# THE USE OF CLINICAL, BEHAVIORAL, AND SOCIAL DETERMINANTS OF HEALTH TO IMPROVE IDENTIFICATION OF PATIENTS IN NEED OF ADVANCED CARE FOR DEPRESSION

Suranga N. Kasthurirathne

Submitted to the faculty of the University Graduate School
in partial fulfillment of the requirements
for the degree
Doctor of Philosophy
in the School of Informatics and Computing,
Indiana University

October 2018

Accepted by the Graduate Faculty of Indiana University, in partial fulfillment of the requirements for the degree of Doctor of Philosophy.

Doctoral Committee

______________________________________
Josette Jones, RN, PhD, Chair

______________________________________
Shaun Grannis, MD, MS

May 30, 2018

______________________________________
Paul Biondich, MD, MS

______________________________________
Saptarshi Purkayastha, PhD

______________________________________
Joshua Vest, MPH, PhD

DEDICATION

To my mother and my wife.

ACKNOWLEDGEMENTS

I wish to thank the members of my dissertation committee for all their guidance, assistance, and support.

Suranga N. Kasthurirathne

THE USE OF CLINICAL, BEHAVIORAL, AND SOCIAL DETERMINANTS OF HEALTH TO IMPROVE IDENTIFICATION OF PATIENTS IN NEED OF ADVANCED CARE FOR DEPRESSION

Depression is the most commonly occurring mental illness the world over. It poses a significant health and economic burden across the individual and community. Not all occurrences of depression require the same level of treatment. However, identifying patients in need of advanced care has been challenging and presents a significant bottleneck in providing care. We developed a knowledge-driven depression taxonomy comprised of features representing clinical, behavioral, and social determinants of health (SDH) that inform the onset, progression, and outcome of depression. We leveraged the depression taxonomy to build decision models that predicted need for referrals across: (a) the overall patient population and (b) various high-risk populations. Decision models were built using longitudinal, clinical, and behavioral data extracted from a population of 84,317 patients seeking care at Eskenazi Health of Indianapolis, Indiana. Each decision model yielded significantly high predictive performance. However, models predicting need of treatment across high-risk populations (ROC's of 86.31% to 94.42%) outperformed models representing the overall patient population (ROC of 78.87%). Next, we assessed the value of adding SDH into each model. For each patient population under study, we built additional decision models that incorporated a wide range of patient and aggregate-level SDH and compared their performance against the original models. Models that incorporated SDH yielded high predictive performance. However, use of SDH did not yield statistically significant performance improvements. Our efforts present significant potential to identify patients in need of advanced care using a limited number of clinical and behavioral features. However, we found no benefit to incorporating additional SDH into these models. Our methods can also be applied across other datasets in response to a wide variety of healthcare challenges.

Josette Jones, RN, PhD, Chair

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

## LIST OF ABBREVIATIONS

American Psychiatric Association (APA)

Area based deprivation indices (ABDI)

Artificial Intelligence (AI)

Beck depression scale (BDI)

Center for Epidemiological Studies-Depression (CESD)

Chronic obstructive pulmonary disease (COPD)

Cornell Scale for Depression in Dementia (CSDD)

Diagnostic and Statistical Manual of Mental Disorders (DSM)

Emergency medical services (EMS)

Hamilton rating scale for depression (HAM-D)

Health Information Exchange (HIE)

Indiana Network for Patient Care (INPC)

Indiana University School of Medicine (IUSM)

Institute of Medicine (IOM)

International Classification of Disease version 10 (ICD-10)

International Classification of Disease version 9 (ICD-9)

K-Nearest Neighbor (KNN)

Kaiser Family Foundation (KFF)

Least Absolute Shrinkage and Selection Operator (LASSO)

Long Short-Term Memory (LTSM)

Logical Observation Identifiers Names and Codes (LOINC)

Natural Language Processing (NLP)

Patient Health Questionnaire (PHQ-9)

Principal component analysis (PCA)

Positive Predictive Value (PPV)

Receiver Operating Characteristic (ROC)

Seasonal affective disorder (SAD)

Social Determinants of Health (SDH)

Unified Medical Language System (UMLS)

United States Preventive Services Task Force (USPSTF)

Web Ontology Language (OWL)
World Health Organization (WHO)

## LIST OF DEFINITIONS

Social Determinants of Health (SDH): Conditions in the environment in which people are born, live, learn, work, play, worship, and age that affect a wide range of health, functioning, and quality-of-life outcomes and risks (Marmot et al., 2008; Office of Disease Prevention and Health Promotion, 2014).

Patients in need of advanced care for depression: Patients suffering from depression whose quality of life and/or health status will degrade if they do not receive specialized treatment above and beyond primary care.

# 1 INTRODUCTION

Mental health issues are widespread and have the potential to influence other health conditions, leading to a significant strain on national resources and impacting overall quality of life. An estimated 43.6 million (18.1%), or roughly one in five Americans age 18 and over have experienced some form of mental illness during the year 2014 (Substance Abuse and Mental Health Services Administration, 2014). The latest available data indicates that approximately 11 million adults, or 4.8% of the U.S. population, had a serious mental illness during the year 2009, alone. In many cases, mental health and substance abuse disorders co-occur; an estimated 7.9 million Americans suffered from co-occurring mental and substance use disorders in 2014 (Substance Abuse and Mental Health Services Administration, 2014).

Depression is the most common type of mental illness the world over (Ferrari et al., 2013). It affects over 26% of the U.S. adult population (Kessler, Chiu, Demler, & Walters, 2005). Depression negatively affects how persons feel, think, and act. It may lead to a variety of emotional and physical problems and can decrease a person's ability to function in society (American Psychiatric Association, 2016). Based on its prevalence and severity, depression can be categorized as follows (National Institute of Mental Health, 2015b):

> (a) Major depression: Severe symptoms that interfere with an individual's ability to work, sleep, study, eat, and enjoy life. Episodes may occur once, but often several times during an individual's lifetime.
>
> (b) Persistent depressive disorder: A depressed mood that lasts for at least two years. Sufferers may experience episodes of major depression, along with periods of less severe symptoms. Forms of depression vary or may develop under unique circumstances. These include, but are not limited to:
>
> - Psychotic depression: Serve depression and psychosis, such as disturbing false beliefs, delusions, or hallucinations.
> - Postpartum depression: Caused by hormonal and physical changes and the overwhelming responsibility of caring for a newborn (Field, 2010).
> - Seasonal affective disorder (SAD): The onset of depression during the winter months, when there is less natural sunlight. SAD generally lifts during spring and summer (Targum & Rosenthal, 2008).
> - Bipolar disorder: A condition that involves experiences of extreme high moods (mania) and extreme low moods (depression).

The impact of depression on American society is quite profound. The most recent data available suggests that depression is a leading cause of disability for people age 15-44 years and is responsible for almost 400 million disability days per year (Colin Mathers, Fat, & Boerma, 2008). In addition, depression is strongly interlinked with many chronic diseases, including diabetes, cancer, cardiovascular disease, asthma, and obesity and can lead to substantially worse health outcomes and reduced quality of life. Persons suffering from severe mental disorders are at risk for a 10-25 year life expectancy reduction, mostly due to chronic physical medical conditions, such as cardiovascular, respiratory, and infectious diseases; diabetes and hypertension; as well as suicide (World Health Organization, 2013c). The direct and indirect consequences of depression can be felt well beyond the healthcare system. The incremental economic burden of depression (covering medical, pharmaceutical, workplace, and suicide-related costs) in the U.S. was evaluated at $210.5 billion in 2010. This is a 21.5% increase from 2005 (Greenberg, Fournier, Sisitsky, Pike, & Kessler, 2015).

However, not all forms or occurrences of depression are equally harmful or serious. As described above, different types of depression can pose different levels of risk. Many individuals who suffer from mild forms of depression may be very unaware of their situation and recover without any assistance. Other less severe cases can be effectively managed by primary care/family practitioners who deliver over a third of all mental-health care services in the U.S. (Wagner et al., 2000; Wells et al., 2000; Williams Jr et al., 2000). However, other cases of depression are more harmful and require more advanced and long-term care (Cuijpers, van Straten, Andersson, & van Oppen, 2008b; Saloheimo et al., 2016a).

As mentioned before, primary care is the first line of defense or treatment against depression. However, U.S. health infrastructure offers numerous services, ranging from inpatient facilities, online support systems, social services, and counseling to treat depression and help improve sufferers' quality of life. These services are provided beyond primary care and are designed to meet specific needs and/or populations at higher risk, thereby improving their overall quality of life and health status. Thus, we define these as methods of advanced care for depression. Patients in need of these services are therefore in need of advanced care for depression.

Identifying cases of depression that require advanced care may be challenging to primary care providers and healthcare team members whose skill sets run broad, rather than deep. Training healthcare teams to successfully identify patients with severe depression would resolve the problem but is unfeasible, given the cost, effort, and time considerations (Force, 2009; Sharp & Lipsky, 2002). Social stigma and ignorance of health issues also encourage depression sufferers to downplay their condition, further increasing difficulty in detection and assessment (Corrigan, 2004). Second, the need for advanced care is also influenced by factors other than disease severity. Factors, such as a patient's background, other clinical conditions, and social and economic status may also contribute toward the need for advanced care. However, providers may lack access to these indicators for their decision-making.

While patients shy away from seeking treatment for behavioral health, they often present to primary care providers seeking care for other medical conditions. In 2012, 8.6 million, or 32.3% of all inpatient stays, involved at least one mental disorder or substance use disorder diagnosis. Only 1.8 million (6.7%) of these inpatient stays were primarily for behavioral health-related challenges (Heslin, Elixhauser, & Steiner, 2012). This implies that, while the identification of patients in need of advanced care is challenging, patients with such needs were actively presenting for care, and thus, primary care treatment facilities are an ideal point of intervention to identify patients in urgent need. Through a better understanding of behavioral health issues and patient backgrounds, primary care providers can facilitate early and comprehensive treatment of depression and direct patients to the care that they need (Muhrer, 2010).

The need to facilitate early and comprehensive treatment of conditions. such as depression, has led to the promotion of integrated care services, which is defined as the seamless provision of a mix of complementary social, behavioral, and medical services by multiple providers. Integrated service delivery has the potential to improve health outcomes by overcoming misalignments between medical and community systems (Kodner & Spreeuwenberg, 2002). Through collaboration and cooperation, integrated services address multiple determinants of health and enable the delivery of preventative and health-promotion-based interventions (Berwick, Nolan, & Whittington, 2008). In treating cases of depression, integrated care services would involve linking patients with advanced care needs with referrals to appropriate mental health, social services, and financial-aid facilities.

But how can providers accurately identify patients in need of integrated care for depression? While depression is common, not all patients require referrals, and many can be effectively treated with minimal effort. Traditionally, providers have used manual screening tools to evaluate the presence and severity of depression. However, these are not optimal, because they tie-up considerable resources and rely heavily on patient-reported outcomes for decision-making (Kerr & Kerr Jr, 2001). These tools utilize only a specific subset of clinical and behavioral data for decision making, while other factors, such as the patient's immediate environment/SDH, are ignored. The cost efficiency of screening approaches are also questionable (Valenstein, Vijan, Zeber, Boehm, & Buttar, 2001).

We seek to investigate alternative approaches to improve the detection of patients in need of advanced care for depression by leveraging existing knowledge, information sources, clinical workflows, and data analytics.

# 2 LITERATURE REVIEW

The study and prevention of depression has won significant interest due to its widespread nature, potential to cause significant harm, and overall impact on the individual and community. Traditionally, clinicians have diagnosed depression and its severity with the assistance of self-administered depression screening tools. Widely used screening tools (together with scientific literature that they first appeared in) include the Beck Depression Scale (Beck, Steer, & Brown, 1996), Center for Epidemiological Studies Depression Scale (Radloff, 1977), Geriatric Depression Scale (Brink, Yesavage, & Lum, 1983), the Patient Health Questionnaire (PHQ-9) (Gilbody, Richards, Brealey, & Hewitt, 2007), and interviewer-administered depression screening tools, such as the Cornell Scale for Depression in Dementia (Alexopoulos, Abrams, Young, & Shamoian, 1988) and the Hamilton Rating Scale for Depression (HAM-D) (Williams, 2001). These scales are significantly influenced by the Diagnostic and Statistical Manual of Mental Disorders (DSM), published by the American Psychiatric Association (APA), which offers a common language and standard criteria for the classification of mental disorders, including depression (American Psychiatric Association, 2013). The DSM has undergone significant revisions since the introduction of DSM-I in 1952. The most recent version, DSM-5, was released in 2013.

A detailed analysis of each of the aforementioned tools and their measuring scales can be found in table 2.1.

Table 2.1. A study of depression measurement tools

| Beck Depression Scale (BDI) | |
|---|---|
| Measurement type | Self-administered |
| Description | Consists of 21 multiple-choice questions. One of the most widely used psychometric tests for measuring depression severity (Beck et al., 1996). |
| Measurement scale | The BDI scale can distinguish between different subtypes of depressive disorders, such as major depression and dysthymia (a less severe form of depression). For people who have been clinically diagnosed, scores from 0 to 9 represent minimal depressive symptoms, 10 to 16 indicate mild depression, 17 to 29 indicate moderate depression, and scores of 30 to 63 indicate severe depression. All scores of 21 or over indicate depression. |

| | |
|---|---|
| Patient Health Questionnaire (PHQ-9) | |
| Measurement type | Self-administered |
| Description | PHQ-9 is a multiple-choice self-report inventory used as a screening and diagnostic tool for mental health disorders of depression, anxiety, alcohol, eating, and somatoform.<br><br>Alternatively, the first two questions of the PHQ-9, also known as PHQ-2, are used to identify the degree to which an individual has experienced depressed mood and anhedonia over the past two weeks. Its purpose is not to establish final diagnosis or to monitor depression severity, but rather to screen for depression (Arroll et al., 2010). |
| Measurement scale | Scores of (0-4) no depression, (5-9) mild, (10-14) moderate, (15-19) moderately severe, and (20-27) severe depression. |
| Hamilton Rating Scale for Depression (HAM-D) | |
| Measurement type | Interviewer administered |
| Description | HAM-D is a multiple-item questionnaire used to provide an indication of depression and serve as a guide to evaluate recovery. Initially considered the "Gold Standard" for rating depression in clinical research, HAM-D has come under criticism, because many scale items are poor contributors to the measurement of depression severity, while other items have poor inter-rater and retest reliability (Bagby, Ryder, Schuller, & Marshall, 2004). |
| Measurement scale | A score of 0-7 is considered normal. Scores of 20 or higher indicate moderate, severe, or very severe depression. Questions 18-20 may be recorded to give further information about depression but are not part of the scale. |
| Cornell Scale for Depression in Dementia (CSDD) | |
| Measurement type | Interviewer administered |
| Description | The Cornell Scale was developed to assess signs and symptoms of major depression in patients with dementia (Alexopoulos et al., 1988). Given that some of these patients may be unreliable, the scale is completed using information elicited via two semi-structured interviews: one for the patient and another for an informant (Alexopoulos et al., 1988). |
| Measurement scale | Each item in the tool is rated on a scale of 0-2 (0=absent, 1=mild or intermittent, 2=severe). A combined score <6 indicates the absence of significant depressive symptoms. Scores of >=10 indicates probable major depression. Scores >18 indicate a definite major depression. |
| Center for Epidemiological Studies-Depression (CESD) | |
| Measurement type | Self-administered |
| Description | CESD is a 20-item measure that asks persons to rate how often they experienced symptoms associated with depression, such as restless sleep, poor appetite, and feeling lonely over the past week (Radloff, 1977). |

| | |
|---|---|
| Measurement scale | Scores for each response option on the CESD scale range from 0 to 3. Patients with an overall score of 16 points or more are considered to be suffering from depression (Radloff, 1977). |

As shown above, BDI, PHQ-9, and CESD are self-administered, while CSDD and HAM-D are interviewer administered. BDI, PHQ-9, and CESD are influenced by different versions of Diagnostic and Statistical Manual criteria for major depressive disorder. BDI consists of 21 items, requires approximately five minutes to complete, and must be purchased (Pearson Education, 2016). In comparison, PHQ-9 consists of nine items, takes about a minute to complete, requires less reading, thanks to its Likert-type scale, and is available free of charge. Despite of these differences, there are strong correlations between results produced by the two tools (Kung et al., 2013). Each of the tools listed in table 5.1 is intended for all audiences, with the exception of CSDD, which is designed specifically for use by elderly patients with underlying cognitive deficits. All tools consist of somatic, cognitive, and affective symptoms. However, HAM-D consists of a larger number of somatic symptoms, and relatively few cognitive and affective symptoms (Hamilton, 1960; Shafer, 2006; Williams, 2001).

