

INFORMATICS STRATEGIES FOR RISK STRATIFICATION
IN POPULATION HEALTH MANAGEMENT

A Dissertation
presented to
the Office of Research and Graduate Studies
at the University of Missouri-Columbia

In Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy

by
LINCOLN SHEETS
Dr. Chi-Ren Shyu, Dissertation Supervisor

MAY 2017

© Copyright by Lincoln Sheets 2017

All Rights Reserved

The undersigned, appointed by the Office of Research and Graduate Studies, have examined the dissertation entitled

INFORMATICS STRATEGIES FOR RISK STRATIFICATION
IN POPULATION HEALTH MANAGEMENT

presented by Lincoln Sheets,

a candidate for the degree of Doctor of Philosophy,

and hereby certify that, in their opinion, it is worthy of acceptance.

Professor Chi-Ren Shyu

Professor Jerry C. Parker

Professor Lori L. Popejoy

Professor Mihail Popescu

Professor Paul Fontelo

“All that I am, or hope to be, I owe to my angel mother.” – Abraham Lincoln

I also owe my being and my works to my father, Dr. Ralph Waldo Sheets; and to the mothers of my children, my many friends and family, and my wonderful mentors.

Thank you.

ACKNOWLEDGEMENTS

I would like to thank Professors Shyu, Parker, Popejoy, Popescu, and Fontelo for their patient and careful guidance of my studies. The work presented here is also indebted to many collaborators, foremost among whom are:

Gregory F. Petroski, PhD

Michael A. Phinney, MS

Kayson Lyttle, BSc

Yan Zhuang, BS

Mohammed Khalilia, PhD

Soo-Yeon Cho, PhD

Julie Jaddoo, MSHI

Ellen Cain, MBA

Raymond Nguyen, MS

Sean Lander, MS

Bin Ge, MA, MD

Hongfei Cao, PhD

Working with these brilliant and pleasant individuals has been a privilege, and the best part of my doctoral education. Any errors that may be discovered in this presentation are mine, not theirs.

TABLE OF CONTENTS

ACKNOWLEDGEMENTS	ii
LIST OF ILLUSTRATIONS	v
LIST OF TABLES	vi
LIST OF ABBREVIATIONS	vii
ABSTRACT	viii
Chapter	
CHAPTER 1. INTRODUCTION	1
<i>1.1 Nature and purpose of the study</i>	1
<i>1.2 Significance of the study</i>	2
<i>1.3 Innovative informatics strategies and applications</i>	4
<i>1.4 Objectives of the study</i>	5
CHAPTER 2. LITERATURE REVIEW	7
<i>2.1 Topics for review</i>	7
<i>2.2 Methodology of the literature review</i>	7
<i>2.3 Results of the review</i>	9
<i>2.4 Interpretation of the review results</i>	11
<i>2.5 Conceptual framework for the present study</i>	16
CHAPTER 3. METHODS	17
<i>3.1 Study population and setting</i>	17
<i>3.2 Identification of the cohorts</i>	18
<i>3.3 Data collection</i>	18

3.4 Risk stratification	18
3.5 Cluster analysis	20
3.6 Contrast mining.....	21
CHAPTER 4. RESULTS	27
4.1 Characteristics of study participants	27
4.2 Validation of the risk stratification tiers	29
4.3 Movement between risk stratification tiers	33
4.4 Inpatient costs by risk stratification tier	36
4.5 Cluster analysis of high-risk patients.....	39
4.6 Contrast mining of high-cost vs. low-cost patients	44
CHAPTER 5. DISCUSSION.....	51
5.1 Primary findings.....	51
5.2 Secondary findings	52
5.3 Implications of study findings	54
5.4 Contributions to informatics	56
5.5 Conclusions	58
APPENDICES	59
Appendix 1. LIGHT ² care coordination protocol by risk stratification tier.....	59
Appendix 2. Individual attributes found in combinations associated with high costs ..	60
BIBLIOGRAPHY	62
VITA.....	77

LIST OF ILLUSTRATIONS

Figure	Page
Figure 1. Actual and Projected Net Medicare Spending, 2010-2026	3
Figure 2. Data flow diagram for literature searches	10
Figure 3. Definitions of risk stratification tiers.....	19
Figure 4. Logarithmic distribution of fiscal year 2013 charges by patient.....	23
Figure 5. Data collection for contrast mining	24
Figure 6. Number of LIGHT ² patients by age and gender	28
Figure 7. Prevalence of monitored chronic conditions in the LIGHT ² population.....	29
Figure 8. Distribution of patients by age and risk tier	30
Figure 9. 2014 healthcare episodes by visit type within 2013 risk tier.....	31
Figure 10. 2014 healthcare charges by 2013 risk tier	32
Figure 11. Movement of a “Healthy” (Tier 1) cohort in the following three years.....	33
Figure 12. Movement of a “Stable” (Tier 2) cohort in the following three years.....	34
Figure 13. Movement of an “Unstable” (Tier 3) cohort in the following three years	34
Figure 14. Movement of a “Complex” (Tier 4) cohort in the following three years	35
Figure 15. Average fiscal 2013 charges per utilization episode by episode type	36
Figure 16. Average charges per admission, US\$.....	38
Figure 17. Average charges per admission by category, US\$	39
Figure 18. Average annual charges by episode type within cluster.....	40
Figure 19. Prevalence of chronic conditions by “Complex” (Tier 4) cluster	41
Figure 20. Receiver operating characteristic curve for the final model.....	49

LIST OF TABLES

Table	Page
Table 1. Comparison groups from patients with lowest non-zero prior-year charges.....	22
Table 2. Average charges per admission by category within risk tier, US\$.....	37
Table 3. Cluster analysis of “Complex” (Tier 4) patients by visit count and length	40
Table 4. Selected attributes of “Complex” (Tier 4) patients by cluster	43
Table 5. Ten (out of 5,188) combinations frequently associated with high costs	45
Table 6. Final model of attributes associated with high costs	48

LIST OF ABBREVIATIONS

AUC: Area Under the Curve

BMI: Body Mass Index

CCW: Chronic Conditions Warehouse

CINAHL: Cumulative Index to Nursing and Allied Health Literature

DBP: Diastolic Blood Pressure

EMR: Electronic Medical Record

GDP: Gross Domestic Product

HbA1c: Hemoglobin A1c

HMG CoA: 3-hydroxy-3-methyl-glutaryl-coenzyme A

ICD9: International Classification of Diseases, 9th revision

LDL: Low-Density Lipoproteins

LIGHT²: Leveraging Information Technology to Guide Hi-Tech and Hi-Touch Care

MeSH: Medical Subject Headings

mg/dL: milligrams per deciliter

mmHg: millimeters of mercury

Rx: prescription

SBP: Systolic Blood Pressure

SSRI: Selective Serotonin Reuptake Inhibitor

SU: subject

UMHS: University of Missouri Health System

US: United States

VIF: Variance Inflation Factor

ABSTRACT

Risk analysis and population health management can improve health outcomes, but improved risk stratification is needed to manage healthcare costs. Analysis of 157 publications on translational implementations of “risk stratification in population health management of chronic disease” showed a consensus that population health management and risk stratification can improve outcomes, but found uncertainty over best methods for risk prediction and controversy over the cost savings. The consensus of another 85 publications on the methodologies of “data mining for predictive healthcare analytics” was that clinically interpretable machine learning techniques are more appropriate than “black box” techniques for structured big data sources in healthcare, and the “area under the curve” of a prediction model’s sensitivity versus one-minus-specificity is a standard and reliable way to measure the model’s discrimination. This study used clinically interpretable machine-learning algorithms, combined with simple but powerful data analytic techniques such as cost analysis and data visualization, to evaluate and improve risk stratification for a managed patient population.

This study retrospectively observed 10,000 mid-Missouri Medicare and Medicaid patients between 2012 and 2014. Cost and utilization analyses, statistical clustering, contrast mining, and logistic regression were used to identify patients within a managed population at risk for higher healthcare costs, demonstrate longitudinal changes in risk stratification, and characterize detailed differences between high-risk and low-risk patients. The two highest risk stratification tiers comprised only 21% of patients but accounted for 43% of prospective charges. Patients in the most expensive sub-cluster of the most expensive risk tier were nearly twice as costly as high-risk patients on average.

Combining contrast mining with logistic regression predicted the most expensive 5% of patients with 84% accuracy, as measured by area under the curve.

All the strategies used in this study, from the simplest to the most sophisticated, produced useful insights. By predicting the small number of patients who will incur the majority of healthcare expenses in terms that are clinically interpretable, these methods can support population health managers in focusing preventive and longitudinal care more effectively. These models, and similar models developed by integrating diverse informatics strategies, could improve health outcomes, delivery, and costs.

CHAPTER 1. INTRODUCTION

1.1 Nature and purpose of the study

The Chronic Care Model (Wagner et al., 2001) proposed improving the effectiveness of interactions between patients and providers as a way of promoting the “Triple Aim” of healthcare (Berwick, Nolan, & Whittington, 2008): (a) better health, (b) better care, and (c) lower costs (Glasgow, Orleans, & Wagner, 2001). These positive outcomes may be anticipated on the basis of improved interactions between informed, activated patients and prepared, proactive providers (Wagner et al., 2001). Evidence also has shown that patients with care coordinators have fewer emergency department and urgent care episodes (Kruse, Zweig, et al., 2010), hospital admissions (Sanderson & Dixon, 2000), and readmissions (Verhaegh et al., 2014).

Care coordination is “... the deliberate organization of patient care activities between two or more participants (including the patient) involved in the patient’s care to facilitate the appropriate delivery of health service...” (McDonald, Vickers, Mohan, Wilkes, & Jackson, 2010). Care coordination is increasingly being used across the healthcare system to improve patient outcomes for populations of patients and is a core element of both the Triple Aim and the Chronic Care Model. In order for care coordinators to manage populations of patients, they must be able to identify the patients who are most in need of their services; one approach to this challenge is risk stratification. Indeed, risk stratification is crucial for effective population health management because it provides care coordinators the opportunity to focus their work on those patients who will benefit the most (McCusker et al., 1999; Suijker, et al., 2012; van

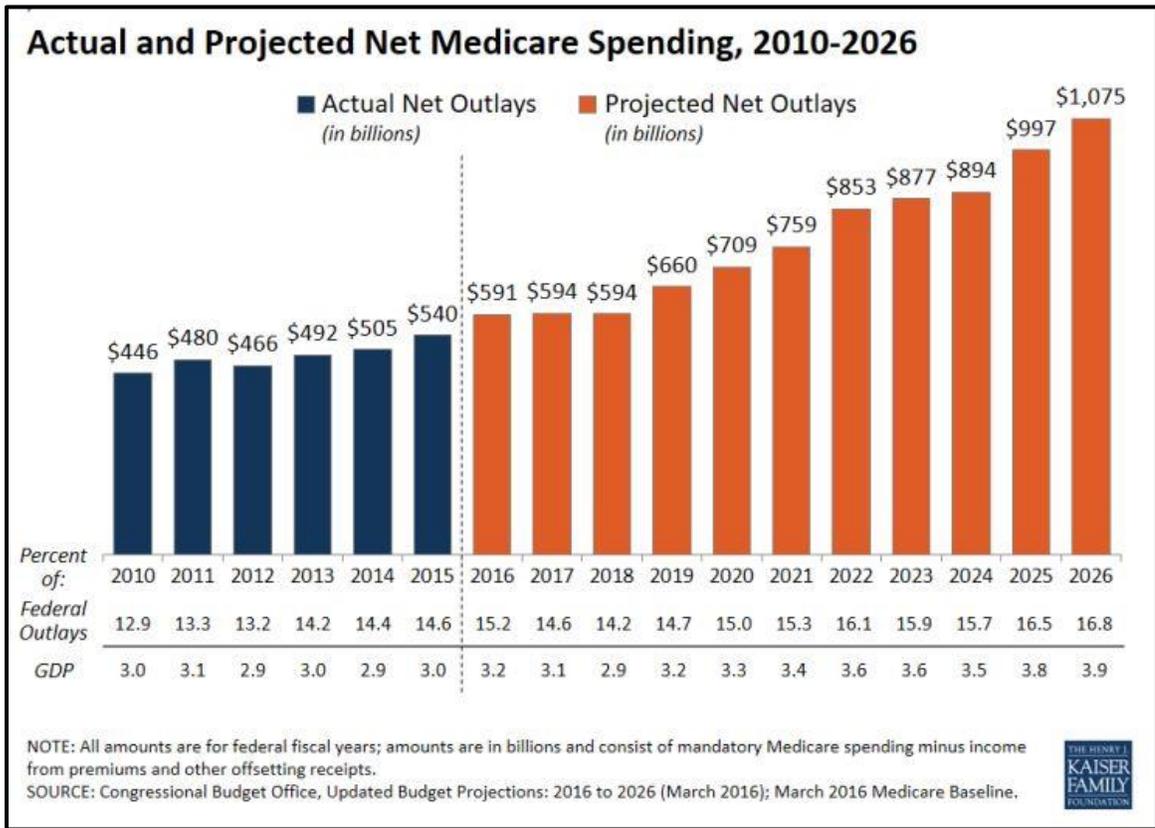
Walraven et al., 2010). Risk stratification is a potentially powerful tool for predicting population health outcomes, and previous healthcare utilization has been shown to be a useful predictor of future healthcare needs (Wherry, Burns, & Leininger, 2014).

However, most predictive indices for health outcomes in community-dwelling adults have less than 70 percent accuracy (O’Caoimha et al., 2015) and relatively few attempts have been made to stratify entire primary care populations. The present study attempts to categorize and predict high-risk members of a managed patient population, using informatics strategies that are simpler, more comprehensive, and more accurate. This introductory chapter will briefly outline the significance of care coordination and the innovation represented by these approaches, and set our three specific aims for their implementation.

1.2 Significance of the study

US healthcare costs continue to rise faster than inflation (Forbes.com LLC, 2015). Figure 1 shows how this “cost curve” continues to rise faster than the US gross domestic product (GDP) and consume more of the US federal budget, putting the long-term solvency of the program at risk (US Centers for Medicare & Medicaid Services, 2013).

Figure 1. Actual and Projected Net Medicare Spending, 2010-2026



(Cubanski & Neuman, 2016)

Half of health care expenses in the United States are incurred by five percent of the population (Cohen & Uberoi, 2013; Stanton & Rutherford, 2005), and 86% of US healthcare resources are consumed in managing chronic diseases (Bodenheimer, Wagner, & Grumbach, 2002; Gerteis, Izrael, Deitz, LeRoy, & Ricciardi, 2014). In order to reduce the public and private burdens of paying for healthcare without reducing quality or access, care delivery must be made more efficient, proactive, and effective (US Congressional Budget Office, 2013).

To achieve the “Triple Aim” of (a) better health outcomes, (b) better health care delivery, and (c) lower costs (Berwick, Nolan, & Whittington, 2008), the Chronic Care

Model seeks to improve interactions between informed, activated patients and prepared, proactive providers (Wagner et al., 2001), including proactive and preventive care (Glasgow, Orleans, & Wagner, 2001). Unfortunately, health care often fails to provide effective coordination of care across a target population (Bodenheimer, Wagner, & Grumbach, 2002; Coleman, Austin, Brach, & Wagner, 2009). When care coordinators don't know which of their patients are at greatest risk for worsening outcomes, they allocate their time by responding to the patient in front of them at the moment (Snyderman & Williams, 2003).

In contrast to less predictable causes of poor health such as traumatic injury, the presence of pre-existing chronic conditions situates patients ideally for proactive intervention, and population health managers are ideally situated to deliver that preventive care if they can focus their work efficiently (Bates, Saria, Ohno-Machado, Shah, & Escobar, 2014). This supports the transition from the traditional “reactive” model of medical care (Snyderman & Williams, 2003) to one of maintaining health and avoiding preventable conditions, and accomplishing the Triple Aim (Amarasingham, Patzer, Huesch, Nguyen, & Xie, 2014). Accomplishing the Triple Aim promises to not only improve health outcomes for healthcare consumers, but improve financial outcomes for those who pay directly or indirectly for healthcare – including everyone who pays taxes, has health insurance, or has a job, savings, or investments.

1.3 Innovative informatics strategies and applications

The number of clinical prediction rules published every decade is expanding at an accelerating rate (Keogh et al., 2014). Electronic health care predictive analytics built on

these clinical predictors can be used to rapidly spot hard-to-identify opportunities to better manage care (Bradley, 2012), but previous predictive analytics based on statistical analysis have had limited usefulness (Kansagara et al., 2011). Part of the reason for this may be limitations in the mathematical tools used to construct the models, because regression analysis and other traditional statistical methods are constrained by the limited number of features that can be used (Kantardzic, 2003). In order to limit the number of model features deductively – that is, before attempting to build and validate a model – some predictive algorithms have focused on specific conditions such as diabetes (Khalid et al., 2014) or hypertension (Sun et al, 2014).

Other predictive analytics have built successful models on demographic and utilization characteristics with only high-level summaries of clinical data (Chechulin, Nazerian, Rais, & Malikov, 2014). However, this strategy fails to exploit the highly detailed clinical history available in electronic medical records (EMR). Data mining algorithms would permit models that surpass these size limits and could leverage the rich data available in the EMR (Witten & Frank, 2005). Advanced informatics strategies, including cluster analysis and data-mining algorithms, are likely to be more appropriate to applying high-dimensional data to this complex problem (He, Mathews, Kalloo, & Huffless, 2014), particularly when designed to produce clinically interpretable results (Fraccaro et al., 2015).

