis Screening Performance in Primary Health Care. *International Journal of Population Data Science*. 2019;4(1). doi:10.23889/ijpds.v4i1.1116

Jordan S, Fazelpour S, Koshiyama A, Kueper J, DeChant C, Leong B, Marchant G, Shank C. Creating a Tool to Reproducibly Estimate the Ethical Impact of Artificial Intelligence. *AI Pulse*. 2019. [Essay]

Selected Presentations

University of Western Ontario Research Computer Science 05/25/2022
Oral: Hybrid Feature- and Similarity-Based Models for Prediction and Interpretation on Large-Scale Observational Data.

5th Emirates Family Medicine Society Congress 03/14/2022
Invited keynote: Artificial Intelligence and its Application in Family Medicine.

North American Primary Care Research Group Conference 11/21/2021

Oral: Identifying Priorities for Artificial Intelligence and Primary Care in Ontario: A Multi-Stakeholder Engagement Event.

North American Primary Care Research Group Conference 11/19/2021
Pre-Conference workshop: Rising Stars in Research - Presented by the NAPCRG Trainee Program.

Trillium Primary Health Care Conference 10/15/2021
Oral: Identifying Priorities for Artificial Intelligence and Primary Care in Ontario: A Multi-Stakeholder Engagement Event.

Gulf Coast Cluster for AI in Healthcare Webinar 09/29/2021
Invited panelist w\ Andrew Bazemore, Steven Lin: Training Clinicians to Use AI.

Setting a Research Agenda for the Use of Artificial Intelligence & Machine Learning in Primary Care co-hosted by the American Board of Family Medicine and Stanford University 03/18/2021
Invited panelist w\ Arlene Bierman, Mark Sendak, Miguel Marino, Sherri Rose: AI/ML & Primary Care: Opportunities to Transform Research.

The Future is Now: AI in Family Medicine Webinar Series co-hosted by the College of Family Physicians of Canada and Upstream Lab 12/10/2020
Invited oral w\ Andrew Pinto: Introduction to AI and Applications to Family Medicine.

North American Primary Care Research Group Conference 11/21/2020
Poster: Identifying Clusters of Conditions and their Sociodemographic Patterns Addressed by Community Health Centers in Ontario.

Power of Population Data Science Webinar Series 02/06/2020
Invited oral w\ Sabrina Wong, Waseem Alsabbagh: Developing a Multi-Jurisdictional, Comparable Measure of Osteoporosis Screening Performance from Administrative Health Data.

North American Primary Care Research Group Conference 11/19/2019
Oral: A Scoping Review on Artificial Intelligence and Primary Care: Where is the Research Field Now and Where Does it Need to Go?

Artificial INTELLIGENce for efficient community based primary health CARE (INTELLIGENT-CARE) CIHR-sponsored Workshop 09/20/2019
Invited oral w\ Daniel J. Lizotte: Introduction to Artificial Intelligence, and Scoping Review on Artificial Intelligence and Primary Care Research.

Trillium Primary Health Care Research Day 06/05/2019
Poster: Artificial Intelligence and Primary Care: What Research Has Been Done and How Do We Move Forward?

Society for Medical Decision Making 40th Annual North American Meeting 10/17/2018

Poster: Impact of Social Determinants of Health Information on Predictive Models for Chronic Kidney Disease in Primary Health Care.

Canadian Student Health Research Forum 06/12-14/2018
Poster: A Topic Modelling Approach to Understanding the Body of Research on Artificial Intelligence & Primary Care. *Received honourable mention in the doctoral student competition*

Fallona Family Interdisciplinary Research Showcase 04/12/2018
Poster: Artificial Intelligence and Primary Care: An Interdisciplinary Scoping Review Protocol & Beyond.

Student Supervision

Mahzabeen Emu Mitacs Business Strategy Internship PhD Student Fellow 2022
Co-supervised w\ Salimur Choudhury for "Advancement of Artificial Intelligence for Family Medicine Research" at the College of Family Physicians of Canada.

Natalie Pallisco UWO Undergraduate Scholar's Electives Program F2020-W2021
Co-supervised w\ Daniel J. Lizotte for "Processing and analysing records from Ontario Community Health Centres to better understand differing needs of urban and rural populations".

Raveen Bahniwal, UWO Master of Public Health Practicum Student S2020
Co-supervised w\ Daniel J. Lizotte and Amanda L. Terry for "Identifying the Connection Between Artificial Intelligence and Primary Care Challenges in Response to the COVID-19 Pandemic".