A majority of these questionnaires are used to quantify depression severity of patients already diagnosed as suffering from depression. Their ability to accurately determine severity is dependent on patient cooperation. As discussed above, patients who are unwilling to pursue treatment due to social stigma or ignorance may not cooperate, thereby leading to misdiagnosis. These tools seek to assess the presence and/or severity of depression, and not necessarily the need for advanced care for depression. Furthermore, administrating these surveys requires additional effort.

Alternatively, other studies have considered the potential for building clinical decision models to detect patients suffering from depression. Such models are advantageous, because they leverage existing patient data, can be easily integrated into existing clinical workflows, and do not pose an additional burden to the healthcare team. Decision models have proven to be successful in predicting illnesses, such as cancer (Carneiro, Nascimento, & Bradley, 2015; Fakoor, Ladhak, Nazi, & Huber, 2013;

Kasthurirathne et al., 2016; West, Mangiameli, Rampal, & West, 2005), ischemic heart disease (Hauskrecht & Fraser, 2000), and leukemia (Corchado, De Paz, Rodríguez, & Bajo, 2009). A number of studies have already attested the ability to predict depression using clinical data. Studies by Trinh et. al. found that EMR data on billing diagnosis, problem lists, and antidepressants in medication lists were able to identify cases of depression with considerable accuracy (Trinh et al., 2011). In a separate study, regression-based models for predicting depression severity and response to treatment using structured diagnosis and medication codes and free-text clinical reports were able to predict a future diagnosis of depression up to 12 months in advance with considerable accuracy (Huang et al., 2014). These studies indicate that clinical data can be used to successfully identify potential cases and severity of depression. The type of decision model most suited for detecting specific conditions also varies, based on the quality/structure of data at hand and specific assumptions made by different algorithms. However, as discussed before, these approaches seek to predict the presence and/or severity of depression or other illnesses, rather than the ability to identify those in need of advanced care.

Many clinical factors can influence a patient's chances of suffering from depression. But what factors indicate the need for advanced care as a measure against worsening health outcomes and reduced quality of life? Research indicates that depression negatively affects outcomes of many chronic illnesses, such as chronic pain (Garbi et al., 2014) and diabetes (Anderson, Freedland, Clouse, & Lustman, 2001; Katon et al., 2004; Lin et al., 2010). Depression also leads to adverse outcomes among those suffering from Coronary artery disease (Blumenthal et al., 2003), cancer (Giese-Davis et al., 2011), asthma (Eisner, Katz, Lactao, & Iribarren, 2005), stroke (Pohjasvaara, Vataja, Leppävuori, Kaste, & Erkinjuntti, 2001), and chronic obstructive pulmonary disease (COPD) (Ng et al., 2007). Unfortunately, there is no scientific literature that evaluates the impact of depression on the outcome of a great many other chronic illnesses. But given the aforementioned research, this seems likely.

However, such factors are not the only determinants to influence the need for treatment. Studies show that behavioral and social (socioeconomic, community, and public health-based determinants) may also influence the onset and severity of many diseases, including depression, and lead to reduced quality of life and adverse outcomes ((Lorant et al., 2003; McPherson & Armstrong, 2006). Behavioral and social determinants affect a person's ability to manage depression, how they react to different situations, and their accessibility to proper care. A literature search revealed evidence of a number of characteristics that could influence severity of depression and lead to adverse outcomes and reduced quality of life (table 2.2).

Table 2.2. Behavioral and social patterns that influence/indicate need of advanced care for depression

| Behavior | Description |
|---|---|
| Eating disorders leading to obesity and anorexia | Obesity is significantly associated with depression in women, but not in men (Onyike, Crum, Lee, Lyketsos, & Eaton, 2003). Furthermore, the comorbidity between anorexia nervosa and major depression is likely due to genetic factors that influence the risk for both disorders (Wade, Bulik, Neale, & Kendler, 2000). |
| Smoking | A lifetime history of major depressive disorder is more than double in smokers than non-smokers (Aubin, Tilikete, & Barrucand, 1995). In addition, depressive symptoms and, in some cases, serious major depression, may ensue when individuals with a history of depression stop smoking (Glassman et al., 1990). |
| Evidence of childhood abuse and household dysfunction | Evidence of childhood abuse and household dysfunction, such as recurrent physical and emotional abuse, contact sexual abuse, and adults with substance abuse problems, as defined by the Adverse Childhood Experiences (ACE) study, may lead to depression (Anda et al., 2002). |
| Substance abuse | Addictions are common in people suffering from mental health problems. Although substance abuse and mental health disorders, such as depression and anxiety, are closely linked, one does not directly cause the other (Lehmann, Hubbard, & Martin, 2001). |
| Self-harm | Self-harm can be a symptom of many psychiatric illnesses, including depression (Hawton, Rodham, Evans, & Weatherall, 2002). |

Other risk factors for increasing the severity of depression and the potential for adverse outcomes are heavily associated with social inequalities derived from the conditions in which people are born, live, work, and age, and the health systems they may access (Lorant et al., 2003; Marmot et al., 2008; Office of Disease Prevention and Health Promotion, 2014). The poor and disadvantaged suffer disproportionately, but those in the middle of the social gradient are also affected (Allen, Balfour, Bell, & Marmot, 2014). Women are twice as likely to suffer from depression than men due to a higher risk of first onset (Kessler, 2003). Race is also a significant factor in one's mental health; large epidemiologic surveys (Blazer, Kessler, & McGonagle, 1994; Kesler et al., 2003) report that, compared with non-Hispanic whites, African-Americans have lower lifetime rates of major depression and equivalent or lower rates of 12-month major depression. Despite this, African-Americans are overrepresented in high-need populations, have reduced access to mental health services, and receive poorer quality care than whites (General & Services, 2001). Studies show that in 2009, African-Americans and Hispanic Americans used mental health services at about one-half the rate of Caucasian Americans and Asian Americans at about one-third the rate (Agency for Healthcare Research and Quality, 2010). Thus, factors, such as race, gender, and income group may also help identify populations in greater need of treatment for depression. In addition, SDH may also restrict a patient's access to care, further increasing the risk of harm. Factors, such as average household income, access to mental health facilities, transportation, crime, and education levels, impact one's potential to obtain treatment for behavioral health needs and thus, the greater risk of harm (Aneshensel & Sucoff, 1996).

Unfortunately, there have been no efforts to systematically categorize social determinants that influence the need for treatment and heightened risk. However, there have been several efforts to identify and categorize social determinants that affect an individual's overall health and quality of life. Researchers from the Kaiser Family Foundation (KFF) proposed a framework that categorizes social determinants into: (a) economic stability, (b) neighborhood and physical environment, (c) education, (d) food, (e) community and social context, and (f) healthcare system-based factors (Heiman & Artiga, 2015). The World Health Organization (WHO) has also proposed a conceptual framework for action on social determinants (World Health Organization, 2010). This framework categorizes indicators into structural determinants, such as socioeconomic position, social class, political context, etc., and intermediary determinants, such as behavioral and biological factors and psychosocial factors. Braveman et al. proposed a third U.S.-specific framework based on upstream social determinants of health (SDH) spread across: (a) medical care, (b) behavioral needs, (c) living and working conditions, and (e) economic and social opportunities and resources. This framework covers social determinants, as well as individual health status and behaviors (Braveman, Egerter, & Williams, 2011). However, these are wholly conceptualized frameworks that have seen little or no practical implementation/adoption.

Another point of conflict lies in how to communicate the urgency of care a patient requires once he or she is identified as in need of care for depression. As shown in table 5.1, many traditional depression screening tools use scoring systems to identify the severity or type of depression. However, it is unclear how a clinician would operationalize such a rating. Clinicians report that these scales do not necessarily translate to "urgency" in terms of medical care delivery. In many cases, clinicians identify patients in need of treatment and refer them to different departments or clinics most suited to provide specialized mental healthcare. In such a scenario, clinicians may be able to recommend that patients in greater need of care be seen urgently (i.e., at the next available appointment) or, if the need for treatment is not as severe, within a specific timeframe. In most cases, a numeric score, such as those provided by traditional screening tools, may do little to influence or assist in care management. Therefore, further investigation is necessary to understand how to best communicate a sense of urgency for depression care, especially in a context that tries to identify the need for treatment in terms of reduced quality of life and the risk of adverse outcomes.

Previous studies that focused on clinical data alone merely indicated whether or not a patient was suffering from depression. However, such evaluations fail to understand the social and behavioral elements that impact health outcomes and need for treatment. As an example, a poor individual living in a remote area with little access to proper health facilities is at higher risk and need of care than a well-to-do individual who lives in an urban area with excellent access to mental health facilities. While healthcare teams would undoubtedly benefit from identifying patients suffering from depression, there is much more advantage in quantifying an individual's risk level and urgency of care, given their socioeconomic, clinical, and behavioral status.

We seek to investigate alternative approaches to improve the detection of patients in need of advanced care for depression by leveraging existing knowledge, information sources, clinical workflows, and data analytics. Our efforts will be broken into three incremental aims, as follows,

- Aim 1: Develop a comprehensive terminology that models the need for advanced care for depression using clinical, behavioral, and social determinants of health.
- Aim 2: Operationalizing the aforesaid terminology, build decision models capable of predicting a patient's need of advanced care for depression using clinical and behavioral data.
- Aim 3: Evaluate whether the inclusion of SDH has a statistically significant impact on improving the performance of decision models built using clinical and behavioral data.

# 3 METHODOLOGY

## 3.1 Aims and Deliverables

### 3.1.1 Aim 1

Develop a comprehensive terminology that models the need for advanced care for depression using clinical, behavioral, and social determinants of health.

### 3.1.2 Deliverables

Based on existing scientific research, we will identify clinical, behavioral, and SDH proven to affect the severity, impact, and progression of depression. By mapping these determinants to appropriate indicators captured across the healthcare domain, we will develop a terminology that models the need for advanced care for depression.

### 3.1.3 Aim 2

Operationalizing the aforesaid terminology, build decision models capable of predicting a patient's need of advanced care for depression using clinical and behavioral data.

### 3.1.4 Deliverables

(a) A series of decision models capable of predicting the need for advanced care for depression using clinical and behavioral features extracted from the aforementioned terminology.

(b) An evaluation of the performance of each decision model.

### 3.1.5 Aim 3

Evaluate whether the inclusion of SDH has a statistically significant impact on improving the performance of decision models built using clinical and behavioral data.

### 3.1.6 Hypothesis

The inclusion of SDH has a statistically significant impact on the performance of decision models that seek to predict the need for advanced care for depression.

## 3.2 Case Study: Eskenazi Health

Eskenazi Health (formerly known as the Wishard Memorial Hospital) is a leading health care provider for Central Indiana (Eskenazi Health, 2016). Eskenazi is staffed by physicians from the Indiana University School of Medicine (IUSM). The hospital provides a wide range of primary and specialty care services within its hospital, various inpatient facilities, and 11 community health centers. Eskenazi Health is widely known for its nationally recognized programs, such as Midtown Community Mental Health and the Primary Care-Center of Excellence in Women's Health; it is the sponsoring hospital for Indianapolis emergency medical services (EMS). The hospital's mission is to advocate, care, teach, and serve, with emphasis on the vulnerable populations of Marion County, Indiana (Eskenazi Health, 2016). Patients presenting at Eskenazi may undergo inpatient or outpatient care, or be referred to other specialist care centers, such as Midtown Community Mental Health, which serves all behavioral health needs. Like many other hospitals nationwide, Eskenazi would benefit from better approaches to identifying patients suffering from depression and refer these patients to appropriate care facilities.

I propose to leverage clinical and behavioral patient data, as well as patient and aggregate-level SDH, such as aggregate geo-coded community, socioeconomic, and public health data extracted from Eskenazi Health and other state and not-for-profit organizations to build clinical-decision models to identify patients' mental health needs.

## 3.3 Aim 1

### 3.3.1 Proposed approach

In our literature review, we presented a strong justification for the role of clinical, behavioral, and patient and aggregate-level SDH in predicting the need for advanced care for depression. Our study identified a number of clinical and behavioral factors that were strongly linked to the need for advanced care. We also presented the KFF framework, a conceptual framework for modeling aggregate-level SDH that had won significant support but showed little evidence of implementation. In this phase of the study, we will build a comprehensive terminology representing advanced care needs for depression using existing clinical, behavioral, and SDH identified via the KFF framework and other knowledge sources.

But how can data elements presented as abstract concepts be mapped to biomedical data and/or indicators collected by the Eskenazi Health system and other organizations?

#### 3.3.2 Operationalizing abstract concepts against the healthcare continuum

- Clinical and behavioral data

  Eskenazi Health is a participating organization of the Indiana Health Information Exchange (IHIE), a large, statewide Health Information Exchange (HIE) (McDonald et al., 2005). Since its inception in 2004, IHIE has grown to cover 93 out of 114 hospitals and more than 14,000 physicians statewide (Overhage, 2013). Due to its participation in IHIE, medical information of any Eskenazi patient seeking treatment at any of these 93 hospitals can be extracted via IHIE. IHIE stores clinical diagnosis in structured format using International Classification of Disease version 9 (ICD-9) codes and other relevant terminologies. However, certain clinical conditions and diagnoses are not limited to a single code and may be classified across multiple clinical codes spread across different terminologies. How can we identify all code sets for specific conditions and diagnoses of interest? We propose to use the UMLS Metathesaurus, a multi-lingual thesaurus that contains millions of biomedical concepts, synonymous names, and relationships across 199 medical dictionaries using information obtained from existing scientific literature (U.S. National Library of Medicine, 2013) to consolidate larger code sets into actionable concepts.

- Social determinants of health

  In our literature search, we researched frameworks for modeling aggregate-level SDH. As a preliminary study, we sought to understand how these conceptual frameworks could be implemented using indicators collected by the Polis Center, a reputed research organization that seeks to understand the communities that people live in, and the Indiana State Department of Health. We found that, while many of the conceptual categories identified via the KFF framework could be populated using existing data, several, such as vocational training, social integration, etc., could not. Thus, we sought suitable proxies to represent these (table 3.1).