1.4 Objectives of the study

This study tests the hypothesis that clinically interpretable machine-learning algorithms, combined with simpler data analytic techniques where appropriate, can

improve risk stratification in population health management of chronic disease. Three specific aims arise from the application of these innovative informatics strategies to risk stratification in chronic disease management:

1. Develop and validate a reproducible system of identifying patients within a managed population who are at increased risk for higher healthcare costs
2. Demonstrate the longitudinal changes in risk stratification within a cohort of well insured primary care patients
3. Characterize the differences between patients at higher and lower risks of increased healthcare utilization and costs in clinically useful terms

CHAPTER 2. LITERATURE REVIEW

2.1 Topics for review

In order to discover the research literature relevant to both the translational and methodological aspects of this study, two searches were conducted. The first literature review focused on the topic of “risk stratification in population health management of chronic disease” and the second one focused on the topic of “data mining for predictive healthcare analytics”.

2.2 Methodology of the literature review

The first literature review, on the topic of “risk stratification in population health management of chronic disease,” searched two databases of medical evidence, MEDLINE and CINAHL. These two databases were selected as the most widely accepted resources for searching the literatures of medicine and nursing (Brazier & Begley, 1996).

For MEDLINE, which uses the MeSH system of subject headings, the search used this query:

((risk assessment[MeSH term])

AND managed care[MeSH term])

AND chronic disease[MeSH term]

while for CINAHL, which has an expanded set of subject headings including terms closer to the defined search topic, the search used this query:

SU risk stratification

AND SU population health management

AND SU chronic disease

(where SU = subject) in CINAHL's SmartText Searching search mode. Both searches were limited to publications in the English language and to research articles. CINAHL was further restricted to peer reviewed articles flagged for interest to "Case Management," while these filters were not available in the MEDLINE interface. After removing duplicate results from the two databases, the search process manually identified and removed publications which were focused on specific disease conditions such as mental or cardiac health.

For the second literature review, on the topic of "data mining for predictive healthcare analytics," SCOPUS was selected as more appropriate than MEDLINE and CINAHL due to its broader coverage outside biomedical disciplines (Falagas, Pitsouni, Malietzis, & Pappas, 2008). This second search used this query:

"data mining"

AND "clinical prediction"

AND (LIMIT-TO(SUBJAREA,"MEDI"))

(where SUBJAREA = subject area and MEDI = medicine). The search process excluded publications not in the English language and restricted results to research articles, then manually identified and removed publications which were focused on specific disease conditions such as mental or cardiac health.

2.3 Results of the review

The search of the translational literature narrowed the MEDLINE results from 54 to 38 publications and the CINAHL results from 1,555,776 to 119 publications, for a combined total of 157. The search of the methodological literature narrowed the SCOPUS results from 225 to 70 publications. Figure 2 summarizes the results of both search processes.

2.4 Interpretation of the review results

The 157 studies discovered on “risk stratification in population health management of chronic disease” supported consensus on at least three broad findings:

1. *Managed care, preventive care, and disease management programs improve health outcomes.* Specific evidence for this finding included reduced costs (Ahmed & Pearce, 2010; Ahmed, Taylor, McDaniel, & Dyer, 2012; Kruse, Zweig, et al, 2010; Schwartz et al, 2010), utilization (Ahmed & Pearce, 2010; Drozda, Libby, Keiserman, & Rundhaug, 2008; Hamar et al., 2010; Kruse, Zweig, et al, 2010), length of stay (Ahmed & Pearce, 2010; Ahmed, Taylor, McDaniel, & Dyer, 2012), and readmissions (Ahmed & Pearce, 2010; Ahmed, Taylor, McDaniel, & Dyer, 2012); and increased patient function and satisfaction (Ahmed & Pearce, 2010; Ahmed, Taylor, McDaniel, & Dyer, 2012; Bohman et al, 2011; Wiley et al, 2015).
2. *Health outcomes for special populations are improved when management is modified to fit their specific needs.* This finding has been duplicated in disabled (Burns, 2009; Hall, Kurth, & Chapman, 2015), homeless (Levitt et al., 2013; McCormack, Hoffman, Wall, & Goldfrank, 2013; Patterson, Nochajski, & Wu, 2014), minority (Eberly, Davidoff, & Miller, 2010; Smith-Gagen, Loux, Drake, & Perez-Stable, 2016), female (Bierman & Clancy, 2000; Kneipp et al., 2013), pediatric (Morris, Schettine, Roohan, & Gesten, 2011), elderly (Ahmed, Taylor, McDaniel, & Dyer, 2012; Black, 2011; Kwak, Kramer, Lang, & Ledger, 2013), and hospice (Brink & Smith, 2008) populations.

3. *Outcomes can be improved even more when supported by risk assessment and predictive analytics.* Despite the wide variation in risk within populations (Chernew, Weissert, & Hirth, 2001), health risk assessment tools (Drozda, Libby, Keiserman, & Rundhaug, 2008; Krist et al, 2016) and claims-based risk modeling (Drozda, Libby, Keiserman, & Rundhaug, 2008) are reliable predictors of subsequent hospital utilization. Identification, assessment, intervention are particularly critical for high-risk (Boult) and end-of-life (Shmueli, Messika, Zmora, & Oberman, 2010) patients, and even moderate adverse selection places managed care plans at financial risk (Zaslavsky & Buntin, 2002).

In addition to this consensus, there were areas of dispute and lack of knowledge in the “risk stratification in population health management of chronic disease” literature, including these three unresolved questions:

1. *Given that managed care and disease management programs improve health outcomes, do they also reduce costs?* Total costs and savings of care management have been studied in many settings including Medicaid (Greene et al., 2008), low-income (Bohman et al., 2011) and geriatric (Ahmed & Pearce, 2010) populations, commercial insurance coverage (Schwartz et al., 2010), and employee-sponsored programs (LeCheminant & Merrill, 2012; Liu et al., 2013; Loeppke, Edington, & Bég, 2010; Loeppke et al., 2008). While some studies have found cost savings (Ahmed & Pearce, 2010; Schwartz et al., 2010) or improved workplace productivity (Bohman et al., 2011; LeCheminant & Merrill, 2012), others have not (Greene et al., 2008; Loeppke, Edington, & Bég, 2010; Loeppke et al., 2008).

Liu et al. (2013) found a cost increase in the first year of an employee-sponsored program, followed by a decrease in the following years. Greene et al. (2008), reviewing eleven Medicaid managed care interventions across the United States, found one intervention that provided a twelve-fold return on its investment, one with a six-fold return on investment, two others with modest returns, four with savings sufficient to offset their costs, and three with negative returns.

2. *How can population health management be implemented sustainably?* Kruse, Zweig, et al. (2010) pointed out that the established US fee-for-service model provided no mechanism for supporting care coordination. A few studies have examined the impact of private versus public funding for population health management, with one finding improvements in costs and outcomes from retail medical clinics (Rohrer, Angstman, & Bartel, 2009) and others finding that for-profit care management produced worse care (Schmuttermaier, Schmitt, King, & Gwynne, 2011) and worse outcomes (Decker, 2011). Other studies have examined the benefits of case-management interventions variously directed by nurse care coordinators (Kruse, Zweig, et al., 2010; Shin et al., 2010), nurses and physicians in collaboration (Schraeder et al., 2008), or by patients themselves (Laragy & Allen, 2015). Williams & Cooper (2008) examined the optimal caseload sizes for improved outcomes and reduced costs, and others have studied the impacts of innovative approaches such as employee-sponsored care management (Kaspin, Gorman, & Miller, 2013) and the use of online primary-care visits (Rohrer et al., 2010).

3. *What available data are most predictive of health outcomes and costs, and how can risk assessment be improved?* Older investigations of this question included evaluations of data sources, such as administrative records (Coleman et al., 1998) and claims data (Zhao, Ash, Ellis, & Slaughter, 2002). More recently, engaging physicians in risk assessment (Springrose, Friedman, Gumnit, & Schmidt, 2010) has been examined as an additional source of data. Improvements have been documented from including the interactions of independent risk factors such as race and age (Sandberg et al., 2009), age and obesity (Rohrer, Takahashi, & Adamson, 2008), or obesity and comorbidities (Twells, Bridger, Knight, Alaghebandan, & Barrett, 2012). Cucciare & O'Donohue (2006) also recommended rigorous statistical analysis of risk-adjustment models.

The 85 studies discovered on “data mining for predictive healthcare analytics” supported consensus on these two findings:

1. *Machine learning is good for big structured data but clinically interpretable methods (such as logistic regression) are more acceptable in clinical decision support than “black boxes”.* Risk stratification, whether presented as risk-prediction models (Kruppa, Ziegler, & König, 2012), clinical decision support (Fraccaro et al., 2015), or simple risk calculators (Mansmann, Rieger, Strahwald, & Crispin, 2016), can support population health management. Machine-learning methods can offer advantages over classical techniques (Kruppa, Ziegler, & König, 2012; Malley, Kruppa, Dasgupta, Malley, & Ziegler, 2012), but Steyerberg et al. (2012) warn that “black box” machine-learning tools (such as

artificial neural networks and support vector machines) that don't produce interpretable rules are less useful for clinical applications.

2. *Incremental increase in “area under the curve” (AUC) of sensitivity versus one-minus-specificity is a good measure of a model’s discrimination, but clinical usefulness must always be considered in the context of pre-test probability, clinical significance, and the relative costs of over-diagnosis and under-diagnosis.* Although Kruppa, Ziegler, & König (2012) and Steyerberg et al. (2012) have offered alternatives to the AUC method of evaluation, they and others (Mansmann, Rieger, Strahwald, & Crispin, 2016) acknowledge that is a useful and widely-used standard of comparison between risk-prediction models. Like other evaluation methods, however, it is vulnerable to common errors in interpretation when considered out of clinical context (Steyerberg et al., 2012).

In addition to this consensus, there were areas of dispute and lack of knowledge in the “data mining for predictive healthcare analytics” literature, including these two unresolved questions:

1. *Are other measures of discrimination better than “area under the curve”?*

Kruppa, Ziegler, & König (2012) proposed predictiveness curves and the Brier score as useful alternatives to the AUC, while Steyerberg et al. (2012) proposed “net benefit” (the net fraction of true positives gained using the new predictor at a single threshold) and “net reclassification index” (the net fraction of reclassifications in the right direction by making decisions based on predictions with the marker compared to decisions without the marker).

2. *How can big data in healthcare be used and shared without risking patient privacy?* Although specific policy recommendations are rare, the critical needs are for shared data governance and patient privacy (Collins & Wagner, 2005), clinician engagement, patient engagement, and shared decision making (Fraccaro et al., 2015), transparency and reproducibility (Ghassemi, Celi, & Stone, 2015), and interdisciplinary collaboration and workflow integration (Marchevsky & Wick, 2004).

2.5 Conceptual framework for the present study

While the study presented here will include some preliminary cost analyses, it is not able to answer questions about overall costs savings, even if limited to the present study population and time period. Neither can it answer policy and economic questions of sustainably implementing population health management or managing, protecting, and sharing research data. It will not attempt to validate newer measures of predictive discrimination, but will limit itself to the widely used “area under the curve” (AUC).

However, by demonstrating simple and effective informatics strategies for identifying patients within a managed population who are at increased risk, this study does contribute to answering the third question discovered in the literature review: what patient attributes are really predictive of higher healthcare costs and worse healthcare outcomes? And, more fundamentally, what informatics strategies can be used to reliably and interpretably discover those attributes among the wealth of data available in electronic medical records?

CHAPTER 3. METHODS

3.1 Study population and setting

In 2012, the Centers for Medicare and Medicaid Services awarded \$13.3 million to the School of Medicine at the University of Missouri for LIGHT² (Leveraging Information Technology to Guide Hi-Tech and Hi-Touch Care), a three year pilot program to examine the use of advanced health information technology and care coordination in a managed population. The LIGHT² program recruited more than 10,000 primary care patients at the University of Missouri Health System who were already enrolled in Medicare or Medicaid, and hired over 20 experienced registered nurses to manage the patients' health care between doctor's visits. The premise explored by LIGHT² was that a combination of advanced information technology and comprehensive health care coordination could improve the health outcomes of Medicare and Medicaid beneficiaries and improve health care delivery, while lowering costs by reducing the number of emergency visits and hospitalizations.

Adult primary care in the University of Missouri Health System (UMHS) is provided by approximately 133 primary care physicians who practice in nine regional clinics; the Department of Family and Community Medicine operates five local outpatient clinics and two in nearby communities, and the General Internal Medicine section of the Department of Medicine operates two local clinics. This community-based, primary care focus is supported by an extensive UMHS tertiary-care system of six hospitals and more than 50 clinics, staffed by approximately 550 university physicians.

3.2 Identification of the cohorts

The LIGHT² study enrolled all participants between early 2013 and early 2015. Because additional patients were enrolled after the beginning of the intervention period, and some patients withdrew or were otherwise lost to followup, the total number of patients in the study varied slightly between analyses conducted at different times. In addition, because some study cohorts were defined by risk characteristics that changed over time, the total number of patients in each cohort varied slightly between analyses conducted at different times.

3.3 Data collection

Diagnoses and utilization histories for all LIGHT² patients between 1 January 2012 and 31 December 2014 were collected retrospectively from the UMHS Health Analytics Library, an analytic data mart created for the LIGHT² study to reflect the relevant contents of the UMHS electronic health records system.

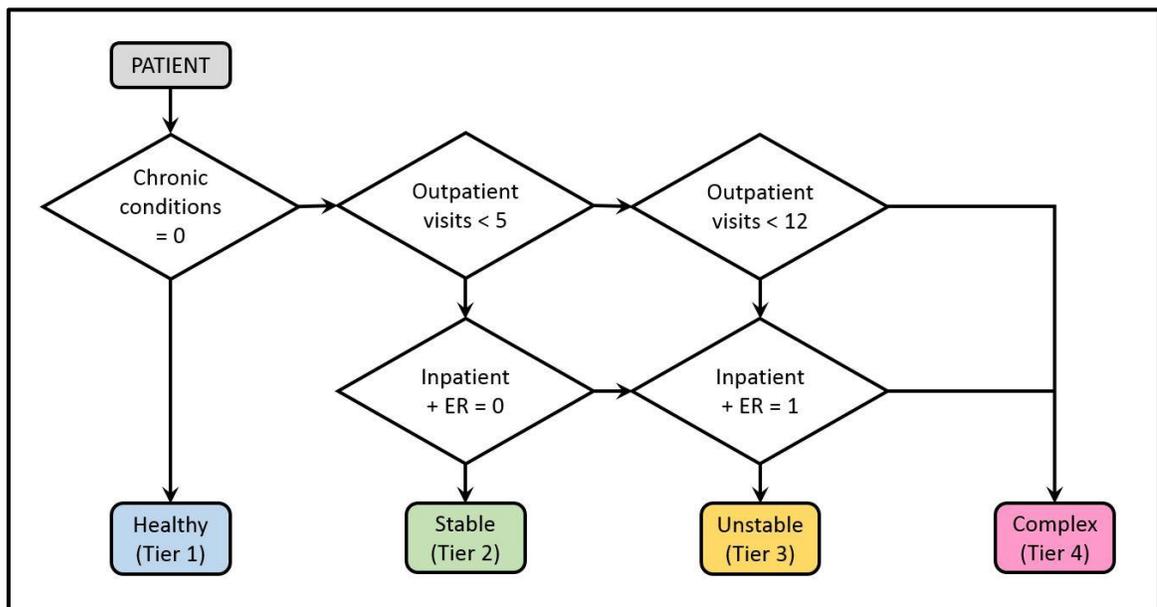
In order to compare costs for different years, UMHS healthcare charges were adjusted to 2014 dollars. Because UMHS charges were adjusted upward by 3% on April 1 of each year during the study period, charges billed from April 1, 2013, to March 31, 2014, were adjusted by multiplying times 1.03. Charges billed from October 1, 2012, to March 31, 2013, were multiplied by (1.03×1.03) or 1.061.

3.4 Risk stratification

Four risk stratification tiers were defined as shown in Figure 3 (Popejoy et al., 2015). Patients who had none of the 27 chronic conditions identified in the Chronic

Conditions Warehouse, or CCW (Goodman, Posner, Huang, Parekh, & Koh, 2013), constituted the lowest-risk or “Healthy” category (Tier 1). Patients with one or more chronic conditions were placed in higher tiers depending on the number of their outpatient clinic visits and hospital episodes during the year prior to analysis. Patients with five to twelve related outpatient visits or one related hospital episode in a year were defined as “Unstable” (Tier 3), based on the investigators’ clinical judgement. Patients with chronic conditions but fewer visits and episodes than “Unstable” patients were defined as “Stable” (Tier 2), and those with more visits or more episodes than “Unstable” patients were defined as “Complex” (Tier 4).