Table 3.1. KFF conceptual indicators, together with practical measures that represent them

| KFF category types | Conceptual indicators | Indicators |
|---|---|---|
| Economic Stability | Employment | Number of employed, unemployment rates, and numbers |
| | Income, expenses, debit, medical bills, support | % of households spending more than 30% on housing, GINI, median family income, median household families w/children <18 in poverty, poverty rate, population with incomes below 125% of poverty level, population with incomes below 185% of poverty level, count x per-capita income/acres |
| Neighborhood and Physical Environment | Housing | Residential building permits, owner-occupied |
| | Transportation | % of trips not in auto |
| | Safety | Juvenile crime, property crime index, violent crime index, collisions per 1000 pop |
| | Parks and playgrounds | Park/greenway/playground within 10 min walk from home |
| | Walkability | Intersections/square mile, % neighborhoods have paved walkways |
| Education | Literacy and language | % of all households with household language of English |
| | Early childhood education | Pct PTQ 3-4 childcares/all childcares, births to mothers with less than high school ed, child-care utilization (PTQ), children under age 6 w/working parents, passing rates IREAD–Grade 3, % of family households with school-aged children within 1 mile of an A or B school by school type (trade public; no private or charter) |
| | Higher education, vocational training | Enrolled in college, attainment (bach degree or higher) |

| | | |
|---|---|---|
| Food | Hunger | Food stamp recipients (SNAP), cash public assistance, % of population within 1 mile of a supermarket/large grocery store |
| | Access to healthy options | % of population within 1 mile of a supermarket/large grocery store |
| Community and social context | Social integration | ? |
| | Support systems | % neighbors are willing to help |
| | Community engagement | Registered voters, voter/registered |
| | Discrimination | ? |
| Healthcare system | Health coverage | Population without health insurance |
| | Provider availability | ? |
| | Provider linguistic and cultural competency | Percent of all households with household language of English |
| | Quality of care | % who feel accepted/respected by their current healthcare provider |

Other patient-level SDH may be extracted from clinical data. We propose to develop a terminology that models advanced care needs for depression using the aforementioned clinical, behavioral, and SDH. Our terminology will be designed using formative principles (Coltman, Devinney, Midgley, & Venaik, 2008) in order to reduce information redundancy for better performance. While we will endeavor to remove redundant or irrelevant indicators from the terminology, not all redundancies may be identified during development. However, such features will continue to be removed if they are identified during decision modeling undertaken as part of aim 2.

## 3.4 Aim 2

### 3.4.1 The problem domain

Currently, patients presenting at Eskenazi with advanced care needs may be referred to mental health services by clinicians or other healthcare team members. Referrals are made via multiple sources. Physicians may refer patients using the referral system in the Gopher G3 medical software (Duke et al., 2014). Alternatively, other members of the healthcare team may use the eClinicalWorks Care coordination for Medical Records (CCMR) tool (eClinicalWorks, 2016) to make referrals. Information, such as type of care required (alcoholism, substance abuse, depression, etc.) and urgency of care (next available, within 3 months, within 1 month, within 2 weeks, and emergency, etc.) may also be documented. All mental health needs identified at Eskenazi Health are referred to the Midtown Community Mental Health Center, which provides comprehensive inpatient and outpatient services for all types of emotional and behavioral health problems, including depression.

We propose to improve the referral process by enabling better detection of patients in need of advanced care for depression by developing decision models using clinical and behavioral indicators extracted from the depression terminology developed during aim 1.

### 3.4.2 Selection of patients for inclusion

We will identify a subset of adult patients (over 18 years of age) who have received at least one outpatient visit for inclusion in our study. This patient sample will include patients who were referred for advanced care for depression, as well as those who were not.

### 3.4.3 Clinical and behavioral data extraction

Clinical and behavioral data of each patient will be extracted from the IHIE using the terminology developed under aim 1. However, given that Eskenazi systems became operational in 2010, will IHIE contain adequate amounts of data for our study? We performed a feasibility study on IHIE data to assess the availability of the aforementioned clinical and behavioral data from patients receiving care at Eskenazi Health. Given technical limitations, our analysis was limited to a subset of relevant ICD-9 codes.

Table 3.2 Availability of relevant clinical and behavioral data in IHIE

| Condition/behavior | ICD9 code/categories used | Approximate no. of patients* |
|---|---|---|
| Diabetes | ICD9:250 Diabetes mellitus | 833 |
| Asthma | ICD9:493 Asthma | 54553 |
| Chronic obstructive pulmonary disease and allied conditions | ICD9 (490-496) Chronic obstructive pulmonary disease and allied conditions | 94399 |

| | | |
|---|---|---|
| Stroke | ICD9:434.91 Cerebral artery occlusion, unspecified, with cerebral infarction | 3554 |
| Eating disorders | ICD9:307.5 Other and unspecified disorders of eating<br>ICD9:278 Overweight, obesity and other hyperalimentation | 231<br><br>46,055 |
| Tobacco abuse | ICD9:15.82 History of tobacco use | 6348 |
| Alcohol abuse | ICD9:305.00 Alcohol abuse, unspecified drinking behavior | 41052 |
| Substance abuse | ICD9:305.90 Other, mixed, or unspecified drug abuse, unspecified | 8845 |
| Self-harming behaviors | ICD9:E958 Suicide and self-inflicted injury by other and unspecified means | 426 |
| History of child abuse | ICD9:V15.41 Personal history of physical abuse | 205 |

#### 3.4.4 Social determinant data extraction

Patient-level SDH will be extracted from clinical data, while geo-coded aggregate-level data covering the domains of education, neighborhood and environment, employment, health levels, and access to various services/facilities will be obtained from the Polis Center (Polis Center, 2018), a not-for-profit that seeks to understand the communities that people live in. Given that Eskenazi Health focuses on the underserved communities of Marion Country, we will only obtain social indicators reported for this county. A preliminary query to these institutions resulted in confirmation that the aforesaid datasets are available and could be shared for use (table 3.2). These data were collected between 2010-2014 and were available at census-tract level.

#### 3.4.5 Clinical decision modeling

Clinical, behavioral, and SDH will be merged into a master dataset. Each row of the dataset will represent a specific patient, while each column will represent various clinical, behavioral, and SDH indicators. Each patient will also be assigned a Boolean outcome variable based on whether or not they have been referred for depression treatment.

#### 3.4.6 Data pre-processing

We will perform appropriate data pre-processing on the master dataset to determine the best approaches to identify patients in need of advanced care for depression. To ensure reproducibility, ease of adoption, and the utilization of existing knowledge and resources, data pre-processing and decision-model building will be performed using publicly available tools, algorithms, and programming languages.

We anticipate that the master dataset will consist of a considerable number of attributes, or "features." However, not all features may contribute equally towards the accuracy of the models. Many features may be redundant, irrelevant, or lack adequate variance to contribute towards predicting the outcomes at-hand. Using such data may lead to overfitting and reduce accuracy (Hawkins, 2004). Therefore, we will consider multiple approaches to filter the master data set and identify only the most important/relevant features for decision-model building. Using smaller feature sets enables: (a) the simplification of models, making them easily interpretable by researchers, (b) shorter training times, and (c) enhanced generalization via the prevention of over-fitting (Jain & Zongker, 1997; Peng, Long, & Ding, 2005).

We will consider the following approaches for identifying the most relevant features:

- Feature selection: The process of identifying a subset of the most relevant features with the least loss of information. Widely known methods of feature selection include information gain, also referred to as Kullback-Leibler divergence (Polani, 2013).
- Feature engineering: The process of combining existing features into composite features that better represent nuances of the dataset under test, resulting in improved model accuracy on unseen data (Turner, Wolf, Fuggetta, & Lavazza, 1998).

We will consider multiple feature engineering approaches, including principle component analysis (PCA) (Jolliffe, 2002), as well as the feasibility of adopting comorbidity indexes to reduce the dimensionality of the feature set. Indexes, such as the Charlson Comorbidity Index, which is used to predict one-year mortality for patients who may suffer from a total of 22 different comorbid conditions, such as heart disease, AIDS, and cancer (Charlson, Pompei, Ales, & MacKenzie, 1987), and the Elixhauser Comorbidity Measure, an alternative measure developed based on a list of 30 comorbidities linked to the ICD-9-CM coding manual, may be considered for evaluation (Sharabiani, Aylin, & Bottle, 2012).

### 3.4.7 Decision-model building

Decision models may be built using two different approaches: (a) classification or supervised algorithms that require training data to group similar instances based on pre-tagged training data, and (b) clustering or unsupervised algorithms that do not require training data, which group similar instances based on available features. Given the reliability of classification algorithms and the availability of training data (i.e., a "gold standard," as indicated by the presence of an outcome variable, "was the patient referred for depression care or not?"), we will attempt to build appropriate decision models using classification algorithms.

We will randomly separate the master dataset into training and test datasets. Decision models may be trained using a number of potential algorithms representing the spectrum of existing classification approaches:

- Regression-based algorithms: Examples: Linear and logistic regression.
- Instance-based algorithms: A family of learning algorithms that perform lazy learning by comparing new problem instances with instances seen in training (Brighton & Mellish, 2002). Examples: k-Nearest Neighbor (KNN) algorithm.
- Decision trees: A class of learning algorithms that map observations about an item to conclusions about the item's target value (Quinlan, 1986). Examples: C4.5 and Random Forrest.
- Bayesian networks: A probabilistic graphical model that represents a set of random variables and their conditional dependencies via a directed acyclic graph (Nielsen & Jensen, 2009). Examples: Naive Bayes.

Dependent and independent variables for our study are as follows:

- Dependent variables: Is the patient in need of advanced care for depression?
- Independent variables: The clinical, behavioral, and SDH selected for decision-model building.

### 3.4.8 Evaluation

Each decision model will be tested with holdout data to evaluate its accuracy. For each patient in the holdout test dataset, the decision model will predict a binary outcome (i.e., does this patient need advanced care for depression, or not) and the predictive probability for this decision. These results would be used to calculate a number of performance metrics, such as Positive Predictive Value (PPV), sensitivity, specificity, overall accuracy, and area under the ROC curve.

## 3.5 Aim 3

We seek to determine if the inclusion of patient and aggregate-level SDH has a statistically significant impact on the accuracy of the decision models. In the previous analysis, we built decision models using a master feature set comprising clinical and behavioral features (model A). To evaluate the contribution of SDH in this evaluation, we will train additional models that include SDH (model B) and evaluate their performances with model A. The two decision models will be tested using test holdout data and be compared based on their overall accuracy. Figure 3.1 represents the potential decision-model building process for aims 2 and 3.

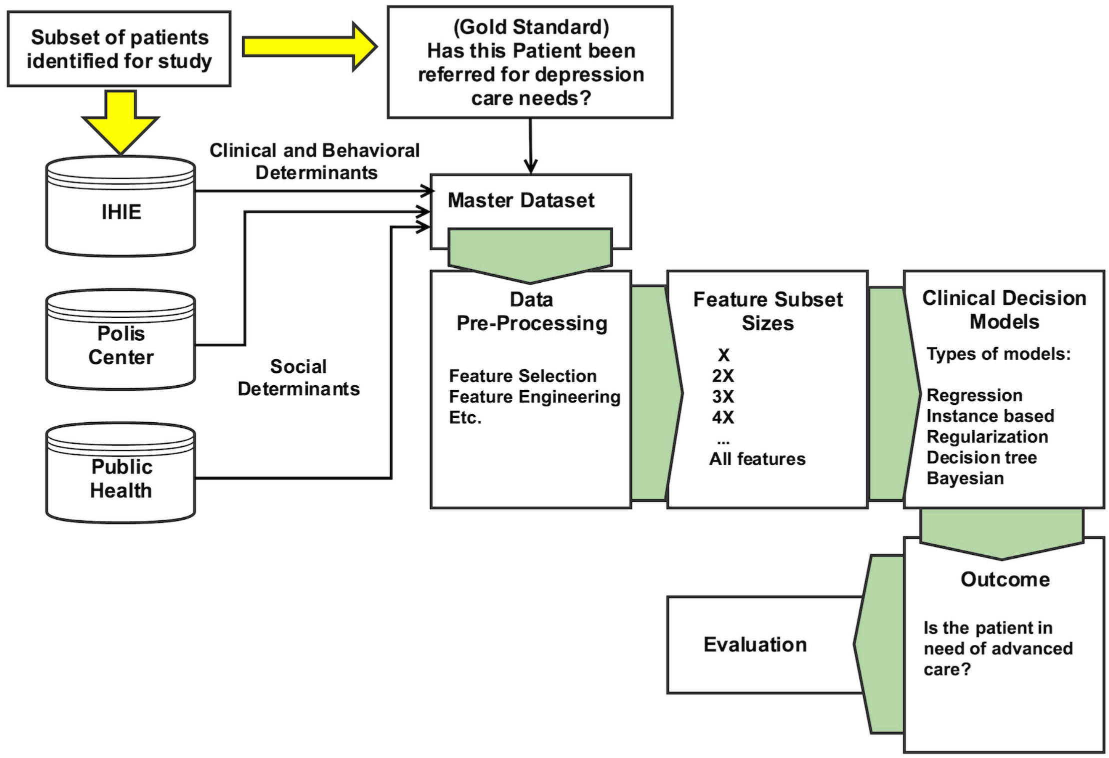



Figure 3.1. The study approach, from the preparation of the master dataset to data pre-processing, decision model building, testing, and outcome evaluation.

### 3.5.1 Evaluation

In aim 2 above, we described challenges in assessing decision model performance and how they may be evaluated. Evaluation for aim 3 will be similar to that prescribed under aim 2. We will calculate the overall accuracy measure and 95% confidence intervals for models A and B. The two models will be compared with one another based on their accuracy and 95% confidence interval, which will be used to determine if there is a statistically significant advantage in using SDH to predict the need for advanced care for depression (Figure 3.2).

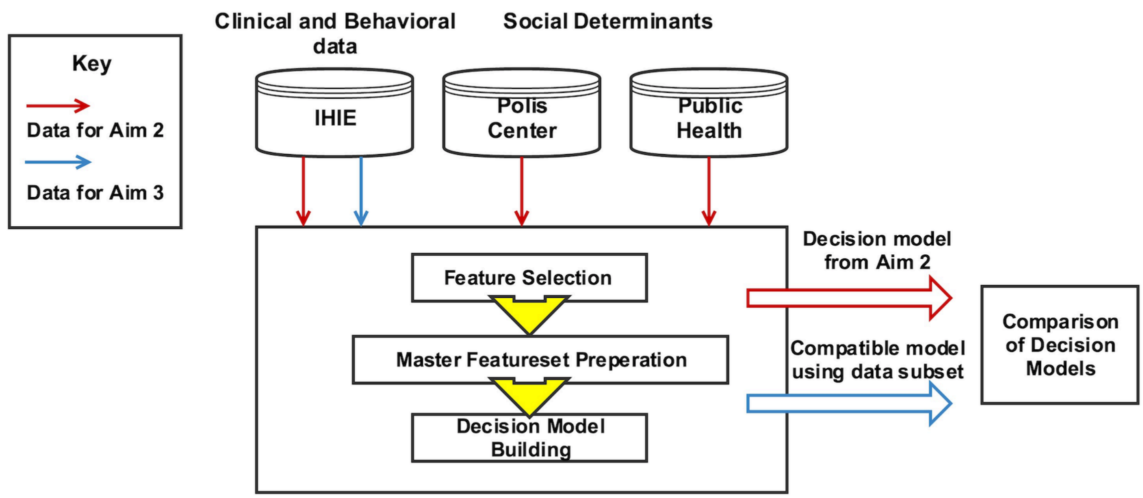



Figure 3.2 Comparison of the performance of decision models built using (a) clinical, behavioral and SDH data, and (b) clinical and behavioral data only

# 4 AIM 1: DEVELOPMENT OF A KNOWLEDGE-DRIVEN MEDICAL TERMINOLOGY FOR BIG DATA-BASED DECISION MODELING FOR NEED OF TREATMENT FOR DEPRESSION

## 4.1 Introduction

Depression has a profound effect on society (Ferrari et al., 2013). It is the most common mental illness, hindering the emotions, thoughts, and behaviors of over 350 million people worldwide (Mathers, Boerma, & Fat, 2008; World Health Organization, 2013a). Depression is also strongly correlated with many chronic diseases, including diabetes (Greenberg et al., 2015), cancer (National Institute of Mental Health, 2015a), cardiovascular disease, asthma, and obesity, and can lead to poor health outcomes and reduced quality of life (Goldberg, 2011). Persons suffering from severe mental disorders are at risk of a 10-25 year life expectancy reduction, mostly due to chronic physical medical conditions, such as cardiovascular, respiratory, and infectious diseases; diabetes; hypertension; and suicide (World Health Organization, 2013b). Studies show that depression is also a leading cause of disability for Americans age 15-44 years, and is responsible for almost 400 million disability days per year (Mathers, 2008). The incremental economic burden of depression (covering medical, pharmaceutical, workplace, and suicide-related costs) in the U.S. was evaluated at $210.5 billion in 2010, a 21.5% increase from 2005 (Greenberg et al., 2015).

### 4.1.1 Types of depression

Depression can be grouped into two primary categories: (a) Major depression: Severe symptoms that interfere with an individual's ability to work, sleep, study, eat, and enjoy life. Episodes may occur once, but often several times during an individual's lifetime; (b) Persistent depressive disorder: A depressed mood that lasts at least two years (Bressert, 2017). A sufferer may have episodes of major depression, along with periods of less severe symptoms, but symptoms must last for up to two years. Forms of persistent depressive disorder may vary or may develop under unique circumstances. These include psychotic depression, postpartum depression, Seasonal Affective Disorder (SAD), and bipolar disorder (National Institute of Mental Health, 2015a).

As a heterogeneous disease (Goldberg, 2011), not all instances of depression are equally harmful or serious. Individuals who suffer from mild forms of depression may be unaware of their situation and recover without assistance. Yet others may be satisfactorily treated in primary care (Cape, Whittington, Buszewicz, Wallace, & Underwood, 2010). However, other cases of depression pose greater risk and require more advanced and long-term care (Cuijpers et al., 2008b; Saloheimo et al., 2016a). Studies show that depression severity and need for care may also be influenced by a wide variety of factors beyond clinical and behavioral indicators. Socioeconomic conditions, such as education, physical environment, feelings of safety, and access to various facilities/services, may also influence the onset and severity of depression, leading to reduced quality of life and adverse outcomes (Lorant et al., 2003; McPherson & Armstrong, 2006; Santos, Kawamura, & Kassouf, 2012). Socioeconomic conditions may also correlate with a person's ability to manage depression, how they react to different situations, and their accessibility to proper care. Thus, such determinants may be invaluable in determining need of care for patients suffering from depression.

Given that depression may be triggered by many conditions and aggravated by various social determinants, it is difficult for healthcare providers to identify patients in need of treatment. Traditionally, providers have used manual screening tools, such as the Patient Health Questionnaire (PHQ-9) (Gilbody et al., 2007), the Cornell Scale for Depression in Dementia (Alexopoulos et al., 1988), and the Hamilton Rating Scale for Depression (HAM-D) (Williams, 2001), to evaluate the presence and severity of depression. However, such tools pose a number of limitations: (a) they encumber considerable resources, (b) rely heavily on a narrow set of patient reported outcomes that may be impacted by fear of stigma for decision-making (Conner et al., 2010; Kerr & Kerr Jr, 2001), and (c) ignore a greater number of clinical and socioeconomic determinants that may influence the onset and severity of depression.

Recent studies demonstrate significant potential for Artificial Intelligence (AI), which leverages the concepts of machine learning, Natural Language Processing (NLP), and knowledge representation for informed medical decision making using clinical, social, and behavioral data (Kasthurirathne, Vest, Menachemi, Halverson, & Grannis, 2017). Given that the onset and severity of depression may be influenced by a variety of clinical and social determinants, predictive models may help overburdened caregivers identify patients in need of care for depression. They improve the accuracy and efficiency of screening by leveraging existing clinical data and may be easily integrated into existing healthcare delivery systems.

But which clinical, behavioral, and social determinants are most relevant for inclusion in such a study? Traditionally, medical terminologies have provided inputs into predictive models (Kasthurirathne, Dixon, et al., 2017a) for various clinical outcomes (Kasthurirathne, Dixon, et al., 2017b). For example, terminologies, such as Berman's Tumor Taxonomy (Berman, 2004), have been used to predict occurrences of cancer with significant success (Kasthurirathne, Dixon, et al., 2017b). Given the diverse nature of depression's presenting symptoms, there is no single terminology that represents the full range of clinical, behavioral, and socio-economic determinants that influence the risk of depression. However, we hypothesize that concepts of relevance do exist and are fragmented across multiple medical terminologies. Given the significant cost and manual effort required to build a condition-specific medical terminology from scratch (Hinz et al., 2010), it is more cost efficient to model such a terminology using medical concepts learned from existing knowledge and terminology sources.

We present a novel, automated approach that combines existing knowledge and terminology systems to compose a knowledge-driven terminology that, coupled with machine-learning approaches, can be used to identify the need of treatment for depression.

### 4.2 Materials and methods

We extracted relevant concepts from the Unified Medical Language System (UMLS) Metathesaurus, a multi-lingual thesaurus that contains millions of biomedical concepts, synonymous names, and relationships across 199 medical dictionaries using information obtained from existing scientific literature (U.S. National Library of Medicine, 2013).

We performed a literature search using Ovid Medline to identify publications on depression and its treatment. Due to the terms of use restricting extraction and processing of entire manuscripts, we extracted only plaintext abstracts for each identified publication. We used Metamap, a Natural Language Processing (NLP)-based tool to map these abstracts against the UMLS Metathesaurus, and thus identify clinical concepts of relevance from 199 medical dictionaries that are part of the UML Metathesaurus (Aronson & Lang, 2010). We used term frequency (tf) to identify the most significant concepts from the abstracts. Next, we removed all concepts that belonged to the semantic categories animal (anim), bird (bird), mammal (mamm), and fish (fish) on the grounds that they were irrelevant for representing human health, and all concepts that belonged to the semantic categories cell component (celc), cell function (celf), cell (cell), nucleotide sequence (nusq), as they referred to cell-level functions that were unrelated to predicting the need of care for depression. We also replaced a number of country names with a generic country tag, race names with a generic race concept, and non-medical occupations with a single-occupation concept.

While the influence of aggregate-level social determinants on depression and overall healthcare delivery is widely known, existing vocabularies are patient-centric and not designed to cover aggregate-level social determinants of health. Thus, we anticipated that the 199 dictionaries included in the UMLS Metathesaurus would not help to identify such concepts. We sought alternative approaches to identify aggregate-level social determinants of health. Researchers from the Kaiser Family Foundation (KFF) proposed a framework that categorizes social determinants of health into subcategories representing an individual's socio-economic environment: (a) economic stability, (b) neighborhood and physical environment, (c) education, (d) food, (e) community and social context, and (f) healthcare system-based factors (Heiman & Artiga, 2015). Given the lack of adequate concepts to model these terms, we used a pre-coordination approach to represent said concepts without our system. These concepts were compiled into a terminology using the Web Ontology Language (OWL) V2 (Motik et al., 2009), a semantic web language that is widely used to represent ontologies. A flowchart presenting our complete approach can be seen in figure 4.1.

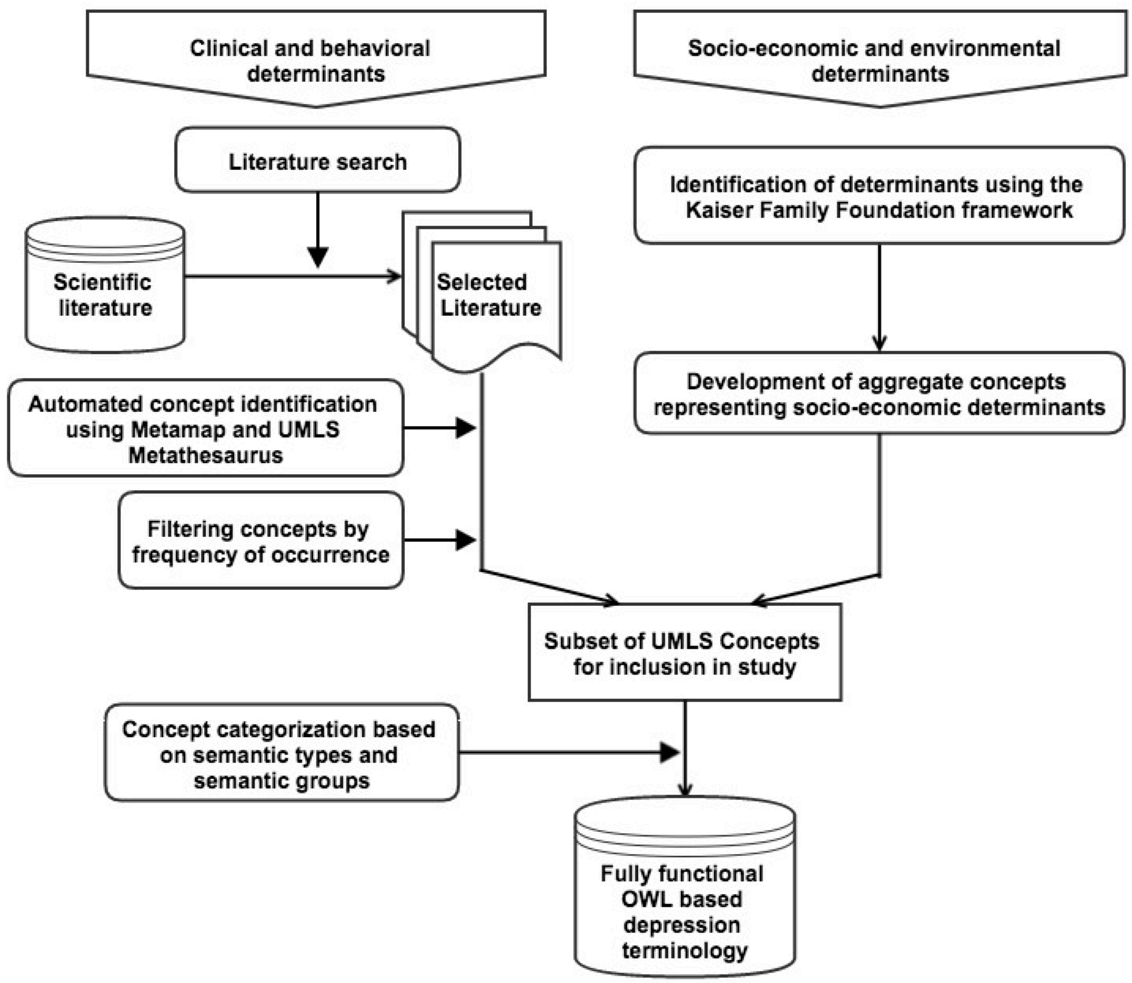



Figure 4.1 The depression terminology development approach

### 4.2.1 Terminology assessment

Researchers have proposed a number of approaches for evaluating terminologies (Brank, Grobelnik, & Mladenić, 2005). However, many of these approaches were challenging: (a) data-driven evaluation approaches call for terminology evaluation by comparing them to existing knowledge sources (usually a collection of textual documents) on the problem domain. However, we cannot adopt such an evaluation approach, as the existing body of scientific research was already used to develop our terminology; (b) taxonomic and semantic evaluation approaches call for analysis of the structural fit of concepts (Brewster, Alani, Dasmahapatra, & Wilks, 2004). However, we cannot adopt this approach, as we rely on the UMLS Metathesaurus as the gold standard for extracting taxonomic and semantic relationships. Thus, the taxonomic and structural accuracy of our terminology would be dependent on the accuracy of the UMLS Metathesaurus; (c) expert-led human evaluation of the depression terminology is also infeasible, given the wide range of concepts representing a large number of expert domains.

As an alternative, we will evaluate the completeness of the depression terminology by comparing it against other existing depression screening tools, and thus, assess its overall coverage and suitability.

### 4.3 Results

Our literature search identified a total of 3,890 publications published in the English language over the last 10 years (table 4.1).

Table 4.1. Literature search to identify factors indicating need of treatment for depression (exp = expression)

| | | |
|---|---|---|
| 1 | exp Depression/or *Long-Term Synaptic Depression/or exp Cortical Spreading Depression/or exp Depression, Postpartum/or exp Depression, Chemical/or depression.mp. | 335,211 |
| 2 | exp Depressive Disorder/or exp Seasonal Affective Disorder/ | 97,664 |
| 3 | exp Bipolar Disorder/ | 36,536 |
| 4 | 1 or 2 or 3 | 386,130 |
| 5 | exp Emergency Medical Services/or exp Palliative Care/ or advanced care.mp. | 165,171 |
| 6 | exp Hospitalization/or exp Hospitals, Psychiatric/or inpatient care.mp. or exp Psychiatric Department, Hospital/ | 224,129 |
| 7 | 5 or 6 | 371,706 |
| 8 | 4 and 7 | 11,277 |
| 9 | limit 8 to (English language and last 10 years) | 4,046 |
| 10 | remove duplicates from 9 | 3,890 |

Metamap analysis of the abstracts of these publications identified a total of 17,019 unique concepts from the UMLS Metathesaurus. By assessing frequency of these concepts, we selected a frequency cutoff threshold of six or more occurrences, which left us with 7,402 concepts for further study. These concepts presented a wide variety of genetic, clinical, demographic, behavioral, and patient-level SDH features that impacted the onset and severity of depression. Next, the concepts were grouped into a hierarchy based on their semantic type. Table 4.2 presents the number of concepts included under each parent concept of our terminology. For a more detailed breakdown, see Appendix 9.1.

Table 4.2. Frequency of UMLS concepts included under each parent concept

| Type | No. of concepts |
|---|---|
| Activities and behaviors | 389 |
| Anatomy | 204 |
| Chemicals and drugs | 604 |
| Concepts and ideas | 2,499 |
| Disorders | 1,250 |
| Genes and molecular sequences | 194 |
| Geographic areas | 162 |
| Living beings | 415 |
| Objects | 213 |
| Occupations | 108 |
| Organizations | 171 |
| Phenomena | 146 |
| Procedures | 1,047 |

These 7,402 concepts, representing clinical and behavioral indicators, were mapped to a total of 45,297 other UMLS concepts from 70 different terminologies of the UMLS Metathesaurus (Appendix 9.2). Assessment of SDH, defined by the Kaiser Family Foundation framework, led to the creation of 56 additional concepts that represented aggregate-level SDH. These concepts were grouped under six different parent concepts (Appendix 9.3).

Our efforts resulted in the development of a valid OWL-based terminology. We evaluated the terminology by comparing it against other conventional depression screening tools. Our analysis indicated complete or high coverage of each of these tools (table 4.3).

Table 4.3. Comparison of terms in depression terminology against other existing depression screening tools

| Depression screening tool | Feature coverage | Features missing from the depression terminology |
|---|---|---|
| PHQ-9 scale<br>(9 questions) | 9/9 (100%) | None |

| | | |
|---|---|---|
| HAM-D scale<br>(21 questions) | 18/21 (85.7%) | Hypochondriasis<br>Diurnal variation<br>Paranoid symptoms |
| Cornell Scale for Depression and Dementia (19 questions) | Mood-related signs: 4/4 (100%)<br>Behavioral disturbance: 4/4 (100%)<br>Physical signs: 3/3 (100%))<br>Cyclic functions: 2/4 (50%)<br>Ideational disturbance: 4/4 (100%) | Early morning awakening<br>Diurnal variation |
| Beck Depression Scale<br>(21 questions) | 21/21 (100%) | None |
| Center for Epidemiological Studies-Depression scale<br>(21 questions) | 21/21 (100%) | None |

## 4.4 Discussion

Our efforts resulted in the development of a well-defined terminology that covered a wide variety of clinical, behavioral, genomic, as well as patient- and aggregate-level SDH that influence the health of an individual suffering from depression. The depression terminology included a total of 7,402 UMLS concepts mapped to 45,297 additional concepts from across 70 other UMLS terminologies, as well as 56 additional concepts that represented aggregate-level SDH developed based on the KFF framework. These concepts, while unmapped to the UMLS Metathesaurus, represented a wide range of indicators representative of an individual's role in society. As presented in table 4.3, concepts in the depression terminology covered a majority of indicators present in existing depression screening tools. The depression terminology was developed using expert knowledge automatically extracted from peer-reviewed scientific literature. Thus, our approach does not require time- and cost-intensive human experts for concept identification.