Figure 3. Definitions of risk stratification tiers



In order to visualize the longitudinal stability of the four risk tiers in the study population, cohorts were retrospectively identified by the risk tier of each patient as of January 1st, 2012. The risk tiers were then retrospectively recalculated bi-weekly for each

patient in the cohort over the following three years. The percentage of each cohort in each tier, as well as the percentages deceased or lost to followup, were calculated and displayed as stacked bar charts. The use of these risk tiers in guiding care coordination for the LIGHT² intervention is outlined in Appendix 1. *LIGHT2 care coordination protocol by risk stratification tier*

In order to analyze the relationships between risk tiers and inpatient utilization charges, the hospital charges for inpatient admissions of patients in each tier were summed for the fiscal years ending in 2013 and 2014. Only admissions for which the primary diagnosis was one of the 27 chronic conditions in the CCW were included in the sum of charges. For each risk tier in each fiscal year, the charges for all chronic-condition-related hospital admissions were averaged, in total and by billing category.

3.5 Cluster analysis

In order to refine the utilization patterns of “Complex” (Tier 4) patients, the study identified 343 patients in this risk tier as of October 1st, 2014 and clustered them by six measures from the previous year: number of outpatient visits, number of emergency visits that did not lead to an inpatient admission, and number of observation stays, number of inpatient admissions, total length of inpatient stays, and number of unplanned readmissions within 30 days. Change in the log-likelihood function was used as the distance measure (Chiu, Fang, Chen, Wang, & Jeris, 2001). The two-step procedure first performs a coarse clustering of the original data resulting in a large number of initial sub-clusters, then combines sub-clusters when the distance between them decreases the log-likelihood function.

Following clustering of the 343 “Complex” (Tier 4) patients into three groups, the study gathered descriptive statistics for patient age, race, gender, marital status, insurance status, and annual hospital charges; and the study measured outcomes in the previous year for medical diagnoses, maximum hemoglobin A1c (HbA1c), maximum systolic blood pressure (SBP) and diastolic blood pressure (DBP), maximum low-density lipoprotein cholesterol (LDL), maximum body mass index (BMI), and screening results for a history of falling, smoking, or hazardous alcohol use.

3.6 Contrast mining

Mining data to contrast two or more conditions, or contrast mining [20], requires comparison groups from comparable populations. Other data reduction techniques such as principal components are less than ideal for several reasons: they do not make explicit use of the known-groups nature of the problem, are not well suited to binary data, and would be computationally impractical with the large number of characteristics considered here. Furthermore, both principal component analysis and factor analysis aim at finding linear combinations of features as opposed to identifying individual features that best discriminate between groups.

This application of contrast mining used multiple comparison groups in order to test the flexibility and robustness of the methodology under varying input conditions. Hospital and clinic charges were selected as the outcome of interest for this study because they are easily measured, continuously distributed, and can be compared comprehensibly between diverse patients or populations. Contrasting the 5% of patients who incurred the highest costs with varying subsets of the other 95% of patients tested whether this

contrast-mining strategy would produce widely divergent results depending on slight variations in the comparison group, or the whether the results would instead be robust and reproducible regardless of those initial conditions.

The study excluded patients with zero healthcare system charges on the grounds that individuals with no recent hospital or outpatient visits may not have current medical histories in the healthcare system EHR. Therefore, the comparison groups comprised each of the lowest non-zero 5%, 10%, 20%, 30%, 40%, and 50% of fiscal year 2013 charges. The definitions of these comparison groups are summarized in Table 1.

Table 1. Comparison groups from patients with lowest non-zero prior-year charges

	5%	10%	15%	20%	30%	40%	50%
Lowest charge in range	<- ----- \$27 ----- >						
Highest charge in range	\$470	\$853	\$1,221	\$1,621	\$2,646	\$4,300	\$6,963
Percentage of all charges	<0.1%	0.2%	0.5%	0.8%	1.8%	3.5%	6.1%

Because contrast mining with the remaining 95% of patients would bias the model to non-high-charge utilizers, matched-size comparison groups were defined by random selection from each of the lowest non-zero comparison groups. Figure 4 shows the distribution of the excluded zero-cost patients and the lowest and highest non-zero charges by patient for that fiscal year.

Figure 4. Logarithmic distribution of fiscal year 2013 charges by patient

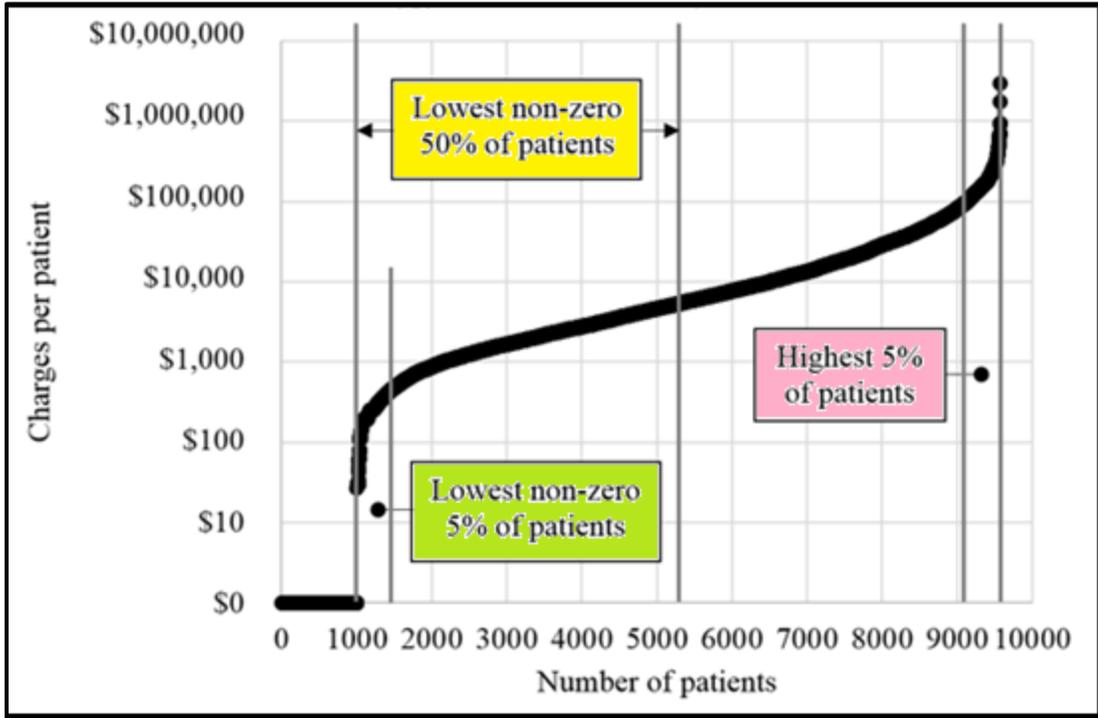
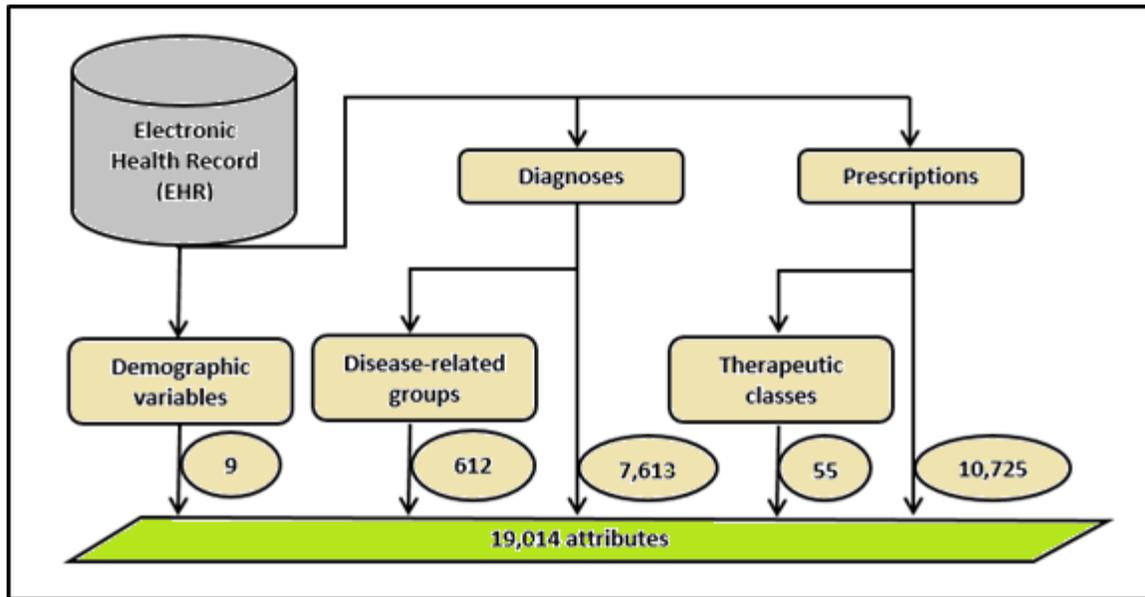


Figure 5 shows the data collected for patients in these groups, including nine demographic attributes: age, gender, race/ethnicity, marital status, English fluency, Medicaid coverage, high prior-year (fiscal year 2012) costs, body mass index (BMI), and history of adherence to prescription instructions. The study additionally categorized the 3,615 ICD9 codes found for these patients into 612 disease-related groups, and the 10,725 prescriptions into 55 higher-level therapeutic classes. All 19,014 attributes were collected for the selected patients at the end of fiscal year 2012, prior to the fiscal year 2013 outcome of interest.

Figure 5. Data collection for contrast mining



In order to process contrast mining algorithms, a distributed association-rule mining tool suite was built on Apache Spark in the Hadoop Distributed File System (Shvachko, Kuang, Radia, & Chansler, 2010). Because contrast mining algorithms require binary values, the study transformed all attributes to true-or-false flags after converting continuous variables to standard categories (US Centers for Disease Control and Prevention, 2016). For example, each categorical variable (i.e., race/ethnicity and marital-status) was transformed to a set of binary values: (a) “race/ethnicity=white-non-Hispanic or not, =Hispanic or not, =African-American or not, =Asian or not, =Native-American or not, =other or not, =unknown or not,” and (b) “marital-status=single or not, =married or not, =divorced or not, =widowed or not.” Gender, a categorical variable with only two values in this population, was not transformed. The two clinical alert flags (i.e., English fluency and prescription adherence) were transformed to “normal” or “abnormal.” The “Apriori” algorithm (Agrawal & Srikant, 1994) was then executed with

a support of 0.2 (i.e., excluding attribute combinations found in less than 20% of patients), and the results were limited to attribute combinations that included the outcome of interest (fiscal year 2013 charges in the top 5% for that year).

In the second step, the algorithm dissected the attribute combinations found frequently (20% or more) in high-cost patients and infrequently in low-cost patients into individual attributes, and treated these contrasting attributes as hypotheses to be tested with multiple regression. Forward selection with $p < 0.05$ was the entry criterion to add attributes to a simplified regression model for each comparison group. Interaction terms were not included. Because the dependent variable was expressed as a binary classifier (high vs. low utilization), the study used logistic regression (Cox, 1958) to construct the risk prediction model. Each candidate predictor generated a Variance Inflation Factor (VIF) resulting from the regression of that variable on the other candidate predictors. None of the VIF values exceed 3.8, substantially less than the standard rule of thumb that a VIF of 10.0 or greater signals instability in the regression coefficients (Myers, 1990). In addition, the study examined influence plots from the final model to see if individual cases exerted extreme influence on the regression coefficients, identifying no remarkable observations.

The discrimination of the resulting prediction was evaluated by testing the predicted outcome against the actual outcome (fiscal year 2013 charges over \$94,895 or not) for the entire study population of 9,581 patients. Discrimination was defined as the c-statistic, or the area under the receiver operating characteristic curve of sensitivity versus one-minus-specificity (Hanley & McNeil, 1982). Each comparison group (lowest non-zero 5%, 10%, 15%, 20%, 30%, 40%, and 50%) was contrast-mined independently

against the 5% of patients with highest fiscal year 2013 charges, and the resulting models were tested independently. The attributes common to all these models also were used to derive a combined model using all fiscal year 2013 observations, which was also tested independently.

CHAPTER 4. RESULTS

4.1 Characteristics of study participants

As shown in Figure 6, 42% of LIGHT² study participants were younger than 65 years, with ages distributed approximately evenly between 20 and 64 years. The ages of the remaining participants at enrollment tapered from a peak of more than 300 participants of age 65 years (more than 3% of the entire study population) to very few with ages over 100 years. At all ages, females outnumbered males in the study population by about 3:2. The peak at age 65 years is explained by the study inclusion criteria of Medicare and/or Medicaid enrollment, because relatively few working-age US adults are enrolled in Medicaid (Coleman A. , 2014) or Medicare (Centers for Medicare & Medicaid Services, 2014), but most retirement-age US adults are enrolled in Medicare (Centers for Medicare & Medicaid Services, 2014).

Figure 6. Number of LIGHT² patients by age and gender

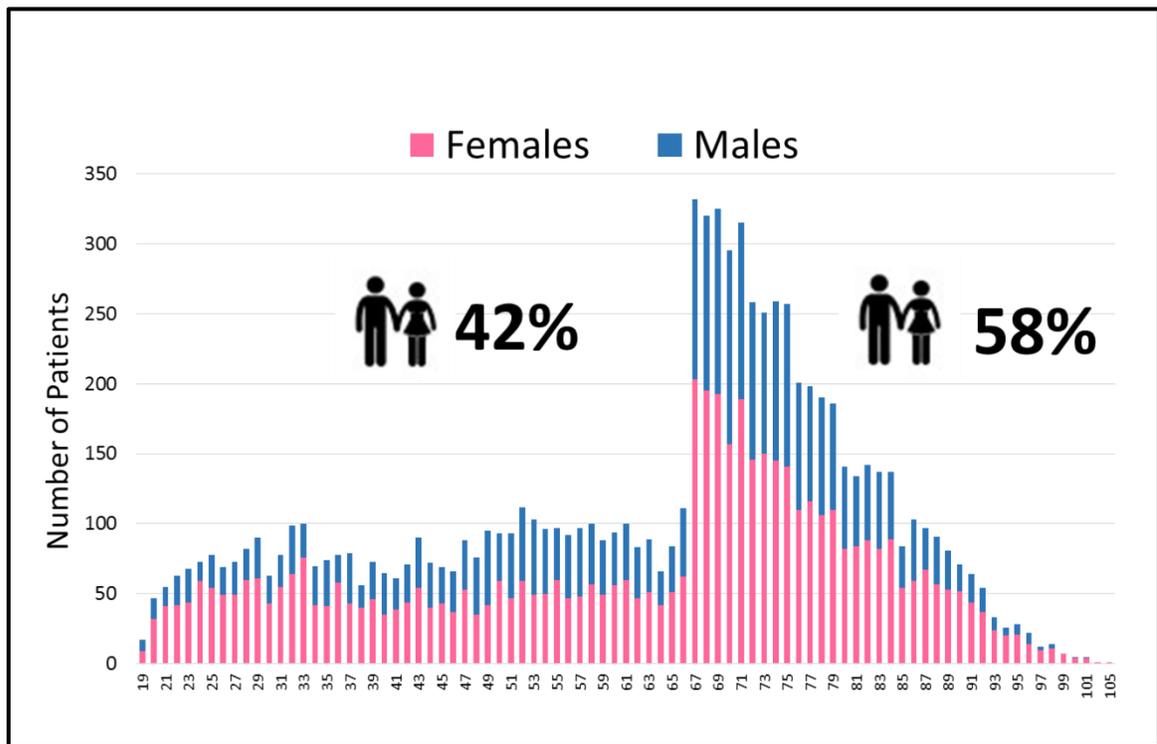
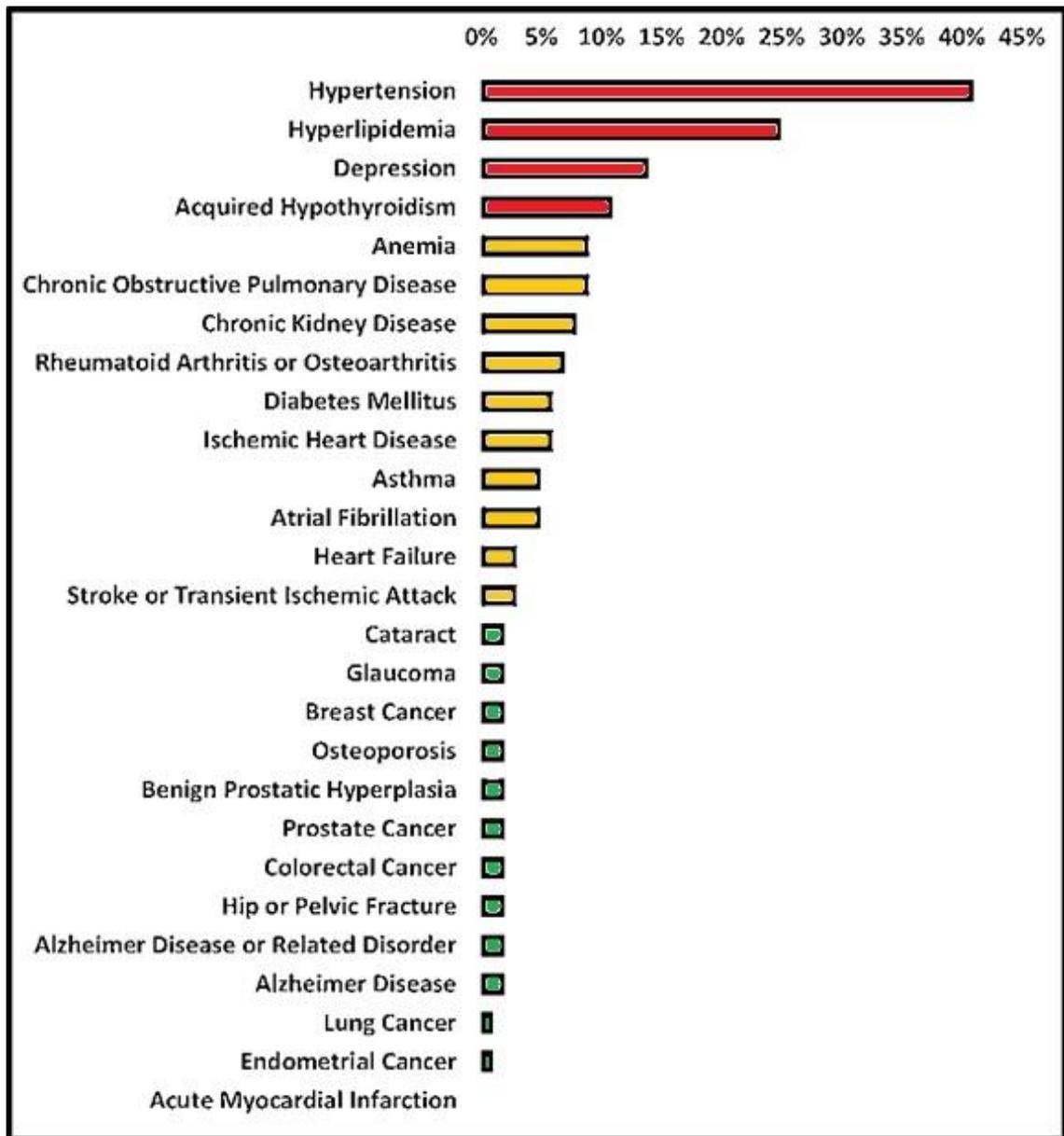


Figure 7 shows the prevalence of the 27 CCW-defined chronic diseases (Goodman, Posner, Huang, Parekh, & Koh, 2013) in the study population. These ranged from hypertension, with a prevalence of 41%, to acute myocardial infarction, which was not found in the study population.