Our efforts were impacted by the lack of terminologies that represented aggregate-level SDH, which prevented our ability to identify their impact on depression from scientific literature. Our terminology also included many concepts that, though not indicative of the need for depression, would be used as responses or measures used to quantify other concepts. In contrast to existing screening tools and risk models, our terminology included a total of 194 concepts on genes and molecular sequences. We believe that these are highly relevant, given the significance of genes on the onset and severity of depression, as well as calls to improve the use of genetics in managing depression (Hyman, 2014).

While our terminology represents a wide range of features indicating risk of depression, its use is dependent on the quality of data at hand. For example, we recognize that not all concepts identified in our terminology may be reported in adequate prevalence for use in machine learning. Thus, implementation of the terminology for machine learning purposes may change based on the healthcare setting. In some instances, we may be required to use different groupings of concepts in order to better represent the dataset being used. However, this should be relatively easy, given the well-defined UMLS concept mappings included in the terminology.

We identified several limitations in our approach. Due to the nature of our approach and the intended application to machine learning, we anticipate that our terminology maximizes capturing all true positives, and thus, may contain a number of false positives. We recognize that while all concepts may be relevant for every use case, for example, concepts, such as country names and cities included under the "geographic areas" parent concept, may be irrelevant. The applicability of other clinical data may also change based on availability of data. However, such concepts could be easily removed from machine-learning models using feature selection approaches (Guyon & Elisseeff, 2003). Thus, we present that our terminology presents "fitness for purpose" for use in machine learning. Second, we sought to identify a subset of UMLS concepts that could model an individual's risk of depression. Thus, our terminology reflects any weaknesses or limitations in the terminologies that are part of the UMLS Metathesaurus. Finally, our terminology leverages a social determinants model intended to represent overall health status, and not specifically social determinants influencing depression.

We believe that the approach used to develop the depression terminology presents a convenient, low-cost effort that requires minimal human effort and expert intervention.

Our approach may also be replicated for healthcare challenges other than depression and will be invaluable in leveraging existing knowledge and terminologies with machine learning for improving automated surveillance and prediction.

Future steps include an application-driven evaluation of the terminology by populating it with clinicals obtained from the Indiana Network for Patient Care (INPC) (McDonald et al., 2005), and social determinants of health obtained from the SAVI Community Information System, the nation's largest provider of spatially-enabled socio-economic determinants (Bodenhamer, Colbert, Comer, & Kandris, 2011) and using it to predict the need for advanced care for depression.

## 4.5 Conclusion

We developed a well-defined terminology that encompasses a wide variety of clinical, behavioral, genomic, as well as patient and aggregate-level social determinants that influence the health of an individual suffering from depression. Given the high dimensionality of this terminology, we hypothesize that our terminology would be highly suited for application in big data-based machine learning efforts to inform depression care. Our terminology would inform feature extraction from big datasets that could then be filtered down for machine learning via various feature selection approaches. Additionally, our methods are replicable and may be duplicated to identify features across a range of other healthcare conditions with low-cost and human expertise requirements. However, features of relevance may change, based on the completeness and accuracy of the dataset to which the terminology is applied.

# 5 AIM 2: PREDICTING NEED OF ADVANCED CARE FOR DEPRESSION USING CLINICAL, BEHAVIORAL, AND DEMOGRAPHIC DATA AND MACHINE-LEARNING APPROACHES

## 5.1 Introduction

Depression is the most commonly occurring mental illness the world over (World Health Organization, 2012). It negatively affects how up to 350 million persons worldwide think, feel, and interact (American Psychiatric Association, 2016).

Depression poses significant health and economic burdens across both the individual and community (Lépine & Briley, 2011). Studies present a strong comorbidity between mental health and medical conditions (Goodell, Druss, Walker, & Mat, 2011). Depression is highly prevalent across patients suffering from chronic conditions, such as chronic obstructive pulmonary disease, chronic heart failure (Yohannes, Willgoss, Baldwin, & Connolly, 2010), and type 2 diabetes (Nouwen et al., 2010). Such patients may suffer up to a 10-25 year reduction in life expectancy due to worsening health conditions and suicide (Hawton, i Comabella, Haw, & Saunders, 2013; World Health Organization, 2016). Depression is also a leading cause of disability for Americans age 15-44 years and is responsible for up to 400 million disability days per year (Mathers et al., 2008). The incremental economic burden of depression covering medical, pharmaceutical, workplace, and suicide-related costs in the U.S. was evaluated at $210.5 billion in 2010, a 21.5% increase from 2005 (Greenberg et al., 2015).

Many healthcare systems screen patients for depression using tools, such as the Beck Depression Scale (Beck et al., 1996), the Patient Health Questionnaire (PHQ-9) (Gilbody et al., 2007), the Cornell Scale for Depression in Dementia (Alexopoulos et al., 1988), and the Hamilton Rating |Scale for Depression (HAM-D) (Williams, 2001). However, there is a disjoint between patients who suffer from depression and patients acutely in need of care for depression. Despite the high prevalence of depression, most forms of depression are relatively mild and may be treated with medication. Only some forms of depression are far more severe and require advanced care above and beyond that provided by primary care providers (Cuijpers, van Straten, Andersson, & van Oppen, 2008a; Saloheimo et al., 2016b). Additionally, traditional depression screening approaches may increase the risk of over-diagnosis and over-treatment of depression across community and primary care settings (Dowrick & Frances, 2013; Mojtabai, 2013; Mojtabai & Olfson, 2011) without contributing to better mental health (Thombs, Ziegelstein, Roseman, Kloda, & Ioannidis, 2014). Recent studies have questioned the benefits of routine screening (Gilbody, Sheldon, & Wessely, 2006; Thombs et al., 2012) and questioned United States Preventive Services Task Force (USPSTF) recommendations to screen adults for depression in primary care settings when staff-assisted depression management programs are available (Thombs et al., 2014).

Due to such limitations, it is more clinically appropriate to develop screening approaches to identify patients suffering from serve cases of depression, patients who cannot be treated at primary care alone, and patients who would suffer from worsening health conditions unless they are provided with specialized care. Previously, we performed knowledge-based terminology extraction of the UMLS Metathesaurus (Bodenreider, 2004) to develop a depression terminology that represented a wide range of clinical, behavioral, and socio-economic determinants that impacted an individual's risk of suffering from depression. In this paper, we leveraged the knowledge-based depression terminology to integrate data obtained from various structured and unstructured clinical and administrative data sources to build decision models that identified patients in need of advanced care for depression.

## 5.2 Materials and Methods

### 5.2.1 Patient group

We extracted a population of 84,317 adult patients (≥ 18 years of age) with one or more outpatient visits at Eskenazi Health, Indianapolis. Indiana, between the years 2011-2016.

### 5.2.2 Patient subset selection

We sought to predict the need for advanced care for depression across: (a) the overall patient population and (b) different groups of high-risk patients. We selected three high-risk patient groups: Group A) patients with a prior diagnosis of depression, Group B) patients with a Charlson Comorbidity Index (Charlson et al., 2008) of >=1, and Group C) patients with a Charlson Comorbidity Index of >=2. Patients with a prior diagnosis of depression were identified as a high-risk group, as their illness may re-emerge or worsen based on other health conditions from which they may suffer. Patients with Charlson indexes >= 1 and >=2 were selected due to the high prevalence of depression across patients suffering from chronic illnesses, and its ability to worsen health outcome for patients. We also identified a fourth group (Group D), which comprised all unique patients identified in groups A-C. By identifying various patient groups, we hoped to: (a) capture as many of the overall number of patients in need of advanced care for depression and (b) compare which patient group was most suitable for use in screening for need of advanced care. Groups A-D were identified by scanning clinical data on each of the 84,317 patients (master patient list) to identify if they had a prior diagnosis of depression and calculate their individual Charlson Comorbidity Index.

#### 5.2.3 Preparation of gold standard

We defined patients in need of advanced care for depression as patients who had received a referral to a certified mental health provider for more specialized treatment for depression. We performed natural language processing of physician order notes to identify patients in the master patient list who had received such a referral/s.

#### 5.2.4 Data preparation

We obtained longitudinal health records on each patient from the Indiana Network for Patient Care (INPC), a statewide health information exchange server (McDonald et al., 2005; Overhage, 2016). The dataset included a wide array of patient data, including patient demographics, diagnosis, and visit data reported in both structured and unstructured form. All clinical data was obtained in the form of structured ICD-9 and ICD-10 codes. We assessed extracted data against the depression terminology and used relationships presented within the UMLS Metathesaurus to identify concepts for inclusion as features. Due to the distributed nature of clinical data available, we decided to categorize ICD-9 and ICD-10 codes.

We tabulated vectors for each patient group under study. In the event that the patient under study had received a referral for depression treatment, the data vector was only comprised of medical data recorded prior to the aforesaid referral. A master data vector encompassing all 84,317 patients was also created using the same approach.

### 5.2.5 Decision model building

We split each of the five data vectors (four patient groups and one master data vector) into random groups of 90% training data and 10% test data. Each training data set was used to train a decision model using the Random Forest classification algorithm (Breiman, 2001). The Random Forest algorithm was selected due to its track record of successful use in decision modeling for healthcare challenges, as well as its ability to perform feature selection. We used Python programming language (V 2.7.6) for all data pre-processing tasks, and the Python scikit-learn package for decision model development and testing (Pedregosa et al., 2011).

### 5.2.6 Analysis

Each decision model was evaluated using the 10% holdout test set. Results produced by each decision model were evaluated using Receiver Operating Characteristic (ROC) values, which measure classifier accuracy. Youden's J Index (Youden, 1950) was used to identify optimal sensitivity and specificity for each decision model.

A flowchart representing our workflow from patient group election to decision model evaluation can be seen in figure 5.1.

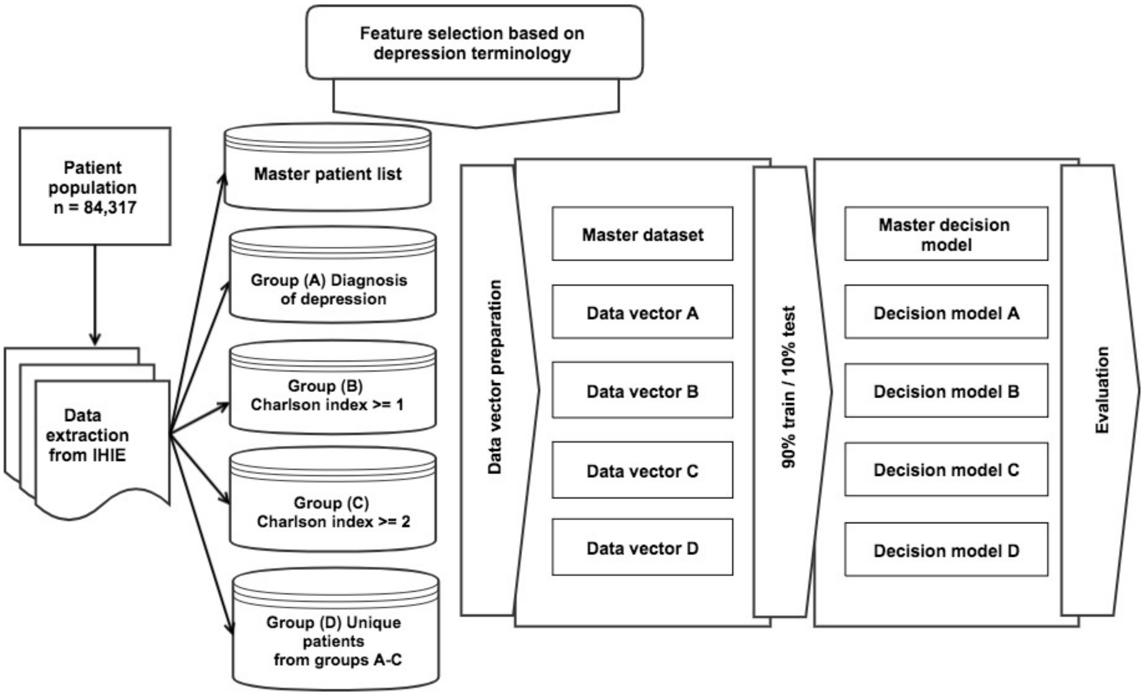



Figure 5.1 A flowchart representing our workflow from patient group election to decision model evaluation

## 5.3 Results

We identified a total of 12,432 patients with a diagnosis of depression (group A), 32,249 patients with a Charlson Index of 1 or greater (group B), and 7,415 patients with a Charlson Index of 2 or more (group C). Overall, these three groups identified a total of 37,560 unique patients (Group D).

The master patient list, as well as each of the four high-risk patient groups, represented an adult, urban population: predominantly female and with high disease burdens (table 5.1). The populations identified by their Charlson indexes were older (mean age >50 years) than the population identified with depression (46.31 mean age). Additionally, populations identified based on Charlson indexes were predominantly African-American. In contrast, the population identified by depression diagnosis were predominantly non-Hispanic whites. As anticipated, the prevalence of depression across a patient population with a Charlson Index of 1 or greater (30.18%) and a patient population with a Charlson Index of 2 or greater (37.25%) was greater than across the master patient list (19%).

Table 5.1. Characteristics of the master patient list/groups of high-risk patients used for decision model building

| Characteristic of interest | Master patient set | Group A | Group B | Group C | Group D |
|---|---|---|---|---|---|
| Definition | All patients | Patients with a prior diagnosis of depression | Patients with a Charlson Index of >= 1 | Patients with a Charlson Index of >= 2 | All unique patients in groups A-C |
| Patient group size (Patient group size as a % of the master patient list) | N= 84,317 | n=12,432 (14.74%) | n=32,249 (38.25%) | n=7,415 (8.8%) | n = 37,560 (44.5%) |
| Need of advanced care for depression (Need of advanced care for depression as a % of patients in each group) | n = 6992 (8.29%) | n = 3,683 (30.04%) | n = 4,016 (12.94%) | n = 1,026 (21.6%) | n = 5,612 (80.26%) |
| Demographics | | | | | |
| Age (mean, sd) | 49.3 (15.6) | 46.31 (14.74) | 51.94 (14.55) | 59.5 (12.33) | 50.3 (14.93) |
| Male gender | 35.1% | 30.22% | 39.8% | 43.98% | 42.03% |
| Race/ethnicity | | | | | |
| White (non-Hispanic) | 25.2% | 44.62% | 33.38% | 37% | 35% |
| African-American (non-Hispanic) | 37.2% | 32% | 42.78% | 47.26% | 40.12% |
| Hispanic or Latino | 19.5% | 11.12% | 10.6% | 4.94% | 7.38% |
| Diagnoses | | | | | |
| Depression (%) | 19% | 100% | 30.18% | 37.25% | 37.51% |
| Charlson Index score (mean, sd) | 0.8 (1.3) | 0.22 (0.75) | 1.89 (1.27) | 3.85 (1.14) | 1.62 (1.35) |
| Hospitalizations (mean, sd) | | | | | |
| ED visits during current month (mean, sd) | 0.21 (1.03) | 0.33 (1.48) | 0.26 (1.15) | 0.31 (1.14) | 0.27 (1.17) |
| ED visits prior to previous months (mean, sd) | 3.73 (14.40%) | 4.69 (18.73) | 8.63 (24.2) | 10.71 (31.36) | 8.03 (23.67) |

Figure 5.2 presents a Venn diagram presenting overlap across the high-risk patient groups identified for the study.