Figure 7. Prevalence of monitored chronic conditions in the LIGHT² population



4.2 Validation of the risk stratification tiers

Of 9,581 patients in the tier validation cohort on October 1, 2013, 63% (n = 6,014) were in Tier 2 while Tiers 1 and 3 comprised 16% each (n = 1,554 for Tier 1; n = 1,555 for Tier 3), with the remaining 5% (n = 458) in Tier 4. More than three-fourths of

the Tier 1 enrollees were younger than 65, but fewer than half of the enrollees in the other tiers were under 65. Figure 8 shows this distribution of the risk stratification tiers.

Figure 8. Distribution of patients by age and risk tier

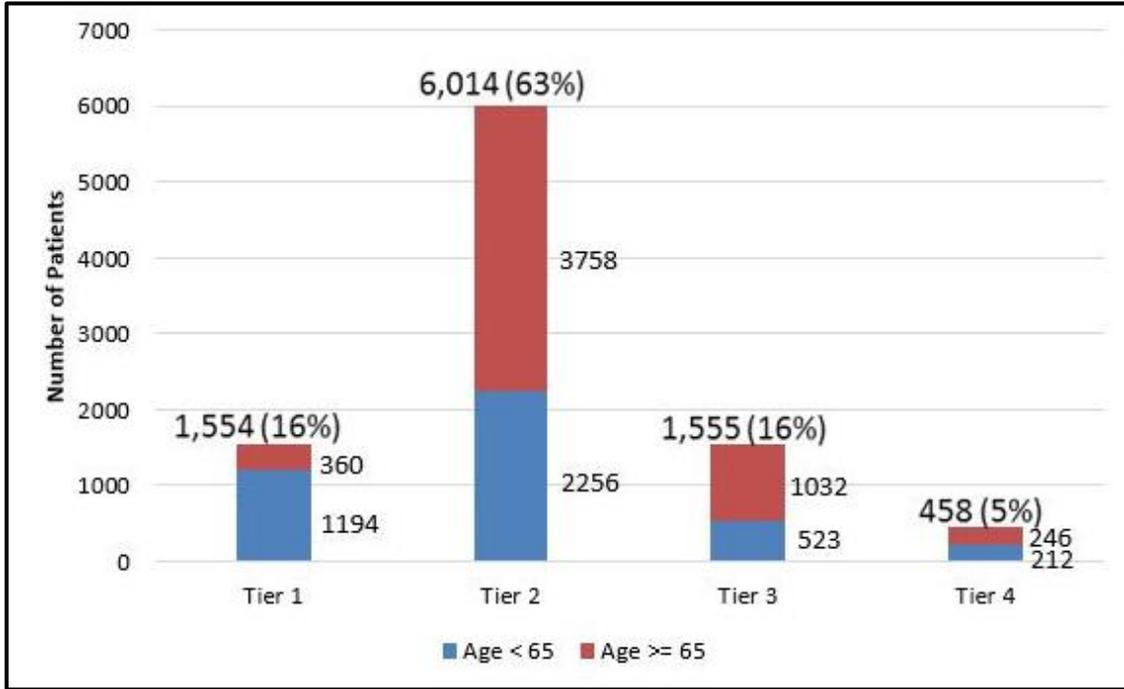


Figure 9 shows the mean number of hospital episodes by type, within each tier, during the 12 months following cohort formation. For all episode types, the overall differences remained significant ($p < 0.001$). “Complex” (Tier 4) patients had significantly more episodes on average than “Unstable” (Tier 3) patients; they, in turn, had significantly more episodes than “Stable” (Tier 2) or “Healthy” (Tier 1) patients ($p < 0.001$ for all comparisons). There were no significant differences between Tier 1 and Tier 2 emergency episodes ($p = 0.279$), but the numbers of observation and inpatient episodes were significantly different between these two tiers ($p < 0.001$ for each type). For all tiers, the median number of episodes of all types was zero.

Figure 9. 2014 healthcare episodes by visit type within 2013 risk tier

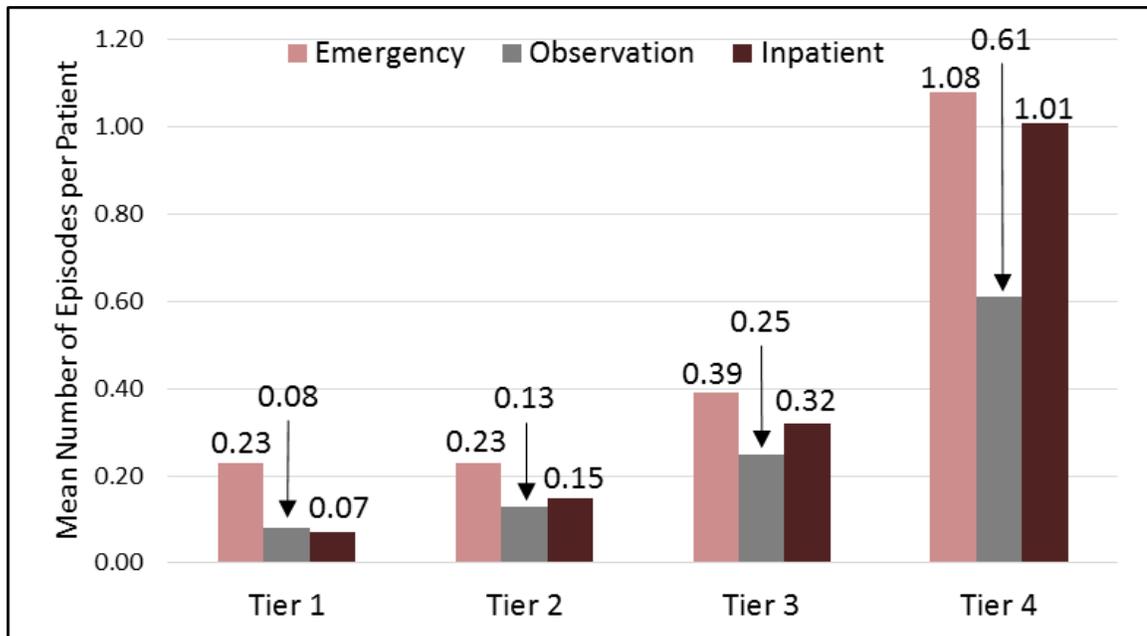
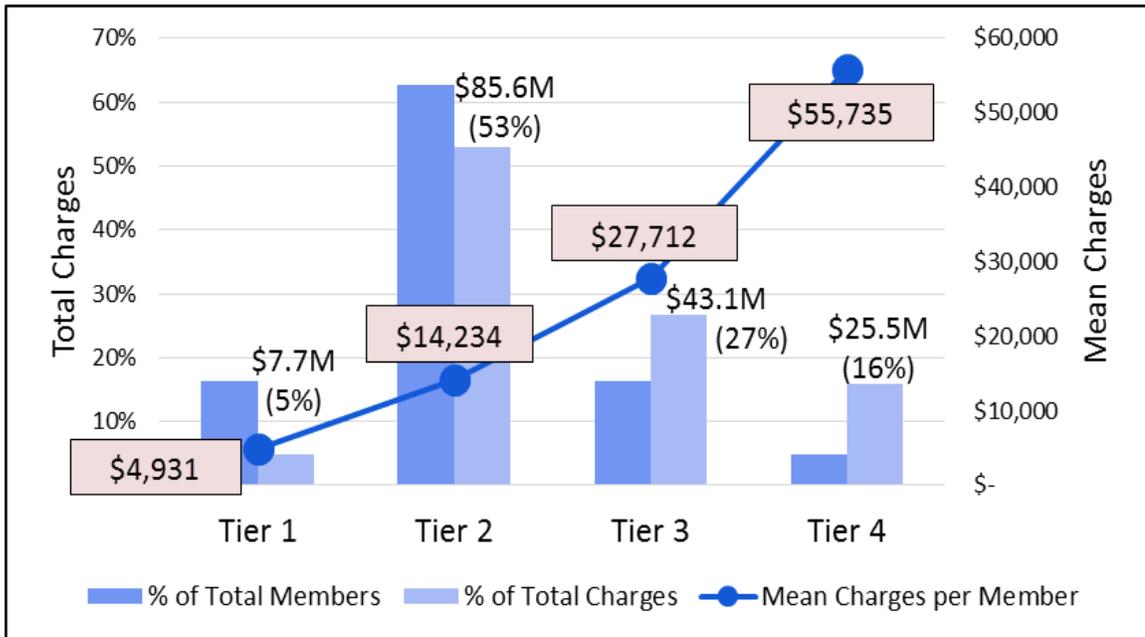


Figure 10 shows the mean and total healthcare charges by tier during the 12 months after cohort formation. “Complex” and “Unstable” patients (Tiers 3 and 4), comprising 21% of the total population, accounted for 43% of total healthcare charges. Overall differences were significant ($p < 0.001$). “Complex” (Tier 4) patients had significantly higher charges on average than “Unstable” (Tier 3) patients, which had significantly higher charges than “Stable” (Tier 2) patients; and these in turn were significantly higher than for “Healthy” (Tier 1) patients ($p < 0.001$ for all comparisons). Median charges in each tier (respectively \$0, \$2,343, \$8,662, and \$20,412) were likewise significantly different overall ($p < 0.001$) and higher in the higher tiers ($p < 0.001$ for all comparisons).

Figure 10. 2014 healthcare charges by 2013 risk tier



The association of higher 2014 utilization and costs with risk tiers based on 2013 utilization supported a hypothesis that risk stratification for individuals would be relatively stable over time. However, the predominance of over-65 patients in all but the lowest risk tier, and the prevalence of chronic conditions in study population, led to a hypothesis that this aging and well insured population would gradually rise in risk of healthcare utilization over the course of a few years. The results of the longitudinal tier movement studies were used to test these secondary hypotheses.

4.3 Movement between risk stratification tiers

“Healthy” (Tier 1) patients had a 56% chance of staying in the same risk tier, and a 27% chance of moving to the next higher tier, as shown in Figure 11. “Stable” (Tier 2) patients, shown in

Figure 12, had a 68% chance of staying in in the same risk tier, and a 13% chance of moving to the next higher tier. “Unstable” (Tier 3) patients had a 48% chance of moving to the next lower tier, and a 26% chance of staying in the same risk tier, as shown in Figure 13. And as shown in Figure 14, “Complex” (Tier 4) patients had a 37% chance of moving to a much lower risk tier, a 25% chance of moving to the next lower tier, and an 18% chance of remaining in the same tier.

Figure 11. Movement of a “Healthy” (Tier 1) cohort in the following three years

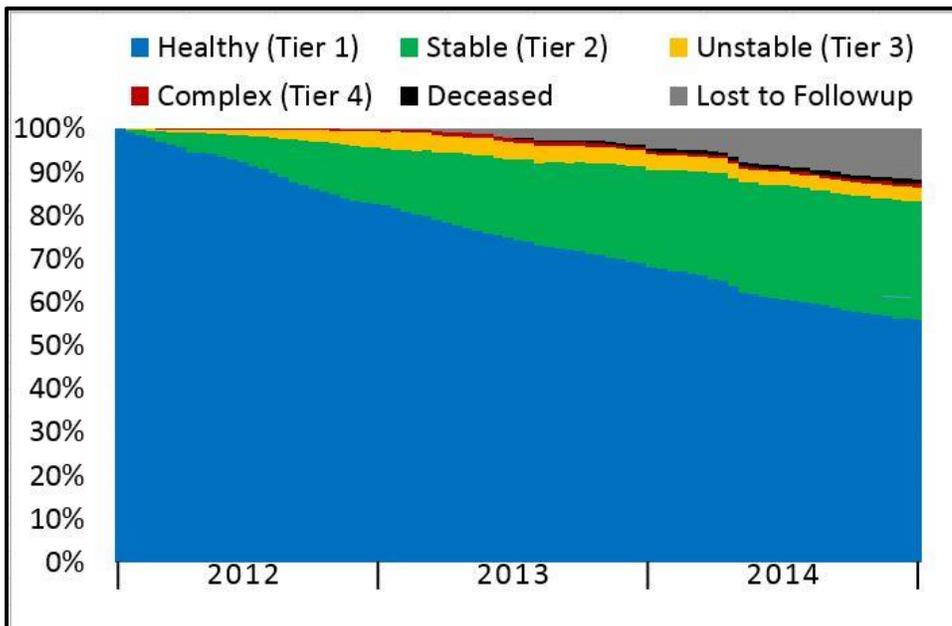


Figure 12. Movement of a “Stable” (Tier 2) cohort in the following three years

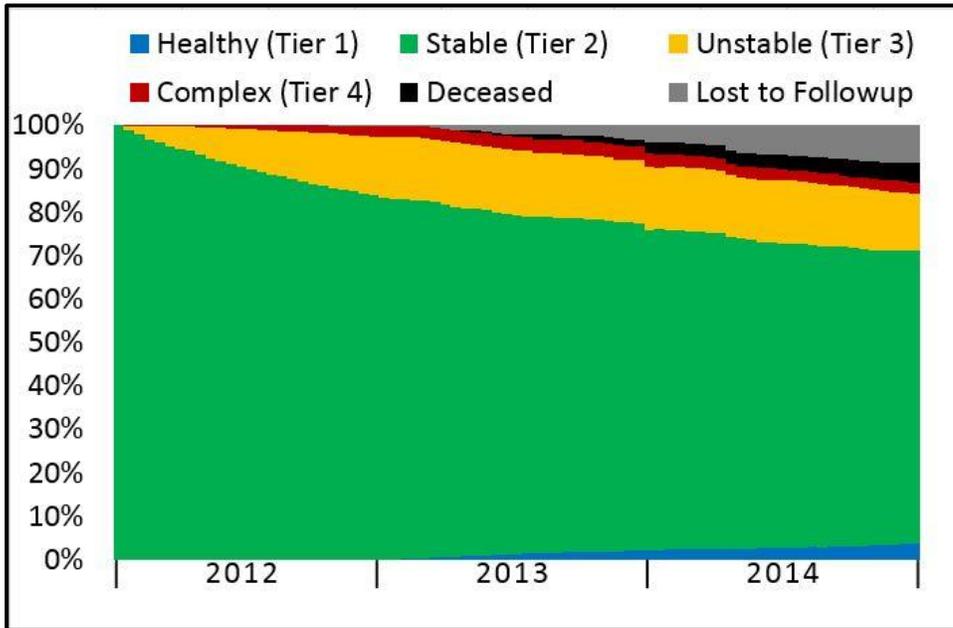


Figure 13. Movement of an “Unstable” (Tier 3) cohort in the following three years