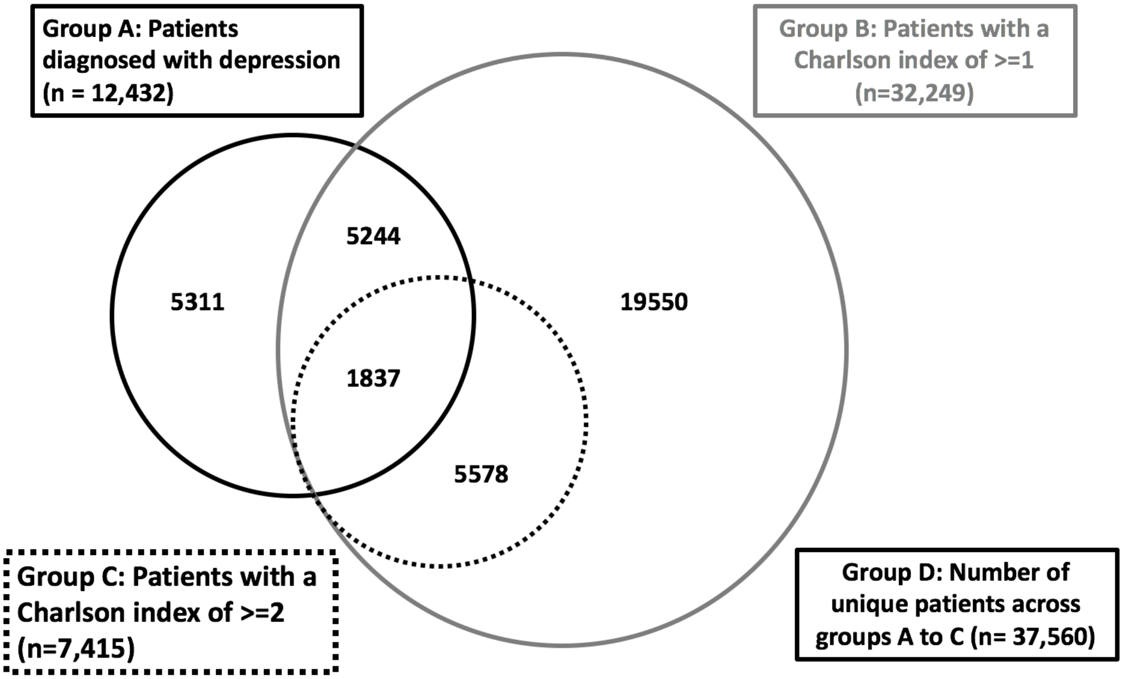



Figure 5.2 Overlap between the patient subsets identified for the study

A total of 6,992 (8.29%) of the 84,317 patients in the master patient list were in need of advanced care for depression. Group A captured 3,683 (52.68%) of these patients. Group B captured 4016 (57.43%) and group C captured 1,026 (14.67%). Overall, all three patient groups were able to identify 5,612 (80.26%) of all patients in need of advanced care for depression.

### 5.3.1 Feature selection using depression terminology

Previously, we developed a knowledge-based depression taxonomy using knowledge-based terminology extraction of the UMLS Metathesaurus. The taxonomy was developed by performing a literature search on Ovid Medline to identify publications that discuss depression and its treatment, and then using Metamap (Aronson & Lang, 2010), a Natural Language Processing (NLP)-based tool to map these abstracts against the UMLS Metathesaurus (U.S. National Library of Medicine, 2016), a large, multipurpose, multi-lingual thesaurus that contains millions of biomedical and health-related concepts, synonymous names, and their relationships across 199 medical dictionaries (U.S. National Library of Medicine, 2015). The most frequently occurring UMLS concepts were compiled into a terminology using the Web Ontology Language (OWL), a semantic web language that is widely used to represent ontologies.

Comparison of patient data against the aforementioned depression terminology resulted in the identification of 1,150 unique concepts for inclusion in each decision model. A detailed list of features included in each of the decision models is presented in Appendix 9.4.

#### 5.3.2 Decision model performance

The decision model predicting need of advanced care across the master population reported a moderate ROC score of 78.87% (optimal sensitivity = 68.79%, optimal specificity = 76.30%). However, decision models to predict need of advanced care across patients' groups A-D performed significantly better. Group A (patients with a prior diagnosis of depression) reported an ROC score of 87.29% (optimal sensitivity = 77.84%, optimal specificity = 82.66%). Group B (patients with a Charlson Index of >= 1) reported an ROC score of 91.78% (optimal sensitivity = 81.05%, optimal specificity = 89.21%). Group C (patients with a Charlson Index of >= 2) reported an ROC score of 94.43% (optimal sensitivity = 83.91%, optimal specificity = 92.18%), while Group D (list of unique patients from groups A-C) reported an ROC score of 86.31% (figure 5.3) (optimal sensitivity = 75.31%, optimal specificity = 76.03%). Precision-recall curves for each decision model are presented in Appendix 9.5.

The top 20 features for each decision model, as ranked using LASSO scores, can be seen in Appendix 9.6. Appendix 9.7 presents the co-occurrence of these top 20 features across each decision model under study. In assessing the top ranked feature selected by LASSO for each decision model, we found significant overlap among the top features for each of the high-risk patient populations. However, Essential (primary) hypertension, Depression, Gender, and Number of outpatient visits appears in the top 20 feature lists for every patient population under test.

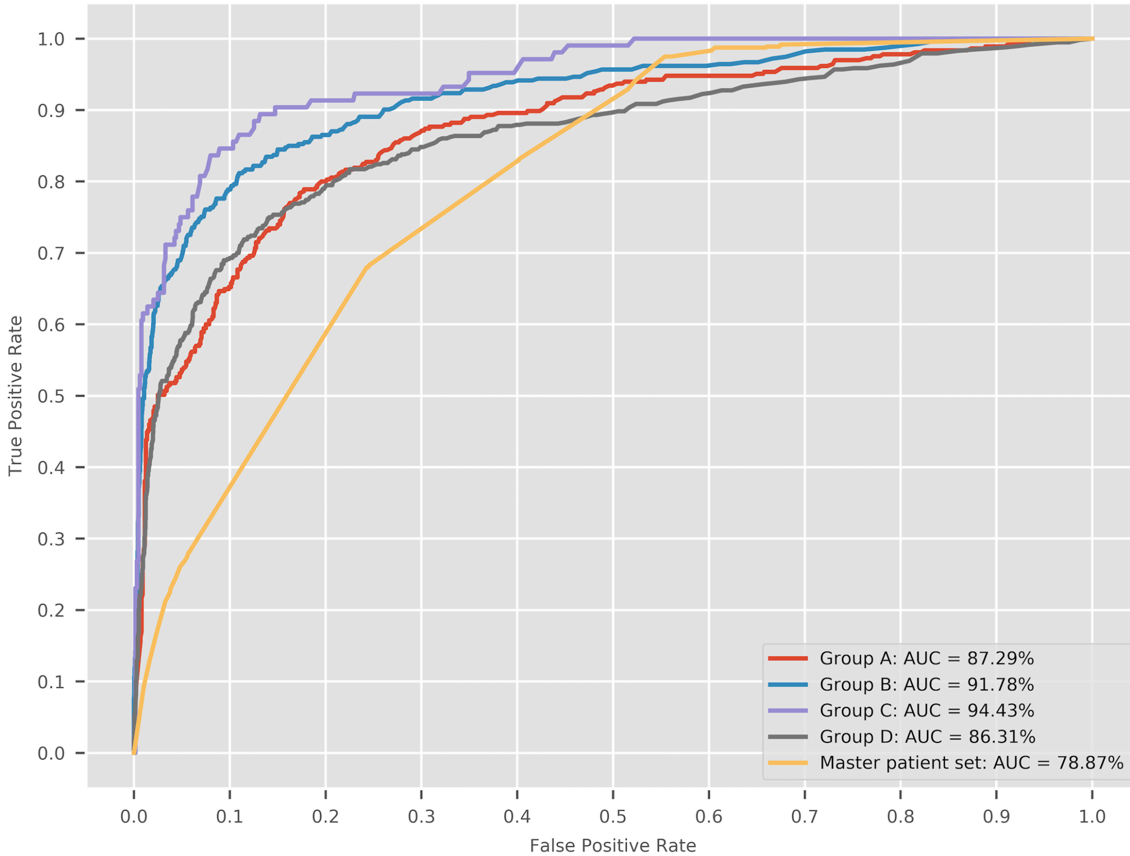



Figure 5.3 ROC curves produced by decision models predicting need of advanced care across each patient group under study

## 5.4 Discussion

The decision model to predict need of advanced care for depression across the entire patient population achieved an ROC score of 78.87%. In comparison, decision models that predicted need of advanced care across four high-risk patient populations performed better, with ROC scores ranging from 86.31%-94.43%. Additionally, optimal sensitivity and specificity for each decision model was significantly high, and demonstrated the models' potential for practical implementation.

We attribute the comparatively low performance of the decision model developed using the master patient set to: (a) the unbalanced nature of the gold standard (Khalilia, Chakraborty, & Popescu, 2011) caused by the relatively low prevalence (8.29%) of patients in need of advanced care in comparison to the overall patient population and (b) the sparsity of clinical data available for some of the patients in the overall patient population, which affected the robustness of features used in the decision model. The high performance of the decision models built using high-risk patient groups could be attributed to the balanced nature of the gold standard due to the higher prevalence of patients in need of advanced care for depression across each of the high-risk patient groups. While various publications have presented approaches to address data imbalance, we did not pursue such as approach, as we wished to focus on demonstrating methods that would be useful to potential implementers who could replicate our efforts across other datasets, which may or may not be imbalanced.

In assessing prediction performance, group C (patients with a Charlson Index >= 2) yielded the highest AUC score (97.43%). Groups A (patients with a diagnosis of depression) and B (patients with a Charlson Index >= 1) reported lesser AUC scores. However, group C captured the least number of patients in need of advanced care in comparison to groups A and B. However, it is noteworthy that none of the decision models developed using high-risk populations could capture all patients in need of advanced care. Overall, all three models could capture only 80.26% of all patients in need of advanced care. The remainder (19.74%) of the patients in need of advanced care did not qualify for any of the three high-risk patient populations that we considered. We hypothesize that a share of the missing 19.74% patients should have fallen into one of the three high-risk patient groups, and thus been eligible for detection.

We present a novel application of machine learning to address a question of significant clinical relevance. We demonstrated the ability to predict need of advanced care for depression across high-risk patients using a host of clinical and behavioral determinants with considerable levels of accuracy. These efforts can be adapted to other challenges, such as improving referrals to wraparound mental health services (Winters & Metz, 2009). Since wraparound services are not delivered by primary care providers (Loeb, Binswanger, Candrian, & Bayliss, 2015), the ability to identify and refer patients in need of such services is extremely important (Jackson et al., 2013). Our efforts yield a highly accurate, automated approach for identifying patients in need of wraparound services for mental health, which is of growing importance to healthcare organizations and incentivized by changing reimbursement policies. By predicting the need for advanced care across various high-risk populations, we offer potential implementers the option of selecting the best screening approach that meets their needs based on any practical limitations. Furthermore, our approach presents the ability to effectively identify the need for advanced care for depression without the risk of over-diagnosis and over-treatment, and without the use of existing screening mechanisms.

In the current study, we adopted a binary (present/absent) outcome flag for each feature used to train decision models. We hypothesize that switching to tabulated counts for each feature will increase the granularity of the feature vector, thereby increasing model performance. Additionally, our study did not consider temporality (time between outcome of interest and clinical condition). Inclusion of temporality measures within the model may further improve performance. Finally, our decision models were developed using patient-based clinical, behavioral, and demographic data. Research presents significant evidence related to the role of various population-based socio-economic and environmental determinants in influencing the progression of mental health and depression (Allen et al., 2014; Patel et al., 2010; Tanner, Martinez, & Harris, 2014; World Health Organization, 2014). We hope to assess the value of including such population measures in improving decision model performance.

We also identified several limitations in our study approach. The patient group used in our study was obtained from the Eskenazi Health system, a safety-net population with significant health burdens. Thus, our models may not generalize to other commercially insured or broader populations. Our clinical and behavioral data was limited to diagnosis data extracted in the form of ICD codes. Using a wider range of clinical data on medical procedures, etc., may further improve decision model performance. None of the three patient groups captured all of the patients in need of advanced care for depression. The closest, group B (patients with a Charlson Index >=1), captured only 78.94% of all patients in need of care. We hypothesize that some patients in need of advanced care for depression were not included in any of the three models, as some of these patients may have missed data elements that would have categorized them as such. Additional efforts to identify different groupings that lead to inclusion of more patients in need of advanced care may be warranted. Our approaches to identifying the most relevant features for model selection resulted in the top ranked features listed in Appendices F and G. However, we note that in the event of highly-correlated variables, the LASSO feature selection tends to select one feature over another, arbitrarily ignoring the second feature (Fonti & Belitser, 2017). This may have resulted in the exclusion of some features that were as relevant, but highly correlated with other features.

Finally, studies present that social determinants of health, such as low-educational attainment, poverty, unemployment, and social isolation, may have a significant impact on depression and the need for treatment (Fryers, Melzer, Jenkins, & Brugha, 2005; World Health Organization, 2014). We propose to expand our models using social determinants of health to assess their impact on decision model performance.

In conclusion, these results show considerable potential for enabling preventative care and can be easily adopted to existing clinical workflows to improve access to wraparound healthcare services.

## 5.5 Conclusion

Our efforts demonstrate the ability to identify patients in need of advanced care for depression across: (a) an overall patient population and (b) various groups of high-risk patients using a wide range of structured and unstructured data that represent a patient's clinical and behavioral determinants. While all models yielded significant performance accuracy, models focused on high-risk patient populations yielded comparatively better results. Furthermore, our methods present a replicable approach for implementers to follow based on their own needs and priorities. However, decision model performance may differ based on the availability of patient data at each healthcare system. These results show considerable potential for enabling preventative care and can be easily adopted to existing clinical workflows to improve access to wraparound healthcare services.

# 6 AIM 3: THE IMPACT OF PATIENT- AND AGGREGATE-LEVEL SOCIAL DETERMINANTS OF HEALTH IN PREDICTING NEED OF ADVANCED CARE FOR DEPRESSION

## 6.1 Introduction

Social determinants of health (SDH) are defined as "the structural determinants and conditions in which people are born, grow, live, work and age" (Marmot et al., 2008). They include factors such as an individual's socio-economic status, education, neighborhood, and physical environment; employment, education, and access to various services (Braveman et al., 2011; Marmot, Allen, Bell, Bloomer, & Goldblatt, 2012; World Health Organization, 2010). SDH plays a significant role in susceptibility (Skinner, 2011), progression, and outcome of disease and illness (Cullen, Cummins, & Fuchs, 2012; Di Monte, 2003; Yu, Lin, & Lin-Tan, 2004). Research indicates that SDH influences the likelihood of smoking and shorter life expectancy (Marmot et al., 2008) and may have multi-generational impacts (Chetty, Hendren, Kline, & Saez, 2014).

There have been significant limitations in the use of social determinants of health (SDH) for improving healthcare delivery. Historically, SDH was collected for use by public health, environmental, and social services, and curated at aggregate geographical levels (Comer, Grannis, Dixon, Bodenhamer, & Wiehe, 2011). Challenges in collecting, managing, and using such data, as well as integrating them with clinical sources, has prevented the use of SDH in informing healthcare delivery. However, these barriers are coming down. Rising interest in the field of precision medicine has highlighted the role of social determinants on an individual's health (Martin Sanchez, Gray, Bellazzi, & Lopez-Campos, 2014) and has led to increasing support for capturing such data in clinical information systems using standard clinical terminology (LaBrec, 2016). Increasing support for interoperability and Health Information Exchanges (HIE) (Vest & Gamm, 2010) are enabling better integration of SDH captured outside the health information eco-system.

The predictive modeling domain poses significant potential for leveraging SDH data in support of organizational planning, research, and healthcare policy activities (Bates, Saria, Ohno-Machado, Shah, & Escobar, 2014; Kansagara et al., 2011; Schneeweiss, 2014). To date, there has been limited efforts to capitalize on SDH for predictive modeling. However, studies that do leverage SDH for predictive modeling are restricted to a minimum number of social determinants that do not present a complete picture of an individual's environment (Torres et al., 2015). Additionally, SDH may not be equally relevant or as impactful across all healthcare challenges. Thus, attempts to demonstrate the value of SDH in improving healthcare delivery must focus on healthcare challenges that are strongly influenced by SDH.

We sought to determine the value of SDH in predictive modeling activities for the treatment of depression, the most common type of mental illness the world over (Ferrari et al., 2013). Depression affects over 26% of the U.S. adult population (Kessler et al., 2005). It is a leading cause of disability for Americans age 15-44 years and is responsible for almost 400 million disability days per year (Mathers, 2008). The incremental economic burden of depression (covering medical, pharmaceutical, workplace, and suicide-related costs) in the U.S. was evaluated at $210.5 billion in 2010, a 21.5% increase from 2005 (Greenberg et al., 2015). The onset, progression, and outcome of depression is significantly impacted by an individual's socio-economic environment (Saveanu & Nemeroff, 2012; Wong, Dong, Andreev, Arcos-Burgos, & Licinio, 2012), making it an ideal use case to evaluate the value of SDH in healthcare decision making.

In aim 2, we leveraged a host of clinical and behavioral data extracted from a statewide HIE to build decision models that predicted need of advanced care for depression. While these models yielded significantly high performance, they were an infeasible solution for healthcare facilities without access to a high volume of clinical data that could only be obtained from a robust HIE. In comparison, SDH data that describe the overall health of an individual and community may be easily available and could be used to augment the accuracy of such decision models. Additionally, models that leverage SDH may be easily interpretable, and thus, be of additional value to the healthcare community. In this paper, we evaluate the ability to: (a) implement conceptual SDH frameworks with patient- and population-centric data for use in decision-model building, (b) leverage SDH data to improve performance of decision models that predict need for advanced care for depression, and (c) identify the minimal number of clinical, behavioral, and SDH features required to build actionable decision models.

## 6.2 Materials and methods

### 6.2.1 Patient subset selection

This study leveraged the same patient sets that were prepared for aim 2. We identified a population of 84,317 adults (≥ 18 years) with >=1 outpatient visit at Eskenazi Health of Indianapolis, Indiana, between the years 2011-2016. Eskenazi Health serves as the public safety-net health system of Marion County, Indiana.

From this master patient population, we identified three patient subsets representing high-risk patient populations: Group A) patients with a prior diagnosis of depression, Group B) patients with a Charlson Comorbidity Index (Charlson et al., 2008) of >=1, and Group C) patients with a Charlson Comorbidity Index of >=2. Patients with a prior diagnosis of depression were identified as being at high-risk for need of advanced care for depression given that their illness may re-emerge or worsen based on other health conditions from which they may suffer. Patients with Charlson indexes >= 1 and >=2 were selected due to the high prevalence of depression across patients suffering from chronic illnesses, and its ability to worsen health outcomes for patients. Groups A-C were identified by scanning clinical data on each of the original 84,317 patients to identify if they had a prior diagnosis of depression and calculate their individual Charlson Comorbidity Index.

### 6.2.2 Preparation of clinical and behavioral datasets

Longitudinal health records on each patient were extracted from the Indiana Network for Patient Care (INPC), a statewide health information exchange server (McDonald et al., 2005; Overhage, 2016). The dataset included a wide array of patient data, including patient demographics, diagnosis, and visit data.

We assessed extracted data against the depression terminology and used relationships presented within the UMLS Metathesaurus to identify concepts for inclusion as features. Due to the distributed nature of clinical data available, we decided to categorize ICD-9 and ICD-10 codes.

We developed data vectors for each of the three high-risk patient groups under study. In the event that the patient under study had received a referral for depression treatment, the data vector was only comprised of medical data recorded prior to the aforesaid referral. A master data vector encompassing all 84,317 patients was also created using the same approach. These datasets would hereafter be referred to as clinical data vectors.

#### 6.2.3 Preparing social determinants of health

We leveraged the Kaiser Family Foundation (KFF) social determinants framework (Heiman & Artiga, 2015) to identify and categorize SDH's of interest. The KFF framework categorizes social determinants into: (a) economic stability, (b) neighborhood and physical environment, (c) education, (d) food, (e) community and social context, and (f) healthcare system-based factors.

SDH data were obtained from two sources: (a) patient-centric data extracted from codified INPC data based on a list of ICD-10 codes representing SDH (LaBrec, 2016) that were mapped to ICD-9 using General Equivalence Mappings (GEM) (Butler, 2007) (Appendix 9.8) and (b) public health and environmental data obtained from the Polis Center, a reputed research organization that seeks to understand the communities that people live in; developers of SAVI, the nation's largest community information system; and the Indiana State Department of Health. We obtained a total of 14 SDH from the INPC and 48 features from the Polis data (Appendix 9.9). The Polis data were aggregated across 224 census tracts and 20 HPA. Data obtained from the Polis Center was linked to patients via the census tract identified from each patient's address. For machine-learning purposes, we reduced the high variability of the SDH data by classifying each feature into smaller bins determined by Sturges' rule (Scott, 2009).

#### 6.2.4 Preparation of gold standard

Patients in need of advanced care for depression were defined as those individuals who had received a referral to a certified mental health provider for more specialized treatment for depression. We performed natural language processing of physician order notes to identify patients in the master patient list who had received such a referral/s.

### 6.2.5 Decision model building

In aim 2, we found that the need for advanced care for each of the groups under testing could be predicted with significantly high accuracy using significant quantities of non-SDH data, alone. However, as discussed previously, not all healthcare facilities are able to produce the volume of clinical and behavioral data used in our previous efforts. We sought to assess the minimal number of features necessary to build accurate models and assessed whether the use of socio-economic data would significantly improve performance.

We identified the most valuable patient-centric features across each of the four data vectors under testing using the Least Absolute Shrinkage and Selection Operator (LASSO) (Tibshirani, 1996). Using this approach, we produced ranked feature lists.

Next, for each dataset, we used the ranked feature lists to build multiple decision models starting from a decision model that included the five top ranking features, iteratively adding on the next important feature until each model was comprised of the 50 top ranking features. The performance of each decision model developed using the top n ranking clinical and behavioral features were evaluated against a decision model that incorporated the top n clinical features and all SDH features. As an example, for patient group A, we began by building a decision model consisting of five patient-centric features. The performance of this model was compared to a decision model built using the same five features plus 50 socio-economic features. Next, we added the next most important patient-centric feature and built a decision model consisting of these six features. This model was compared to a decision model built using the same six features plus 50 socio-economic features. We continued building models until we had included the top 50 most relevant patient-centric features. Thus, for patient group A, we built a total of 90 ([50-5]*2) decision models.

Each training dataset was used to build a decision model using the random forest classification algorithm. We selected random forest (Breiman, 2001) for decision-model building due to its proven track record of use in decision modeling for healthcare challenges.

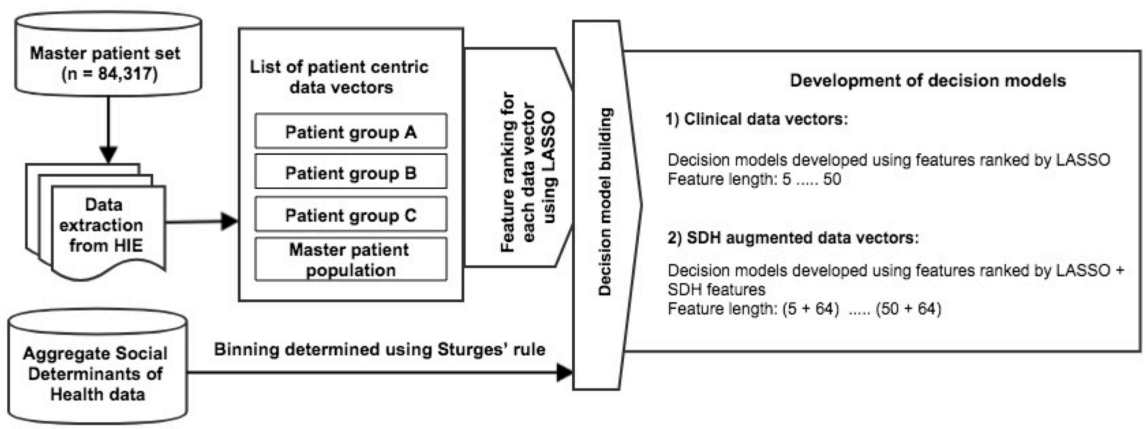



Figure 6.1 Workflow depicting the study approach, from data collection to decision-model building

All data cleaning, decision-model development, and testing was performed using python and scikit-learn software (Pedregosa et al., 2011).

#### 6.2.6 Analysis

Each decision model was trained and tested using 10-fold cross validation (Zhang, 1993). For each decision model, we calculated accuracy in the form of F1-scores, together with 95% confidence intervals.

### 6.3 Results

The master patient list, comprised of 84,317 patients, represented a predominantly female adult population with high disease burdens (table 5.1). The patient population was distributed across 37 zip codes and 224 census tracts. Group A (prior diagnosis of depression) comprised 12,432 patients. Group B (Charlson Index of 1 or greater) comprised 32,249 patients, while group C (Charlson Index of 2 or greater) comprised 7,415 patients. Overall, each of the patient subgroups under test was also evenly distributed geographically (Appendix 9.10). In evaluating the data obtained from the INPC, we found that a considerable number of patients had social data recorded in the form of ICD-9 and ICD-10 codes 9 (Appendix 9.8).

### 6.3.1 Decision model performance

The performance of decision models for each patient population under test are presented in the form of F1 scores and 95% confidence intervals against the feature set size used. (figures 6.2 to 6.5).

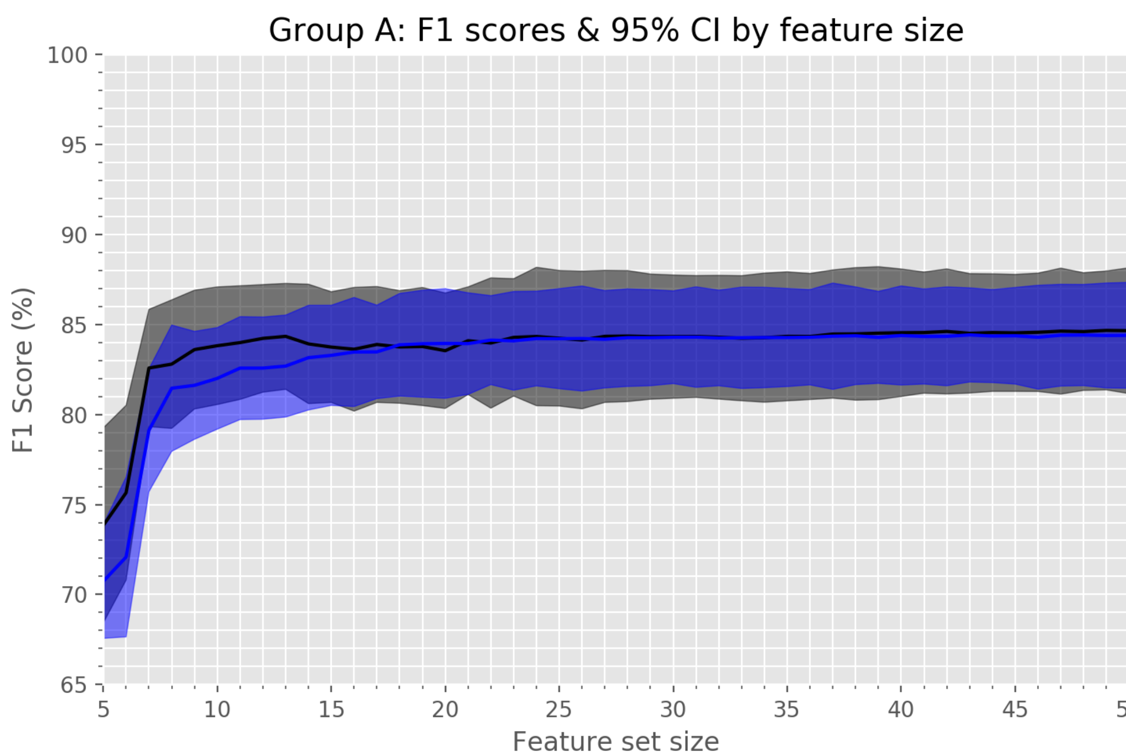



Figure 6.2. Group A: F1-scores and 95% confidence intervals reported by feature set size

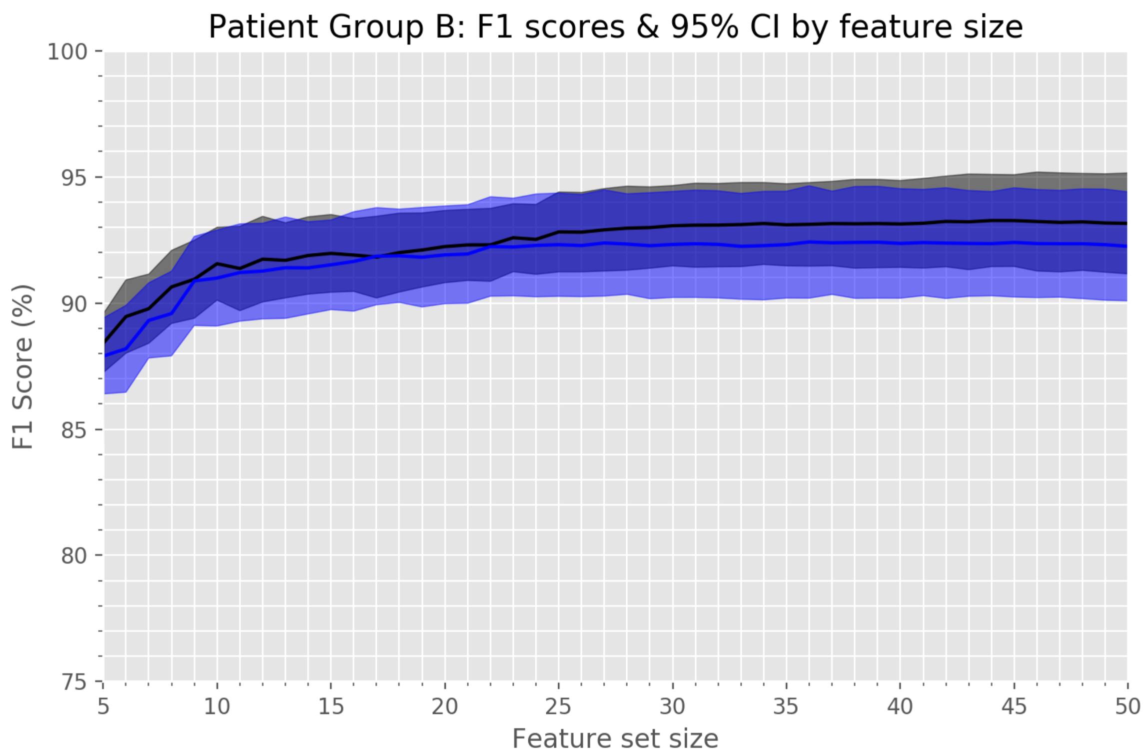



Figure 6.3. Group B: F1-scores and 95% confidence intervals reported by feature set size