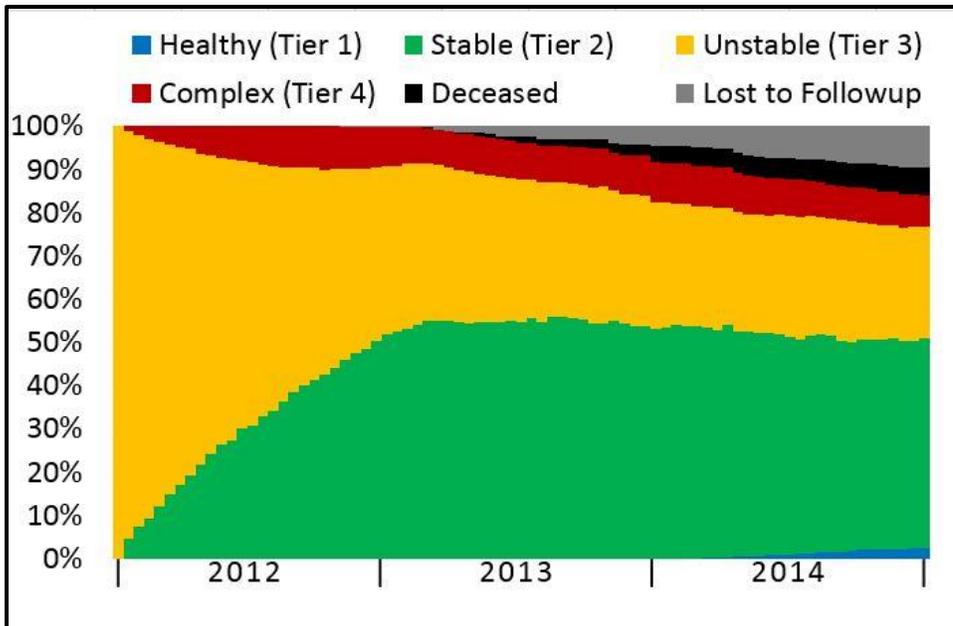
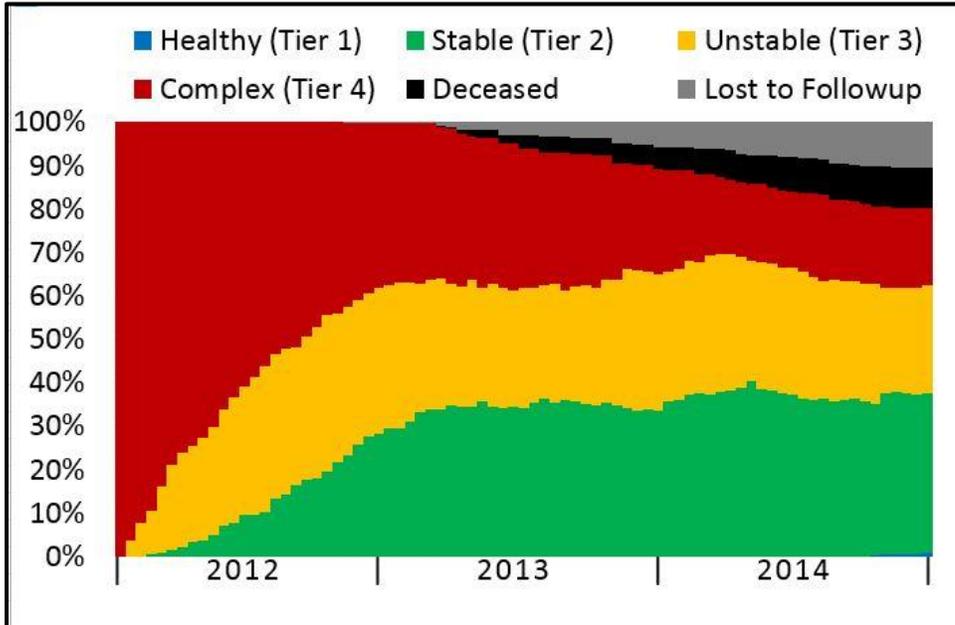


Figure 14. Movement of a “Complex” (Tier 4) cohort in the following three years



As shown in Figure 11 through Figure 14 (above), the rate of death increased steadily in higher risk tiers, at 1%, 4%, 7%, and 9% respectively; and the loss to followup was 12%, 9%, 10%, and 11% respectively.

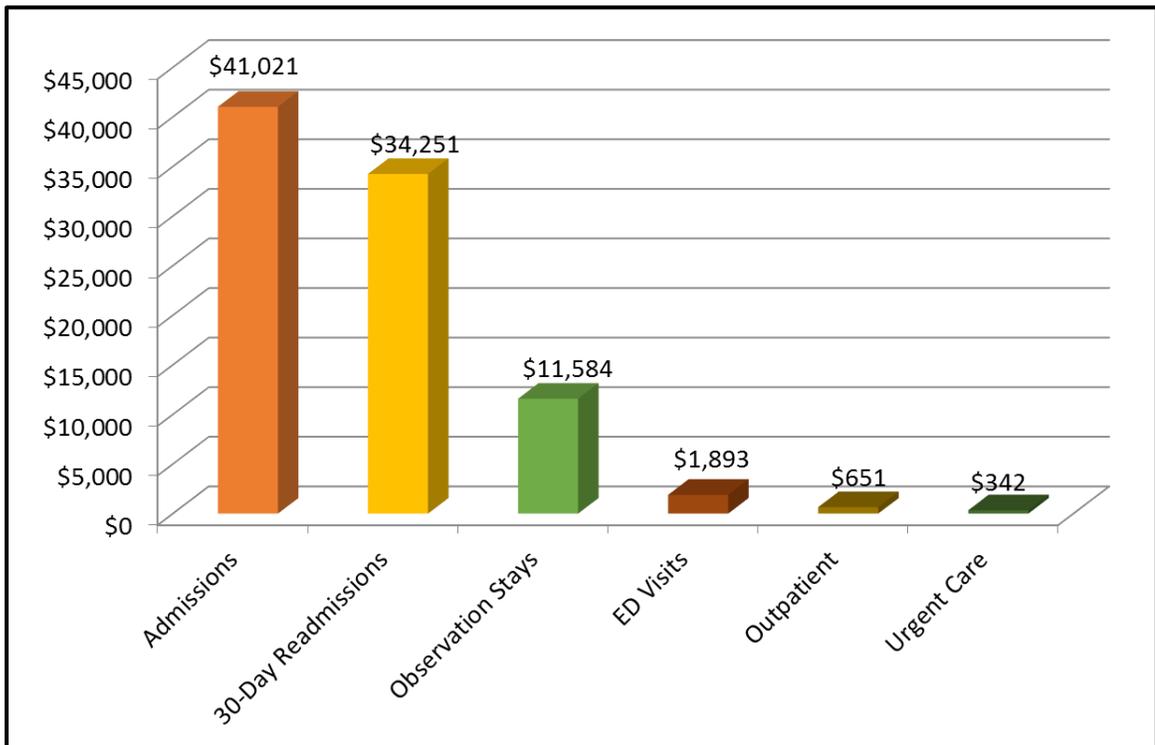
Although risk tiers based on chronic disease diagnoses and recent healthcare utilization were predictive of healthcare utilization and charges in a managed population, patients in this population were not likely to rise in risk over the course of the study as expected. Instead, patients in lower risk tiers (“Healthy” and “Stable” patients) at the beginning of the study period were most likely to remain in lower risk tiers, and those in higher initial risk tiers (“Unstable” and “Complex” patients) were most likely to move to lower risk tiers. In a time frame of three years, this return to stability was a more important influence on healthcare utilization than risk or aging. Because these unexpected findings highlighted the fact that the correlation of risk tiers with specific high-cost outcomes was unknown, more detailed analyses were performed to test the hypothesis

that similar utilization episodes for higher-risk patients were not only more frequent (as required by the definitions of the risk tiers) but also more expensive.

4.4 Inpatient costs by risk stratification tier

Analysis of charges by charges by risk tier focused on inpatient admissions because these episodes are more than 20 times as costly as emergency visits and more than 60 times as costly as outpatient visits, as illustrated in Figure 15.

Figure 15. Average fiscal 2013 charges per utilization episode by episode type



In both fiscal years ending in 2013 and 2014, average charges for chronic-condition-related inpatient admissions of “Complex” (Tier 4) patients were near the

lowest for all risk tiers, at US \$10,046 and \$10,123 for the two years, as shown in Table 2.

Table 2. Average charges per admission by category within risk tier, US\$

Fiscal Year	Surgery	Supplies	Pharmacy	Other	Nursing	Lab	Implants	Imaging	Emergency	Clinic	Cardiology	Total
Tier 1												
2013	1,004	626	544	460	3,808	374	603	236	173	8	324	8,162
2014	833	623	502	150	2,671	428	330	118	266	4	1,241	7,164
Tier 2												
2013	1,476	914	905	583	3,863	747	2,382	407	307	15	343	11,954
2014	1,769	1,075	1,185	869	5,144	839	2,747	420	270	12	505	14,836
Tier 3												
2013	1,533	1,063	1,110	818	4,919	755	2,468	387	267	14	299	13,633
2014	1,539	944	862	739	4,503	712	2,183	323	282	18	313	12,423
Tier 4												
2013	764	575	1,076	754	4,394	854	615	379	323	27	285	10,046
2014	856	472	881	665	4,468	909	937	336	357	21	222	10,123

The only lower average charges were for “Healthy” (Tier 1) patients at \$8,162 and \$7,164. The highest average charges were for “Stable” (Tier 2) patients at \$11,954 and \$14,835, and “Unstable” (Tier 3) patients at \$13,633 and \$12,423, as shown in Table 2 (above) and Figure 16.

Figure 16. Average charges per admission, US\$

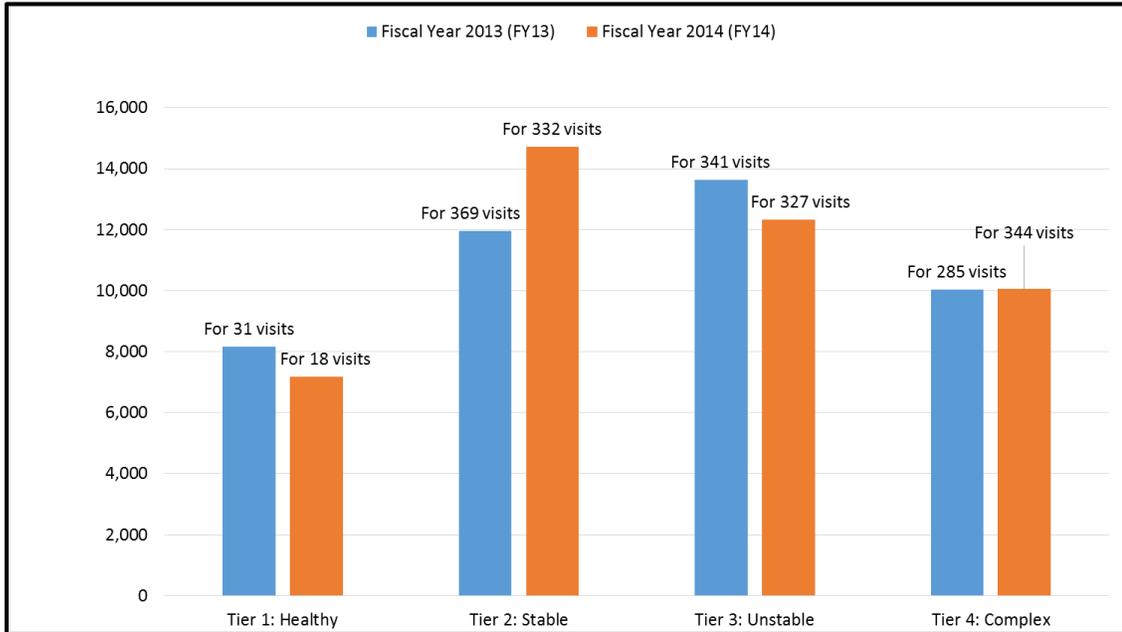
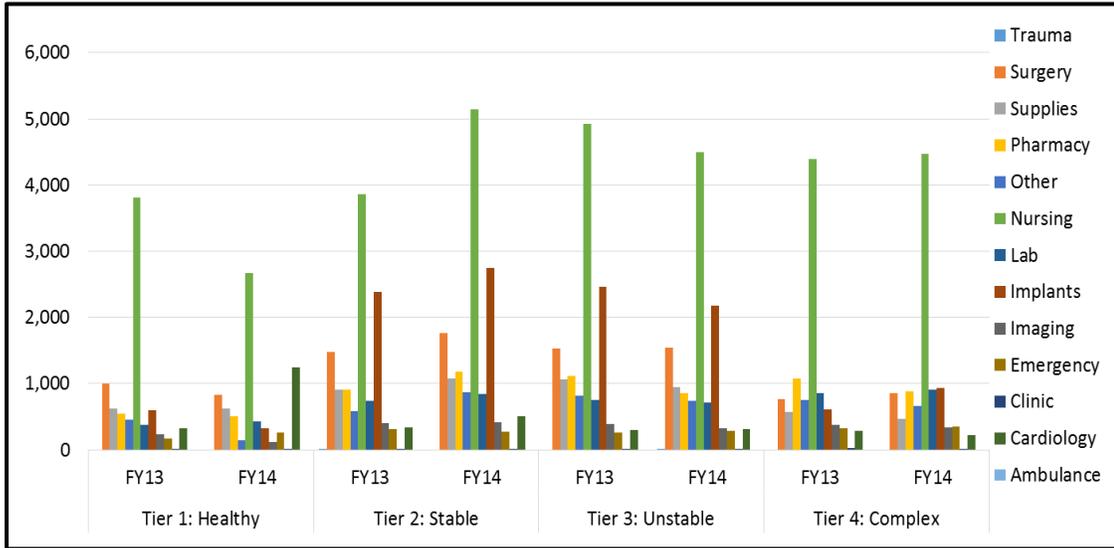


Table 2 (above) and Figure 17 subtotal these charges by thirteen billing categories. In all risk tiers for both fiscal years, nursing is the largest single billing category. However, the difference in average nursing charges for “Complex” patients (\$4,431) and Tiers 2 and 3 patients (\$4,605 combined average) is only 4%. Charges in the implants, surgery, and supplies billing categories accounted for 93% of the difference between “Complex” patient charges and charges for “Stable” and “Unstable” patients, as shown in Figure 17.

Figure 17. Average charges per admission by category, US\$



Initial validation of the tier definitions showed that “Unstable” (Tier 3) and “Complex” (Tier 4) risk stratifications were associated with significantly higher utilization, but detailed longitudinal and cost analyses disproved the study’s initial assumptions that high risk would persist or rise over time and lead to more expensive utilization episodes. In order to resolve these unexpected findings, more sophisticated data science techniques were used to characterize the “Complex” patients, including cluster analysis and contrast mining.

4.5 Cluster analysis of high-risk patients

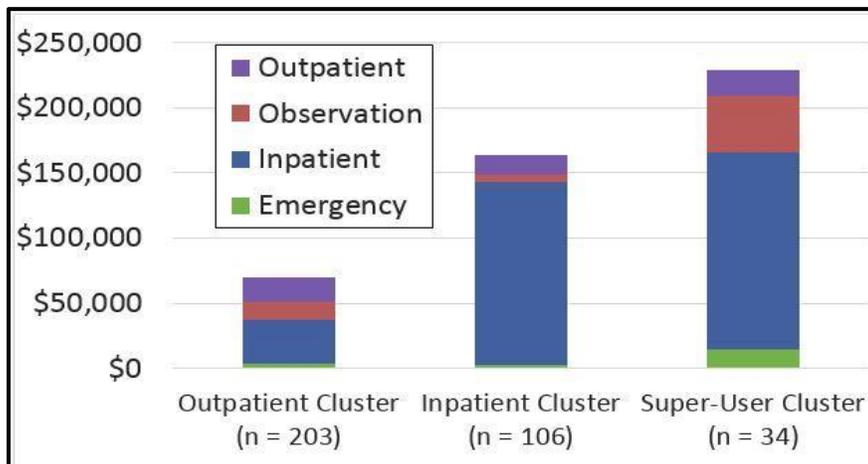
Clustering of the 343 patients with “Complex” (Tier 4) utilization histories found three distinct patterns, as shown in Table 3: frequent outpatient visits for 59% of the group, frequent inpatient admissions for another 31%, and frequent utilization of all types except outpatient visits for the remaining 10% of “Super-Users”.

Table 3. Cluster analysis of “Complex” (Tier 4) patients by visit count and length

	All Tier 4	Cluster 1	Cluster 2	Cluster 3
Average number of visits by type	n = 343	n = 203	n = 106	n = 34
(or Length of stay in days)		(59%)	(31%)	(10%)
Outpatient visits	8.76	10.75	5.67	6.44
Emergency visits	0.62	0.61	0.24	1.82
Observation stays	0.56	0.58	1.17	1.71
Inpatient admissions	1.12	0.50	1.94	2.26
(Inpatient length of stay, days)	8.70	3.28	15.55	19.74
Unplanned 30-day readmissions	0.48	0.05	0.61	2.56

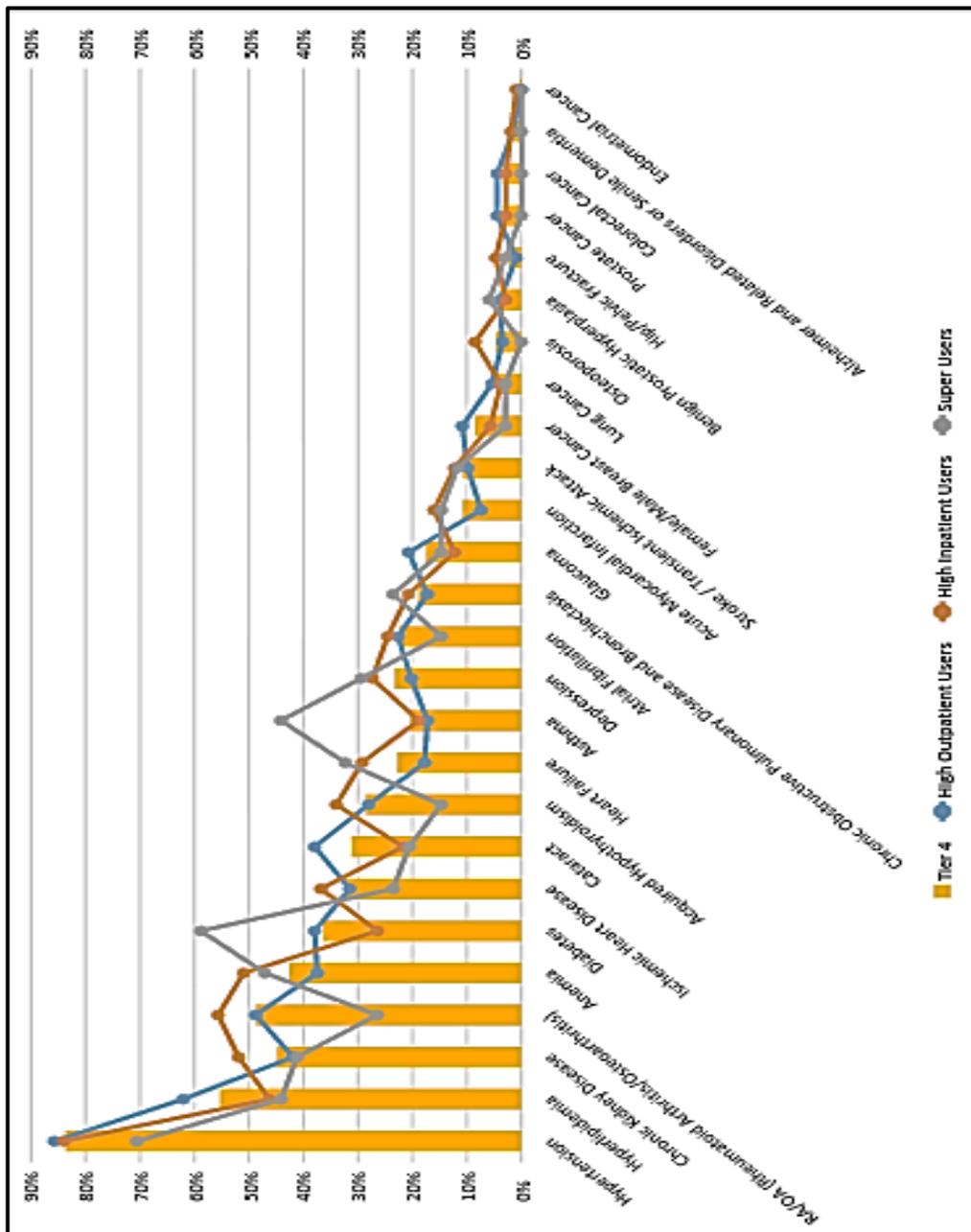
Average annual healthcare charges varied more than threefold between the three clusters within “Complex” (Tier 4) patients, from \$69,893 for the “Outpatient” cluster to \$163,700 for the “Inpatient” cluster and \$229,037 for the “Super-User” cluster, as shown in Figure 18.

Figure 18. Average annual charges by episode type within cluster



As shown in Figure 19, “Super-Users” were more likely than other “Complex” (Tier 4) patients to have diagnoses of diabetes or asthma, but less likely to have arthritis, acquired hypothyroidism, atrial fibrillation, or osteoporosis.

Figure 19. Prevalence of chronic conditions by “Complex” (Tier 4) cluster



As shown in Table 4, “Super-Users” were more likely than other “Complex” (Tier 4) patients to be have Medicaid coverage, have poorly controlled blood glucose (HbA1c > 9%), have positive histories for falling, smoking, or hazardous drinking, or have serious mental illness (schizophrenia, major depression, or bipolar depression). The rate of diastolic hypertension (DBP \geq 90 mmHg) was slightly but significantly ($p < 0.05$) lower in the “Outpatient” cluster than the “Inpatient” cluster, but neither was different from the “Super-User” cluster in this measure; and the rate of systolic hypertension (SBP \geq 140 mmHg) was not different between any of the three clusters.

Gender and race were not significantly ($p < 0.05$) different between the three clusters, nor were hyperlipidemia (LDL \geq 100 mg/dL), obesity (BMI \geq 30), or multi-morbidity (number of chronic conditions). “Super-Users” were significantly younger and more likely to be single than other “Complex” (Tier 4) patients. This may have been confounded by the higher rate of Medicaid coverage among “Super-Users,” because LIGHT² eligibility required either Medicaid coverage or Medicare coverage (with a typical age of 65 or older).

Table 4. Selected attributes of “Complex” (Tier 4) patients by cluster

Attribute	All Tier 4 patients	“Outpatient” users	“Inpatient” users	“Super-Users”	p-value
<i>Demographic Attributes</i>					
Age (years)	66.7 ± 17	68.0 ± 16	69.5 ± 16	50.0 ± 18	0.000
Marital Status:					.020
Married	147 (43%)	96 (47%)	39 (37%)	12 (35%)	
Divorced/Separated	62 (18%)	35 (17%)	24 (23%)	3 (9%)	
Single	68 (20%)	35 (17%)	19 (18%)	14 (41%)	
Widowed	66 (19%)	37 (18%)	24 (23%)	5 (15%)	
Medicaid	97 (28%)	47 (23%)	32 (30%)	18 (53%)	0.001
<i>Health Outcomes</i>					
HbA1c > 9%	41 (22%)	21 (19%)	9 (16%)	11 (50%)	0.003
SBP ≥ 140 mmHg	315 (96%)	182 (95%)	100 (98%)	33 (97%)	0.184
DBP ≥ 90 mmHg	284 (87%)	156 (81%)	97 (94%)	31 (94%)	0.003
LDL ≥ 100 mg/dL	74 (22%)	44 (38%)	24 (36%)	6 (31%)	0.864
BMI ≥ 30	196 (61%)	112 (61%)	61 (60%)	23 (70%)	0.580
Chronic conditions	6.2 ± 2.7	6.1 ± 2.6	6.2 ± 2.6	6.3 ± 3.5	0.921
<i>Mental Health and Behavior</i>					
Fall history	139 (41%)	57 (28%)	62 (59%)	20 (59%)	0.000
Smoking	95 (28%)	46 (23%)	30 (29%)	19 (58%)	0.000
Hazardous drinking	33 (11%)	11 (6%)	14 (16%)	8 (28%)	0.001
Serious mental illness	93 (27%)	46 (23%)	32 (30%)	15 (44%)	0.023

While cluster analysis was successful at identifying important sub-groups within the high-cost patients, it was not appropriate for characterizing the differences between high-cost and low-cost patients, which were defined deductively by the study's original risk-stratification schema. Contrast mining was selected as the most effective strategy for this final step.

4.6 Contrast mining of high-cost vs. low-cost patients

Contrast mining of 19,014 clinical attributes from the first year of electronic medical records (EMR) data for 479 high-cost patients and comparison groups with low-cost patients (ranging from the lowest 5% to the lowest 50%) identified 5,188 attribute combinations frequently found (support of 20% or more) in patients with high costs in the second year, but infrequently in other patients. A sample of ten of these combinations is illustrated in Table 5.

Table 5. Ten (out of 5,188) combinations frequently associated with high costs

Attribute Combination	Support
Narcotic analgesics, Analgesics, Platelet aggregation inhibitors	0.21
Anti-hyperlipidemic agents, Analgesics, HMG CoA reductase inhibitors	0.39
Antidepressants, ICD9=311 (Depressive disorder), Antihistamines	0.20
Beta-adrenergic blocking agents, Cardioselective beta blockers, Nutritional products	0.29
Narcotic analgesics, Respiratory agents, Nutritional products	0.20
Race=White, Salicylates, Antiplatelet agents, Platelet aggregation inhibitors, Age=65to84	0.25
Antiplatelet agents, Analgesics, Beta-adrenergic blocking agents, Platelet aggregation inhibitors	0.33
Vitamins, Gastrointestinal agents, Salicylates, Nutritional products, Antiplatelet agents	0.20
Narcotic analgesics, Anxiolytics/sedatives/hypnotics	0.25
Narcotic/analgesic combinations, Gastrointestinal agents, Laxatives	0.23

Not all combinations were infrequent in all comparison groups, but at least 5,178 of the 5,188 contrasted in all seven analyses. These 5,188 contrasting combinations were made up of 67 unique attributes, listed in Appendix 1. LIGHT2 care coordination protocol by risk stratification tier

Coordination Domain	Tier 1	Tier 2	Tier 3	Tier 4
----------------------------	---------------	---------------	---------------	---------------

Negotiate responsibility	-	-	X	X
Communication	-	-	X	X
Facilitate transitions	-	-	-	X
Assess needs and goals	-	-	X	X
Create a proactive plan of care	-	-	X	X
Monitor and respond to change	-	-	-	X
Support self-management	-	-	X	X
Link to community resources	-	-	-	X
Align resources with needs	-	-	-	X

Appendix 2. Individual attributes found in combinations associated with high costs Logistic regression of the 67 contrasting attributes found eleven attributes to be significantly ($p < 0.05$) associated with high costs. This final regression model is detailed in Table 6.

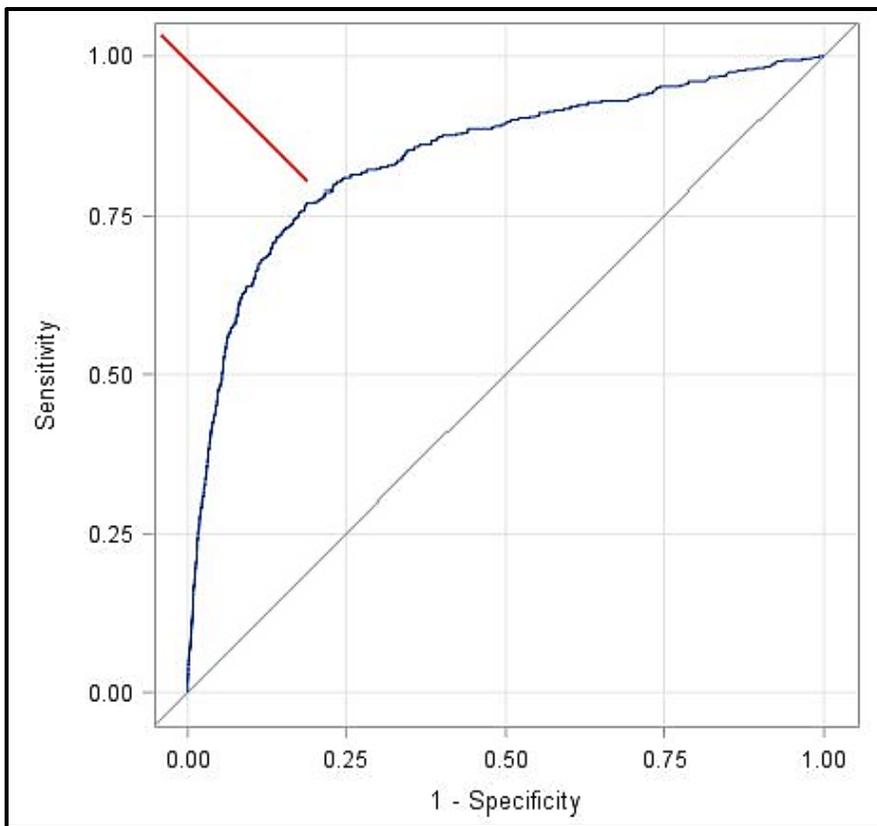
Table 6. Final model of attributes associated with high costs

Attribute	Coef- ficient	p-value	Odds ratio	95% Confidence limits	
<i>Diagnoses</i>					
ICD9=311, Depressive disorder	0.5568	<0.0001	1.707	1.343	2.168
ICD9=401.9, Unspecified essential hypertension	0.3967	0.0007	1.423	1.128	1.795
ICD9=414, Ischemic heart disease	0.5939	<0.0001	1.828	1.386	2.411
ICD9=715, Osteoarthritis	1.0479	<0.0001	2.769	2.192	3.499
<i>Demographic Attribute</i>					
Obesity (BMI \geq 30)	2.3520	<0.0001	9.496	7.530	11.976
<i>Prescription Types</i>					
Anti-infectives	0.4136	0.0060	1.504	1.117	2.025
Benzodiazepines	0.2975	0.0139	1.307	1.026	1.665
Beta-adrenergic blocking agents	0.2832	0.0148	1.314	1.047	1.649
Quinolones	0.4916	0.0087	1.674	1.158	2.421
Respiratory agents	0.3030	0.0063	1.340	1.076	1.668
Selective serotonin reuptake inhibitor (SSRI) antidepressants	-0.4062	0.0019	0.655	0.506	0.847

* Intercept = -4.2585 with p < 0.0001

As shown in Figure 20, the c-statistic of the resulting model was 0.8436, with a 95% confidence interval of (0.8227, 0.8645). By assuming sensitivity and specificity errors to be equally important, an optimal threshold for the model was calculated to minimize the distance to the upper left corner of the receiver operating characteristic graph.

Figure 20. Receiver operating characteristic curve for the final model



This distance from the upper left corner of the graph to the nearest point of the curve was calculated as

$$\sqrt{(1 - sensitivity)^2 + (1 - specificity)^2}$$

(Zhou, Obuchowski, & McClish, 2002) and tuning the model to this threshold produced a sensitivity of 0.770, a specificity of 0.812, a positive predictive value of 0.202, and a negative predictive value of 0.983. While the positive predictive value of 20% and negative predictive value of 98% appear low and high, respectively, they are reasonably useful given a population in which only 5% of patients are truly positive for high cost, and 95% of patients are negative. For example, a positive predictive value of 20% would result in five patients receiving the intervention of care management for every patient actually destined to incur high costs without intervention. This over-treatment penalty may be reasonable because care management is both extremely safe and cost-effective, and because the 98% negative predictive value of the model would direct population health managers away from nearly all patients who will not incur the highest 5% of costs without the intervention. These examples demonstrate the utility of mining the rich data available in the EMR to predict the small number of patients who will incur the majority of healthcare expenses, which support population health managers in focusing preventive and longitudinal care more effectively.

CHAPTER 5. DISCUSSION

5.1 Primary findings

The most important results of these analyses point to the effectiveness of these informatics strategies in identifying patients at risk of high cost and high utilization. The two highest risk stratification tiers, “Unstable” (Tier 3) and “Complex” (Tier 4) comprised only 21% of patients but accounted for 43% of prospective charges. Within “Complex” (Tier 4) patients, cluster analysis found a small group, the most expensive 5% of the most expensive 10%, were nearly twice as costly per patient as “Complex” patients on average. Finally, combining contrast mining with logistic regression predicted the most expensive 5% of patients with 84% accuracy. All the strategies used in this study, from the simplest to the most sophisticated, produced useful insights.

Risk stratification tiers, based on healthcare utilization and charges over the previous twelve months, provided highly useful information about what is likely to occur going forward. This finding confirms the expectation that past utilization is a good predictor of future utilization. When viewed through a population health lens, the extraction of readily available retrospective data can provide care coordinators with useful information that allows them to focus their efforts on those patients whose care needs are most expensive, and who may require more intensive management. The risk stratification tiers examined in this study were defined on the basis of diagnoses that are included in the Chronic Conditions Data Warehouse (Goodman, Posner, Huang, Parekh, & Koh, 2013) and on utilization categories that are in general use in healthcare systems.

Because these are easily reproducible, the LIGHT² risk stratification framework is potentially applicable to other healthcare systems.

Clustering of the 343 patients with “Complex” (Tier 4) utilization histories found three distinct patterns: frequent outpatient visits for 59% of the group, frequent inpatient admissions for another 31%, and frequent utilization of all types for the remaining 10% of “Super-Users”. Average annual healthcare charges varied more than threefold between the three clusters within Tier 4, from \$69,893 for the “Outpatient” cluster to \$163,700 for the “Inpatient” cluster and \$229,037 for the “Super-User” cluster.

The last strategy, a novel two-step combination of EMR data mining with multiple logistic regression, yielded a manageably small number of clinical attributes which accurately predicted the 5% of patients who incurred nearly 50% of healthcare expenses. The model presented here has the virtue of simplicity and interpretability while still achieving an accuracy of 84%, higher than the value of 70% reported in comparable models (O’Caoimha et al., 2015).

5.2 Secondary findings

Initial validation of the risk stratification tiers discovered that there was no difference in mean number of emergency department episodes between “Healthy” (Tier 1) and “Stable” (Tier 2) patients; and the differences in outpatient and inpatient visits between these groups were relatively small. Because the cost and utilization patterns of patients in these risk tiers are not notably different, it may be that time and resources devoted to keeping “Stable” (Tier 2) patients from evolving into the much more costly “Unstable” (Tier 3) utilization pattern would be a useful care coordination strategy.

In a related finding, but contrary to the study's secondary hypothesis that the members of an aging and well insured population would gradually rise in risk of healthcare utilization over the course of three years, patients in lower risk tiers ("Healthy" and "Stable" patients) at the beginning of the study period were most likely to remain in lower risk tiers after three years. Those in higher initial risk tiers ("Unstable" and "Complex" patients) were most likely to move to lower risk tiers after three years. The consistently higher death rate in higher risk tier cohorts, however, provides an additional validation of the predictive utility of the risk tiers.

Another finding that reversed a secondary study hypothesis was that charges for admissions of "Complex" patients were less expensive on average than those for "Stable" (33% higher overall) and "Unstable" (30% higher overall) patients. Most of the difference seems to be explained by "Complex" patients getting less surgery and fewer implants than "Stable" and "Unstable" patients. This may be because complex-care patients were often less appropriate candidates for elective surgery. While the primary findings confirm that using risk stratification to help predict future healthcare utilization and charges is a valuable emerging technique in practice management, these unexpected secondary findings demonstrate that costs and utilization must be understood in some detail in order to discover the knowledge hidden in the data.

Patients in the "Super-Users" cluster of the "Complex" (Tier 4) patients, the costliest 10% of the costliest 5%, had significantly worse mental health, poverty, and behavior risks. This group was more likely than any other to have diabetes and asthma, but less likely to have arthritis and osteoporosis, acquired hypothyroidism, and atrial fibrillation. These protective findings may be related to the younger average age of the

group, which may be a consequence of a greater proportion of Medicaid beneficiaries in this cluster.

All of the four diagnoses identified by contrast mining are among the ten most expensive medical conditions in the U.S. in 2013 (US Agency for Healthcare Research and Quality, 2016): (a) ischemic heart disease (second most expensive), (b) depression (third), (c) osteoarthritis (fifth), and (d) hypertension (eighth). Of the prescription types found to be associated with high utilization, beta-adrenergic blocking agents may be indicative of ischemic heart disease (second most expensive); benzodiazepines may be indicative of depression (third), and respiratory agents may be indicative of chronic obstructive pulmonary disease (sixth). The partial congruence of the sample model with the medical conditions known to be most expensive validates the generalizability of these findings, while demonstrating the potential for other, novel discoveries (i.e., a nearly ten-fold increase in the odds of high costs associated with obesity, increased risks associated with anti-infectives in general and quinolones specifically, and risk reduction associated with SSRI antidepressants).

5.3 Implications of study findings

All these findings provide meaningful answers to the question that arose from a review of the literature, “What predicts healthcare utilization and how can we improve prediction?” From deductive definitions of risk stratification tiers, the simplest and most direct approach, it emerged that past utilization is a powerful predictor of future utilization, despite the tendency of sick patients to return to their baseline health. More detailed analyses of costs and utilization identified important sub-groups within the most

complex and expensive patients, and discovered distinct and sometimes paradoxical patterns of utilization. Applying complex algorithms to mine the deep wealth of EMR data on patients found even more predictors, some confirmatory and some unexpected.

One important design limitation that affects all analyses here is that data were collected from a single healthcare system, and some utilization of other providers may be missing. This could affect both the risk stratification of patients and the number and cost of their admissions after stratification. Another limitation resulting from the enrollment of all eligible patients during the study period is that no control group was formulated for direct testing of the LIGHT2 treatment effect. The opportunity to compare outcomes in a control group would have enabled this study to examine the effects of the LIGHT² intervention on patient risks and detect confounding variables.

Several study findings pointed to the effects of socioeconomic factors such as Medicaid eligibility, but detailed analysis of health disparities was impossible using the data collected here. The use of charges as a proxy for costs limited precision of the financial analyses, particularly because nursing charges are calculated as fixed overhead for patient care. The finding that nearly 3% of the study population did move from higher tiers to “Healthy” (Tier 1) may reflect an error in data representation, because this category was defined as an absence of chronic conditions. Finally, the study sample included only adults who received Medicare or Medicaid services in central Missouri, which may limit the generalizability of these findings to the wider US and global populations. Data from additional years and expanded data sources would refine these conclusions, and these analysis should be validated by replication in other populations.

The most important directions for future development of this research are the use of Medicare and Medicaid claims data for more complete collection of utilization and cost histories, and the analysis of additional data sources for social and geospatial determinants. The finding that obesity drastically increased the odds of high healthcare utilization should be examined further by correlating the co-morbidities of obesity with utilization patterns. The formulation of a comparable control group, with patients of similar risk propensity in a similar setting, would allow important analyses of the effect of the overall LIGHT2 intervention on patient health outcomes, care delivery, and costs. This experimental design might also be structured to detect the effects of payer models on patient risk and calculate pre-test likelihoods for more accurate risk prediction. In addition, hidden Markov models may be useful in understanding the tiers or “states” through which individual patients move to result in the “net” tier counts presented here, and more detailed billing data are needed to confirm the interactions of surgeries, implants, and supply costs in “Complex” (Tier 4) admission costs. Cluster analysis and contrast mining, which have been applied successfully to the “Complex” (Tier 4) sub-population, should also be applied to patients in all risk tiers and clustering analyses should be conducted on contrast-mining results. These methods and sequence mining strategies are needed to find the characteristics of patients in each tier that move to lower or higher tiers, as well as those who remain in their initial tiers.

5.4 Contributions to informatics

This study used accessible EMR data to create simple and effective risk stratification tiers, and strategically applied advanced data-mining and statistical

methodologies to characterize and predict healthcare utilization patterns in finer detail. Because “black box” machine-learning algorithms were excluded from the methodology, the results were clinically interpretable and could support clinical decision support. The study also integrated simple data visualization and cost analyses to improve the use of predictive analytics in managed care populations.

The coefficients of the final regression model can be used to calculate a relative score for all patients in a population (see Table 6). For this model, possible scores would range from a minimum of -4.6647, for a patient with SSRI antidepressants (coefficient = -0.4062 added to the intercept of -4.2585) and no other predictors, to a maximum of +2.4777, for a patient with all predictors except SSRI antidepressants. This can be normalized to an index of 0 to 100 using this formula:

$$Index = (Intercept - Min + \sum_1^n Coefficient_n \cdot presence\ of\ Attribute_n) \frac{100}{(Max - Min)}$$

(Gelman & Hill, 2007) where “Max” is the maximum possible score generated by the given model (in this case, 2.4777) and “Min” is the minimum possible score (in this case, -4.6647). This index gives an approximate “risk” of high utilization in the upcoming year, and patient interventions could be prioritized using this relative risk. Alternatively, clinical alerts could be triggered for patients with indices exceeding a given threshold. By adjusting the threshold of the scoring system, the sensitivity and specificity of the model could be tuned to identify only as many high-risk patients as can be managed. However, because population health management is a low-risk and relatively low-cost intervention, clinical use may benefit from greater sensitivity even at the price of lower specificity.

The results of this study support the hypothesis that clinically interpretable machine-learning algorithms, combined with simpler data analytic techniques where appropriate, can improve risk stratification in population health management of chronic disease. By validating a simple and reproducible risk stratification system, demonstrating longitudinal changes in risk stratification, and characterizing the differences between high-risk and low-risk patients in clinically interpretable terms, these informatics strategies have the potential to support population health managers in understanding and predicting the healthcare needs of their patients.

5.5 Conclusions

A combination of simple and sophisticated strategies in health informatics was able to identify patients within a managed population at risk for higher healthcare costs, demonstrate longitudinal changes in risk stratification, and characterize detailed differences between high-risk and low-risk patients. By producing clinically interpretable rules to predict the small number of patients who will incur the majority of healthcare expenses, these methods can support population health managers in focusing preventive and longitudinal care more effectively. These models, and similar models developed by integrating diverse informatics strategies, could be used by population health managers to further the “Triple Aim” of better health outcomes, better healthcare delivery, and lower costs (Berwick, Nolan, & Whittington, 2008).

APPENDICES

Appendix 1. LIGHT² care coordination protocol by risk stratification tier

Coordination Domain	Tier 1	Tier 2	Tier 3	Tier 4
Negotiate responsibility	-	-	X	X
Communication	-	-	X	X
Facilitate transitions	-	-	-	X
Assess needs and goals	-	-	X	X
Create a proactive plan of care	-	-	X	X
Monitor and respond to change	-	-	-	X
Support self-management	-	-	X	X
Link to community resources	-	-	-	X
Align resources with needs	-	-	-	X

Appendix 2. Individual attributes found in combinations associated with high costs

Age = 25 to 44	Antiarrhythmic agents
Age = 45 to 64	Anticonvulsants
Age = 65 to 84	Antidepressants
Race/ethnicity = White/non-Hispanic	Antidiabetic agents
Female	Antiemetic anti-vertigo agents
Male	Antihistamines
Obesity	Anti-hyperlipidemic agents
Taking Rx as prescribed	Anti-infectives
Taking Rx not as prescribed	Antiplatelet agents
Medicaid	Antipsychotics
Prior high cost	Anxiolytics, sedatives and hypnotics
ICD9 = 250 (Diabetes mellitus)	Benzodiazepine anticonvulsants
ICD9 = 272.4 (Hyperlipidemia)	Benzodiazepines
ICD9 = 311 (Depressive disorder)	Beta-adrenergic blocking agents
ICD9 = 401.1 (Benign essential hypertension)	Bronchodilators
ICD9 = 401.9 (Unspecified essential hypertension)	Calcium channel blocking agents
ICD9 = 414 (Ischemic heart disease)	Cardioselective beta blockers
ICD9 = 715 (Osteoarthritis)	Cardiovascular agents
Analgesics	Dermatological agents
Angiotensin converting enzyme inhibitors	Diuretics
	Gamma-aminobutyric acid analogs
	Gastrointestinal agents

HMG CoA reductase inhibitors	Nutraceutical products
Hormones/hormone modifiers	Nutritional products
Iron products	Platelet aggregation inhibitors
Laxatives	Proton pump inhibitors
Minerals and electrolytes	Quinolones
Miscellaneous analgesics	Respiratory agents
Miscellaneous anxiolytics, sedatives and hypnotics	Salicylates
Muscle relaxants	Skeletal muscle relaxants
Narcotic/analgesic combinations	SSRI antidepressants
Narcotic analgesics	Thiazide and thiazide like diuretics
Nonsteroidal anti-inflammatory agents	Vitamin and mineral combinations
	Vitamins

BIBLIOGRAPHY

- Agrawal, R., & Srikant, R. (1994). Fast algorithms for mining association rules in large databases. *Proceedings of the 20th International Conference on Very Large Data Bases*, 487-499.
- Ahmed, N., & Pearce, S. (2010). Acute care for the elderly: A literature review. *Population Health Management*, 219-225.
- Ahmed, N., Taylor, K., McDaniel, Y., & Dyer, C. B. (2012). The role of an Acute Care for the Elderly unit in achieving hospital quality indicators while caring for frail hospitalized elders. *Population Health Management*, 236-240.
- Amarasingham, R., Patzer, R. E., Huesch, M., Nguyen, N. Q., & Xie, B. (2014). Implementing electronic health care predictive analytics: Considerations and challenges. *Health Affairs (Millwood)*, 1148-1154.
- Bates, D. W., Saria, S., Ohno-Machado, L., Shah, A., & Escobar, G. (2014). Big data in health care: Using analytics to identify and manage high-risk and high-cost patients. *Health Affairs (Millwood)*, 1123-1131.
- Berwick, D. M., Nolan, T. W., & Whittington, J. (2008). The triple aim: Care, health, and cost. *Health Affairs (Project Hope)*, 759-769.
- Bierman, A. S., & Clancy, C. M. (2000). Making capitated Medicare work for women: Policy and research challenges. *Women's Health Issues*, 59-69.
- Black, K. (2011). Professional and personal factors associated with gerontological practice: Implications for training and education. *Educational Gerontology*, 982-994.

- Bodenheimer, T., Wagner, E. H., & Grumbach, K. (2002). Improving primary care for patients with chronic illness: The Chronic Care Model, Part 2. *Journal of the American Medical Association*, 1909-1914.
- Bohman, T. M., Wallisch, L., Christensen, K., Stoner, D., Pittman, A., Reed, B., & Ostermeyer, B. (2011). Working well - the Texas demonstration to maintain independence and employment: 18-month outcomes. *Journal of Vocational Rehabilitation*, 97-106.
- Bøttcher Berthelsen, C., & Kristensson, J. (2015). The content, dissemination and effects of case management interventions for informal caregivers of older adults: A systematic review. *International Journal of Nursing Studies*, 988-1002.
- Boult, C., Pualwan, T. F., Fox, P. D., & Pacala, J. T. (1998). Identification and assessment of high-risk seniors. *American Journal of Managed Care*, 1137-1146.
- Bradley, P. (2012). Predictive analytics can support the ACO model. *Healthcare Finance Management*, 102-106.
- Brazier, H., & Begley, C. M. (1996). Selecting a database for literature searches in nursing: MEDLINE or CINAHL? *Journal of Advanced Nursing*, 868-875.
- Brink, P., & Smith, T. (2008). Determinants of home death in palliative home care: Using the interRAI palliative care to assess end-of-life care. *American Journal of Hospice & Palliative Medicine*, 263-270.
- Burns, M. E. (2009). Medicaid managed care and cost containment in the adult disabled population. *Medical Care*, 1069-1076.
- Centers for Medicare & Medicaid Services. (2014). *2014 CMS Statistics*. Rockville, MD: US Department of Health & Human Services.

- Chechulin, Y., Nazerian, A., Rais, S., & Malikov, K. (2014). Predicting patients with high risk of becoming high-cost healthcare users in Ontario (Canada). *Healthcare Policy*, 68-79.
- Chernew, M. E., Weissert, W. G., & Hirth, R. A. (2001). Heterogeneity of risk in a managed home health care population. *Medical Care*, 1002-1013.
- Chiu, T., Fang, D., Chen, J., Wang, Y., & Jeris, C. (2001). A robust and scalable clustering algorithm for mixed type attributes in large database environment. *Proceedings of the Seventh ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 263-268.
- Cohen, S. B., & Uberoi, N. (2013). *Differentials in the concentration in the level of health expenditures across population subgroups in the US, 2010*. Rockville, MD: Agency for Healthcare Research and Quality.
- Coleman, A. (2014). *Missouri Medicaid Basics, Summer 2014*. St. Louis: Missouri Foundation for Health.
- Coleman, E. A., Wagner, E. H., Grothaus, L. C., Hecht, J., Savarino, J., & Buchner, D. M. (1998). Predicting hospitalization and functional decline in older health plan enrollees: Are administrative data as accurate as self-report? *Journal of the American Geriatric Society*, 419-425.
- Coleman, K., Austin, B. T., Brach, C., & Wagner, E. H. (2009). Evidence on the Chronic Care Model in the new millennium. *Health Affairs (Project Hope)*, 75-85.
- Collins, B., & Wagner, M. (2005). Early experiences in using computerized patient record data for monitoring charting compliance, supporting quality initiatives and

- assisting with accurate charging at Allina Hospitals & Clinics. *International Journal of Medical Informatics*, 917-925.
- Cox, D. R. (1958). The regression analysis of binary sequences (with discussion). *Journal of the Royal Statistical Society, B*, 215–242.
- Cubanski, J., & Neuman, T. (2016, July 20). *The Facts on Medicare Spending and Financing*. Retrieved from The Henry J. Kaiser Family Foundation: <http://kff.org/medicare/issue-brief/the-facts-on-medicare-spending-and-financing/>
- Cucciare, M. A., & O'Donohue, W. (2006). Predicting future healthcare costs: How well does risk-adjustment work? *Journal of Health Organization Management*, 150-162.
- Decker, F. H. (2011). Profit status of home health care agencies and the risk of hospitalization. *Population Health Management*, 199-204.
- Drozda, J. P., Libby, D., Keiserman, W., & Rundhaug, P. (2008). Case management decision support tools: Predictive risk report or health risk assessment? *Population Health Management*, 193-196.
- Eberly, T., Davidoff, A., & Miller, C. (2010). Managing the gap: evaluating the impact of Medicaid managed care on preventive care receipt by child and adolescent minority populations. *Journal of Health Care for the Poor and Underserved*, 92-111.
- Falagas, M. E., Pitsouni, E., Malietzis, G. A., & Pappas, G. (2008). Comparison of PubMed, Scopus, Web of Science, and Google Scholar: Strengths and weaknesses. *Federation of American Societies for Experimental Biology Journal*, 338-342.

- Forbes.com LLC. (2015, January 8). Behind the numbers: Are healthcare costs really getting better? *Forbes*.
- Fraccaro, P., O'Sullivan, D., Plastiras, P., O'Sullivan, H., Dentone, C., Di Biagio, A., & Weller, P. (2015). Behind the screens: Clinical decision support methodologies - A review. *Health Policy and Technology*, 29-38.
- Gelman, A., & Hill, J. (2007). *Data Analysis Using Regression and Multilevel/Hierarchical Models*. New York: Cambridge University Press.
- Gerteis, J., Izrael, D., Deitz, D., LeRoy, L., & Ricciardi, R. (2014). *Multiple chronic conditions chartbook*. Rockville, MD: Agency for Healthcare Research and Quality.
- Ghassemi, M., Celi, L. A., & Stone, D. J. (2015). State of the art review: The data revolution in critical care. *Critical Care*, 118-126.
- Glasgow, R. E., Orleans, C. T., & Wagner, E. H. (2001). Does the Chronic Care Model also serve as a template for improving prevention? *Milbank Quarterly*, 579-612.
- Goodman, R. A., Posner, S. F., Huang, E. S., Parekh, A. K., & Koh, H. K. (2013). Defining and measuring chronic conditions: Imperatives for research, policy, program, and practice. *Preventing Chronic Disease*, 1-16.
- Greene, S. B., Reiter, K. L., Kilpatrick, K. E., Leatherman, S., Somers, S. A., & Hamblin, A. (2008). Searching for a business case for quality in Medicaid managed care. *Health Care Management Review*, 350-360.
- Hall, J. P., Kurth, N. K., & Chapman, S. L. (2015). Medicaid managed care: issues for beneficiaries with disabilities. *Disabilities & Health Journal*, 130-135.

- Halpin, H. A., McMenamin, S. B., Schmittziel, J., Gillies, R. R., Shortell, S. M., Rundall, T., & Casalino, L. (2005). The routine use of health risk appraisals: Results from a national study of physician organizations. *American Journal of Health Promotion, 34-38*.
- Hamar, B., Wells, A., Gandy, W., Haaf, A., Coberley, C., Pope, J., & Rula, E. (2010). The impact of a proactive chronic care management program on hospital admission rates in a German health insurance society. *Population Health Management, 339-345*.
- Hanley, J. A., & McNeil, B. J. (1982). The meaning and use of the area under a receiver operating characteristic (ROC) curve. *Radiology, 29-36*.
- He, D., Mathews, S. C., Kallou, A. N., & Huftless, S. (2014). Mining high-dimensional administrative claims data to predict early hospital readmissions. *Journal of the American Medical Informatics Association, 272-279*.
- Kansagara, D., Englander, H., Salanitro, A., Kagen, D., Theobald, C., Freeman, M., & Kripalani, S. (2011). Risk prediction models for hospital readmission: A systematic review. *Journal of the American Medical Association, 1688-1698*.
- Kantardzic, M. (2003). *Data mining: Concepts, models, methods, and algorithms*. New York: John Wiley & Sons.
- Kaspin, L. C., Gorman, K. M., & Miller, R. M. (2013). Systematic review of employer-sponsored wellness strategies and their economic and health-related outcomes. *Population Health Management, 14-21*.

- Keogh, C., Wallace, E., O'Brien, K. K., Galvin, R., Smith, S. M., Lewis, C., . . . Fahey, T. (2014). Developing an international register of clinical prediction rules for use in primary care: A descriptive analysis. *Annals of Family Medicine*, 359-366.
- Khalid, J. M., Raluy-Callado, M., Curtis, B. H., Boye, K. S., Maguire, A., & Reaney, M. (2014). Rates and risk of hospitalisation among patients with type 2 diabetes: Retrospective cohort study using the UK General Practice Research Database linked to English hospital episode statistics. *International Journal of Clinical Practice*, 40-48.
- Kneipp, S. M., Lutz, B. J., Levonian, C., Cook, C., Hamilton, J. B., & Roberson, D. (2013). Women's experiences in a community-based participatory research randomized controlled trial. *Qualitative Health Research*, 847-860.
- Krist, A. H., Glasgow, R. E., Heurtin-Roberts, S., Sabo, R. T., Roby, D. H., Gorin, S. N., . . . Fernandez, M. E. (2016). The impact of behavioral and mental health risk assessments on goal setting in primary care. *Translational Behavioral Medicine*, 212-219.
- Kruppa, J., Ziegler, A., & König, I. R. (2012). Risk estimation and risk prediction using machine-learning methods. *Human Genetics*, 1639-1654.
- Kruse, R. L., Parker Oliver, D., Mehr, D. R., Petroski, G. F., Swenson, D. L., & Zweig, S. C. (2010). Using mortality risk scores for long-term prognosis of nursing home residents: Caution is recommended. *Journals of Gerontology Series A, Biological Sciences and Medical Sciences*, 1235-1241.

- Kruse, R., Zweig, S., Nikodim, B., LeMaster, J., Coberly, J., & Colwill, J. (2010). Nurse care coordination of older patients in an academic family medicine clinic: 5-year outcomes. *Journal of Clinical Outcomes Management*, 209-215.
- Kwak, J., Kramer, B. J., Lang, J., & Ledger, M. (2013). Challenges in end-of-life care management for low-income frail elders: A case study of the Wisconsin Family Care Program. *Research on Aging*, 393-419.
- Laragy, C., & Allen, J. (2015). Community aged care case managers transitioning to consumer directed care: More than procedural change required. *Australian Social Work*, 212-227.
- LeCheminant, J. D., & Merrill, R. M. (2012). Improved health behaviors persist over two years for employees in a worksite wellness program. *Population Health Management*, 261-266.
- Levitt, A. J., Mitchell, K., Pareti, L., DeGenova, J., Heller, A., Hannigan, A., & Gholston, J. (2013). Randomized trial of intensive housing placement and community transition services for episodic and recidivist homeless families. *American Journal of Public Health*, S348-S354.
- Liu, H., Harris, K. M., Weinberger, S., Serxner, S., Mattke, S., & Exum, E. (2013). Effect of an employer-sponsored health and wellness program on medical cost and utilization. *Population Health Management*, 1-6.
- Loeppke, R., Edington, D., & Bég, S. (2010). Impact of the prevention plan on employee health risk reduction. *Population Health Management*, 275-284.
- Loeppke, R., Nicholson, S., Taitel, M., Sweeney, M., Haufle, V., & Kessler, R. (2008). The impact of an integrated population health enhancement and disease

- management program on employee health risk, health conditions, and productivity. *Population Health Management*, 287-296.
- Mansmann, U., Rieger, A., Strahwald, B., & Crispin, A. (2016). Risk calculators—methods, development, implementation, and validation. *International Journal of Colorectal Disease*, 1111-1116.
- Marchevsky, A. M., & Wick, M. R. (2004). Evidence-based medicine, medical decision analysis, and pathology. *Human Pathology*, 1179-1188.
- McCormack, R. P., Hoffman, L. F., Wall, S. P., & Goldfrank, L. R. (2013). Resource-limited, collaborative pilot intervention for chronically homeless, alcohol-dependent frequent emergency department users. *American Journal of Public Health*, S221-S224.
- McCusker, J., Bellavance, F., Cardin, S., Trepanier, S., Verdon, J., & Ardman, O. (1999). Detection of older people at increased risk of adverse health outcomes after an emergency visit: The ISAR screening tool. *Journal of the American Geriatrics Society*, 1229-1237.
- McDonald, G., Vickers, M. H., Mohan, S., Wilkes, L., & Jackson, D. (2010). Workplace conversations: Building and maintaining collaborative capital. *Contemporary Nurse*, 96- 105.
- Morris, L. S., Schettine, A. M., Roohan, P. J., & Gesten, F. (2011). Preventive care for chronically ill children in medicaid managed care. *American Journal of Managed Care*, e435-442.
- Myers, R. H. (1990). *Classical and modern regression with applications, second edition*. Boston: PWK Kent.

- Newton, J., & Browne, L. (2008). How fair is fair access to care? *Practice*, 235-249.
- O’Caoimha, R., Cornallya, N., Weathersa, E., O’Sullivan, R., Fitzgeralda, C., Orfilad, F., . . . Molloya, D. W. (2015). Risk prediction in the community: A systematic review of case-finding instruments that predict adverse healthcare outcomes in community-. *Maturitas*, 3-21.
- Patterson, K. L., Nochajski, T., & Wu, L. (2014). Neighborhood outcomes of formally homeless veterans participating in the HUD-VASH program. *Journal of Community Practice*, 324-341.
- Popejoy, L. L., Jaddoo, J., Sherman, J., Howk, C., Nguyen, R., & Parker, J. C. (2015). Monitoring resource utilization in a health care coordination program. *Professional Case Management*, 310-320.
- Rohrer, J., Angstman, K., & Bartel, G. (2009). Impact of retail medicine on standard costs in primary care: A semiparametric analysis. *Population Health Management*, 333-335.
- Rohrer, J., Angstman, K., Adamson, S., Bernard, M., Bachman, J., & Morgan, M. (2010). Impact of online primary care visits on standard costs: A pilot study. *Population Health Management*, 59-63.
- Rohrer, J., Takahashi, P., & Adamson, S. (2008). Age, obesity, and medical visits in family medicine. *Population Health Management*, 255-259.
- Sandberg, T., Wilson, A., Rodin, H., Garrett, N., Bargman, E., Dobmeyer, D., & Plocher, D. (2009). Improving the imputation of race: Evaluating the benefits of stratifying by age. *Population Health Management*, 325-331.

- Sanderson, C., & Dixon, J. (2000). Conditions for which onset or hospital admission is potentially preventable by timely and effective ambulatory care. *Journal of Health Services Research & Policy*, 222-230.
- Schmuttermaier, J., Schmitt, D., King, C., & Gwynne, A. (2011). Whole of client health care in a gridlocked system: An insider dialogue between the theory and practice of community case management. *Home Health Care Management & Practice*, 36-49.
- Schraeder, C., Fraser, C., Clark, I., Long, B., Shelton, P., Waldschmidt, V., & Lanker, W. (2008). Evaluation of a primary care nurse case management intervention for chronically ill community dwelling older people. *Journal of Nursing & Healthcare of Chronic Illnesses*, 407-417.
- Schwartz, S., Ireland, C., Strecher, V., Nakao, D., Wang, C., & Juarez, D. (2010). The economic value of a wellness and disease prevention program. *Population Health Management*, 309-317.
- Shin, S., Kim, M., Kim, E., Lee, H., Park, C., & Kim, H. (2010). Medical aid service overuse assessed by case managers in Korea. *Journal of Advanced Nursing*, 2257-2265.
- Shmueli, A., Messika, D., Zmora, I., & Oberman, B. (2010). Health care costs during the last 12 months of life in Israel: Estimation and implications for risk-adjustment. *International Journal of Health Care Finance & Economics*, 257-273.
- Shvachko, K., Kuang, H., Radia, S., & Chansler, R. (2010). The Hadoop distributed file system. *IEEE 26th Symposium on Mass Storage Systems and Technologies*.

- Smith-Gagen, J., Loux, T., Drake, C., & Perez-Stable, E. J. (2016). How does managed care improve the quality of breast cancer care among Medicare-insured minority women? *Journal of Racial and Ethnic Health Disparities*, 496-507.
- Snyderman, R., & Williams, R. A. (2003). Prospective medicine: The next health care transformation. *Academic Medicine*, 1079-1084.
- Springrose, J., Friedman, F., Gumnit, S., & Schmidt, E. (2010). Engaging physicians in risk factor reduction. *Population Health Management*, 255-261.
- Stanton, M. W., & Rutherford, M. K. (2005). *The high concentration of U.S. health care expenditures*. Rockville, MD: Agency for Healthcare Research and Quality.
- Steyerberg, E. W., Pencina, M. J., Lingsma, H. F., Kattan, M. W., Vickers, A. J., & van Calster, B. (2012). Assessing the incremental value of diagnostic and prognostic markers: A review and illustration. *European Journal of Clinical Investigation*, 216-228.
- Suijker, J. J., Burman, B. M., ter Riet, G., van Rijn, M., de Haan, R. J., de Rooij, S. E., & Moll van Charante, E. P. (2012). Comprehensive geriatric assessment, multifactorial interventions and nurse-led care coordination to prevent functional decline in community-dwelling older persons: Protocol of a cluster randomized trial. *BMC Health Services Research*, 85.
- Sun, J., McNaughton, C. D., Zhang, P., Perer, A., Gkoulalas-Divanis, A., Denny, J. C., . . . Malin, B. A. (2014). Predicting changes in hypertension control using electronic health records from a chronic disease management program. *Journal of the American Medical Informatics Association*, 337-344.

- Twells, L. K., Bridger, T., Knight, J. C., Alaghebandan, R., & Barrett, B. (2012). Obesity predicts primary health care visits: A cohort study. *Population Health Management, 29-36*.
- US Agency for Healthcare Research and Quality. (2016, May 26). *Total expenses and percent distribution for selected conditions by type of service: United States, 2013*. Retrieved from Agency for Healthcare Research and Quality: http://meps.ahrq.gov/mepsweb/data_stats/tables_compendia_hh_interactive.jsp?_SERVICE=MEPSSocket0&_PROGRAM=MEPSPGM.TC.SAS&File=HCFY2013&Table=HCFY2013_CNDXP_C&_Debug=
- US Centers for Disease Control and Prevention. (2016, May 16). *Health data interactive*. Retrieved from Centers for Disease Control and Prevention: <http://www.cdc.gov/nchs/hdi.htm>
- US Centers for Medicare & Medicaid Services. (2013). *Trustees report shows reduced cost growth, longer Medicare solvency*. Rockville, MD: Centers for Medicare & Medicaid Services.
- US Congressional Budget Office. (2013). *Health-related options for reducing the deficit: 2014 to 2023*. Washington, DC: US Congressional Budget Office.
- van Walraven, C., Dhalla, I. A., Bell, C., Etchells, E., Stiell, I. G., Zarnke, K., . . . Forster, A. J. (2010). Derivation and validation of an index to predict early death or unplanned readmission after discharge from hospital to the community. *Canadian Medical Association Journal, 551-557*.
- Verhaegh, K. J., MacNeil-Vroomen, J. L., Eslami, S., Geerlings, S. E., de Rooij, S. E., & Burman, B. M. (2014). Transitional care interventions prevent hospital

- readmissions for adults with chronic illnesses. *Health Affairs (Millwood)*, 1531-1539.
- Wagner, E. H., Glasgow, R. E., Davis, C., Bonomi, A. E., Provost, L., & McCulloch, D. (2001). Quality improvement in chronic illness care: A collaborative approach. *Joint Commission Journal on Quality Improvement*, 63-80.
- Wherry, L. R., Burns, M. E., & Leininger, L. J. (2014). Using self-reported health measures to predict high-need cases among Medicaid-eligible adults. *Health Services Research*, 2147-2172.
- Wiley, J. A., Rittenhouse, D. R., Shortell, S. M., Casalino, L. P., Ramsay, P. P., Bibi, S., & Alexander, J. A. (2015). Managing chronic illness: Physician practices increased the use of care management and medical home processes. *Health Affairs (Project Hope)*, 77-86.
- Williams, A., & Cooper, B. (2008). Determining caseloads in the community care of frail older people with chronic illnesses. *Journal of Nursing & Healthcare Of Chronic Illnesses*, 60-66.
- Witten, I. H., & Frank, E. (2005). *Data mining: Practical machine learning tools and techniques*. Burlington, MA: Morgan Kaufmann.
- Zaslavsky, A. M., & Buntin, M. J. (2002). Using survey measures to assess risk selection among Medicare managed care plans. *Inquiry*, 138-151.
- Zhao, Y., Ash, A. S., Ellis, R. P., & Slaughter, J. P. (2002). Disease burden profiles: An emerging tool for managing managed care. *Health Care Management Science*, 211-219.

Zhou, X. H., Obuchowski, N. A., & McClish, D. K. (2002). *Statistical methods in diagnostic medicine*. New York: John Wiley and Sons.

VITA

Lincoln Sheets is a physician and medical researcher who serves as Director of Health Informatics for the School of Medicine at the University of Missouri. He is a Research Assistant Professor in the Department of Health Management and Informatics, and is a doctoral candidate in Health Informatics at the University of Missouri Informatics Institute. His research combines evidence-based medicine with information technology and business strategy, in order to improve health outcomes in underserved rural and minority communities of the United States and around the world.

Dr. Sheets received his medical degree in 2011 from the University of Missouri, where his study on health disparities was named “Outstanding Medical Student Research Publication” for that year. During his 2012 post-doctoral fellowship at the National Library of Medicine he authored four publications on the use of mobile technology for clinical decision support and co-authored a highly cited review of healthcare applications for smartphones. His 2014-2016 fellowship at the University of Missouri Institute for Clinical and Translational Science, on the application of translational informatics to support patient care coordination, has led to three first-author publications and numerous peer-reviewed presentations, including the “Best Student Presentation” of the Missouri Informatics Symposium for three years in a row. He is past president of the American Medical Student Association local chapter and the Missouri Informatics Institute Graduate Students Association, is active in the American Medical Association and the American Medical Informatics Association, and reviews for several academic journals.

Before entering medical school, Lincoln enjoyed a brief career in professional ballet, then earned bachelor’s and master’s degrees and 20 years of experience in software engineering. As a medical student, he organized medical missions in Africa and Latin America, and helped create a student-run free clinic in Columbia, Missouri. He now

lives and serves full-time at St. Francis House, a Catholic Worker community that provides shelter and hospitality to the homeless residents of Columbia, and he is a member and leader of the Columbia Friends (Quaker) Meeting. He volunteers in public advocacy and legislative lobbying locally and nationally. His free-time interests include dance, triathlon, cinema, mountaineering, European languages, martial arts, and spending time with family at home in Columbia and in his hometown of Springfield, Missouri.