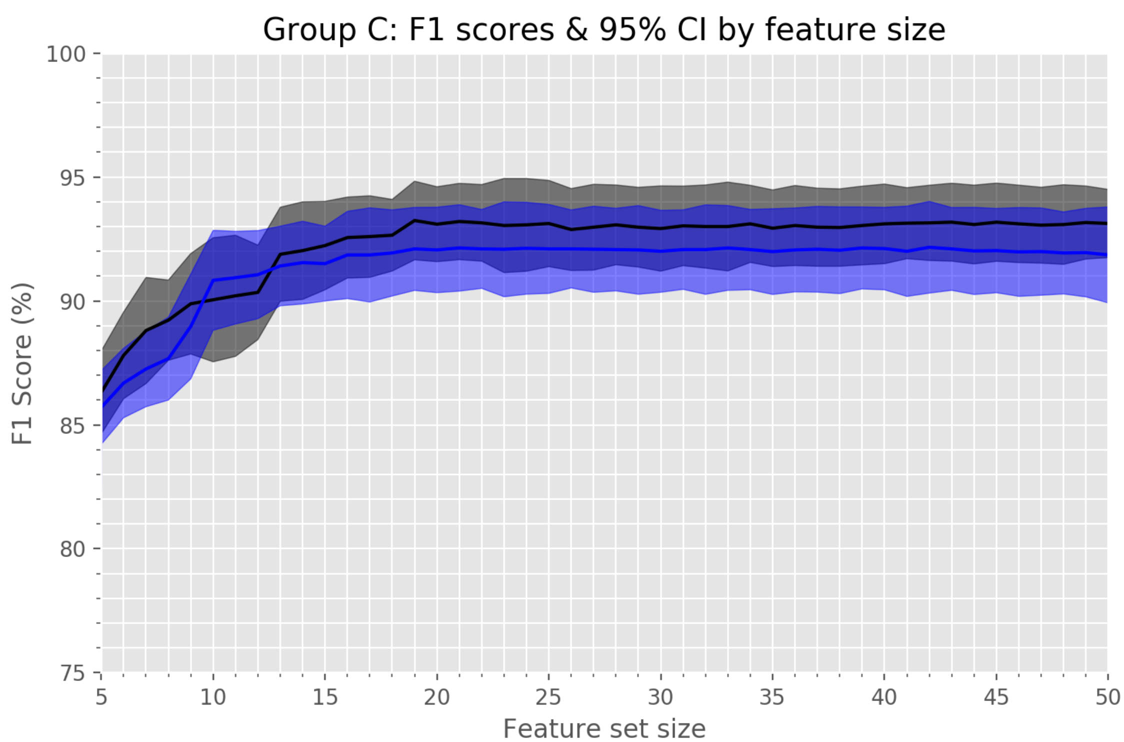



Figure 6.4. Group C: F1-scores and 95% confidence intervals reported by feature set size

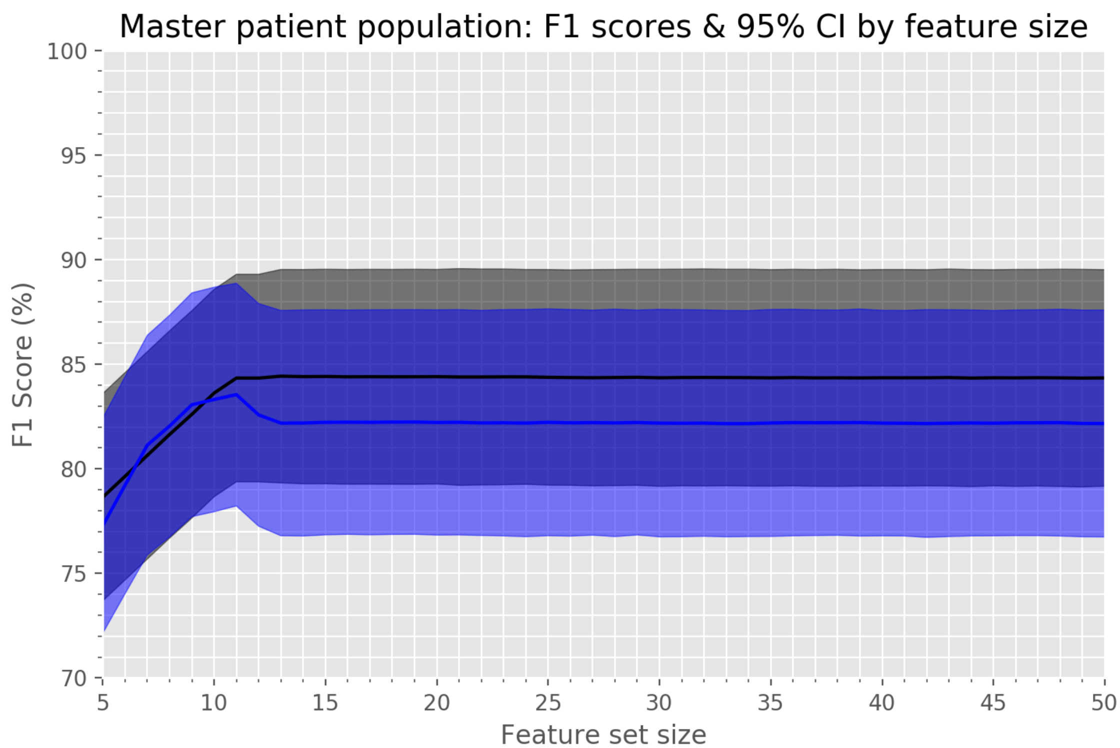



Figure 6.5. Master patient population: F1-scores and 95% confidence intervals reported by feature set size

For group A, F-1 scores for decision models developed using clinical features selected in ascending order of importance started with 73.83% for decision models using just one feature, and rose to ~84% in models using top ~10 clinical features, at which point they remained constant. Models augmented with SDH features followed a similar pattern, with f1 scores starting at 70.73% for decision models with just one clinical feature, rising to ~85%, and then remaining constant as further features were increased. There was no statistically significant change in model performance at ~83% f1- score.

For group B, F-1 scores for decision models developed using clinical features selected in ascending order of importance, starting with 88.40% for decision models using just one feature, rising to ~92% in models using top ~20 clinical features, at which point they remained constant. Models augmented with SDH features followed a similar pattern, with f1-scores starting at 87.16% for decision models with just one clinical feature, rising to ~92% in models using the top ~15 features, and then remaining constant as further features were increased. There was no statistically significant change between models using clinical or SDH data.

For group C, F-1 scores for decision models developed using clinical features selected in ascending order of importance started with 86.30% for decision models using five features, rose to ~84% in models using top ~15 clinical features, at which point they remained constant. Models augmented with SDH features followed a similar pattern, with f1-scores starting at 85.70% for decision models with just five clinical features, rising to ~92% in models using the top ~15 features, and then remaining constant as further features were increased. There was no statistically significant change between models using clinical or SDH data.

For the master patient population, F-1 scores for decision models developed using clinical features selected in ascending order of importance started from x% for decision models using five features, rose to ~92% in models using the top ~11 clinical features, at which point they remained constant. Models augmented with SDH features reported f1-scores starting at y% for decision models with just five clinical features, and peaking at ~84% in models using the top ~11 features. At this point, F-1 scores fell back to ~82%, and remained constant as further features were increased. However, there was no statistically significant change between models using clinical or SDH data.

The top 20 features for each decision model, as ranked using LASSO scores, can be seen in Appendix 9.6. Appendix 9.7 presents the co-occurrence of these top 20 features across each decision model under study.

## 6.4 Discussion

This study indicates the ability to predict the need for advanced care for depression across: (a) different groups of high-risk patients and (b) the overall patient population with significant accuracy using only a minimal number of clinical, behavioral, and social determinants of health. By building decision models using varying numbers of clinical and behavioral features ranked by descending order of importance, we found that decision model performance did not improve after the inclusion of the first ~15 top-ranked clinical features (Appendix 9.6). The inclusion of SDH into decision models built using clinical and behavioral data did not lead to any statistically significant performance improvements.

We hypothesize that the inclusion of patient and aggregate-level SDH did not lead to performance improvements due to several limitations: (a) weaknesses in our study population, which represented a high-risk population with high disease burdens spread across a relatively constrained safety net geography. A larger, more socially diverse patient population may represent greater variance, and thus, lead to greater improvements in decision model performance. (b) Aggregate socio-economic features were assigned based only on an individual's current address. We did not consider changes in an individual's residence over time, nor changing exposures to different socio-economic environments caused by an individual's occupation and/or movement during their day-to-day life. The aggregate SDH indicators used in this study may be improved by obtaining SDH data aggregated at a finer level of granularity than census tract and HPA level. Our results may also benefit from the inclusion of additional data sources, such as those available via Zillow (programmableweb.com, 2017), which may capture more up-to-date data elements; (c) a significant part of the clinical data used in our study was collected via ICD-9 codes, which did not support a wide range of socio-economic data. A more up-to-date clinical dataset captured using only ICD-10 codes may incorporate more patient-centric socio-economic data, and thus lead to improvements in decision-model performance.

Our study leveraged only a subset of potential clinical, behavioral, and SDH features. Other patient level features may be extracted from sources, such as wearable devices and activity trackers (Almalki, Gray, & Sanchez, 2015; Evenson, Goto, & Furberg, 2015). Finally, our decision models did not consider data temporality/sequence of each event in performing prediction. Adopting a sequence-based deep learning algorithm, such as Long Short-Term Memory (LTSM) networks (Lipton, Kale, Elkan, & Wetzel, 2015) that identified patterns in sequential data, may improve decision-model performance.

Our efforts present a comprehensive effort to implement a conceptual SDH framework by extracting SDH from patient-centric and aggregate-data sources captured within and outside the healthcare domain and leverage them for machine learning purposes. While the use of SDH indicators did not result in improvements in this specific scenario, our approach may be replicated across other populations and different socio-economic features, where the use of SDH may yield significant improvements. Our methods included implementation of the conceptual KFF framework on SDH. However, our methods may also be applied to other conceptual SDH frameworks, such as the Danaher framework (Danaher, 2011), the Rural Community Health & Well-Being Framework (Annis, Racher, & Beattie, 2004), the World Health Organization's conceptual framework for action on SDH (Solar & Irwin, 2010), and the Area-based deprivation indices (ABDI) (Schuurman, Bell, Dunn, & Oliver, 2007), which may be more suitable for other healthcare challenges based on location and data available for use. Additionally, there is a need for more comprehensive research on how existing SDH frameworks can be leveraged to capture the element of temporality, covering changes in an individual's: (a) residence over time, (b) day-to-day movement across different geographies, and (c) occupational exposures, thus making them more suitable for use in machine-learning tasks.

## 6.5 Conclusion

It is possible to integrate patient and aggregate-level SDH with patient-level clinical and behavioral data for effective decision-model building for predicting the need for advanced care for depression. Our efforts constitute a comprehensive plan to implement a conceptual SDH framework by extracting SDH from patient-centric and aggregate data sources captured within and outside the healthcare domain, and leveraging them for machine-learning purposes. Decision models built using the aforementioned data yielded significantly high performance measures using only a relatively small number (~15) of most relevant features. The use of SDH did not lead to any statistically significant performance improvements in any of the decision models under test. However, our methods can be applied to a multitude of other healthcare questions, where the inclusion of SDH may yield improved results. There is an urgent need for research and implementation efforts around capturing and sharing patient- and population-centric SDH, as well as better approaches to include the element of temporality in modelling SDH for machine-learning purposes.

# 7 DISCUSSION

We sought to investigate alternative approaches to improve the detection of patients in need of advanced care for depression by leveraging existing knowledge, information sources, clinical workflows, and data analytics. Our efforts were be broken into three incremental aims.

In aim 1, we developed a comprehensive depression terminology that covered a wide range of clinical, behavioral, and SDH that influenced the need for advanced care for depression. The depression terminology was developed by performing an extensive literature search for publications that discuss depression and its treatment, using Metamap to extract UMLS concepts associated with the most frequently recurring conditions from the UMLS Metathesaurus. The depression terminology comprised 7,402 concepts representing patient-level clinical, behavioral, and social determinants of health that were mapped to 45,297 UMLS concepts from across 70 different terminologies, as well as 64 additional concepts representing aggregate-level social determinants of health. Together, these concepts represented a wide range of features that influenced the onset, progression, and outcome of depression.

In aim 2, we leveraged the depression terminology to extract comprehensive longitudinal patient data from the INPC, a large, statewide HIE. By analyzing a population of 84,317 patients, we identified several high-risk populations for predicting need of advanced care for depression: (a) patients with a prior diagnosis of depression (n=12,432), (b) patients with a Charlson Index >= 1 (n=), and (c) patients with a Charlson Index of >=2 (n=7,415), as well as the overall patient population (n=84,317). We leveraged the depression terminology developed in aim 1 to extract a wide range of clinical data covering both chronic and acute conditions, as well as behavioral data. We. built a series of decision models that predicted need of advanced treatment for depression using these features.

In aim 3, we assessed the potential to improve the performance of decision models built in aim 2 by integrating a host of patient- and aggregate-level SDH into existing datasets. However, since the models developed in aim 2 had already achieved significantly high levels of performance, we assessed the contribution of SDH by building incremental decision models consisting of the top n clinical and behavioral features and comparing their performance against models built using the same number of features plus SDH. We found that: (a) performance of decision models plateaued after ~15 clinical features, and (b) inclusion of patient and aggregate-level SDH did not lead to statistically significant performance improvement in any of the models.

Overall, our efforts contribute to best practices on the use of data-driven decision making and existing medical knowledge to inform actionable, cost-efficient approaches to influence access to healthcare delivery. Through the use of existing medical knowledge, as derived from the depression terminology for decision-model building, we seek to close the gap on the chasm between healthcare and machine-learning domains, thereby leading to an approach that is acceptable to all stakeholders. Our efforts are timely given: (a) increasing interest in big data and precision medicine, which has led to the availability of a glut of data sources of varied value to healthcare (Gligorijević, Malod-Dognin, & Pržulj, 2016; Kasthurirathne, Biondich, Mamlin, Cullen, & Grannis, 2017) and (b) falling costs of computational storage, memory, and processing bandwidth (McAfee, Brynjolfsson, Davenport, Patil, & Barton, 2012), as well as a rapidly evolving machine learning domain (Schmidhuber, 2015) that has led to a multitude of neural network and deep learning-based solutions that seek to address healthcare challenges (Liang, Zhang, Huang, & Hu, 2014; Miotto, Wang, Wang, Jiang, & Dudley, 2017).

In terms of big data and precision medicine, our findings help inform the identification of the most pertinent data sources based on accessibility, completeness, and value to precision medicine. However, much of these improvements come at the cost of clarity and simplicity; neural networks are exceedingly complex and difficult to interpret (Bengio, Courville, & Vincent, 2013; Schmidhuber, 2015), which drives a wedge between rich sources of medical data and healthcare workers who stand to benefit from them. In contrast, our machine-learning efforts prioritized interpretability and simplicity of decision models. These qualities promise healthcare practitioners a more acceptable path to benefit from these technological advances. The use of existing knowledge in the form of published research also addresses the research-to-implementation gap (Brownson, Colditz, & Proctor, 2017), enabling stakeholders to benefit from breakthroughs in research at a faster pace.

Our focus on informing referrals to supplementary/wraparound services are in line with changing reimbursement policies that incentivize preventive health care (Casalino et al., 2003; McGinnis & Foege, 1993). Additionally, our approaches demonstrate the value of comprehensive data collection and sharing for medical decision making, from patient referrals and previous visit history to aggregate-level SDH. The use of non-traditional administrative data, as well as public health/socio-economic data demonstrates, diverge from the usual confines of the "clinical" domain and build a platform for collaboration among a wider range of stakeholders.

In aims 2 and 3, we attempted to predict the need for advanced care for depression across: (a) various subsets of high-risk patient populations and (b) the overall patient population. In each case, we found that decision models assessing high-risk patient subsets yielded better results than models built using the overall patient population. We hypothesize that the overall patient population failed to achieve better performance due to: (a) missing clinical observations leading to sparse data vectors, which may decrease machine-learning performance and (b) a relatively low prevalence of need for advanced care for depression across the overall population, leading to an imbalanced gold standard for machine learning (Prati, Batista, & Monard, 2009).

Our approach developed decision models for a number of high-risk populations. These populations represent easily identifiable groups based on a subset of easily discernable clinical conditions and disease burdens. However, it is unclear whether these populations are the most clinically relevant given the data at hand. Thus, we recommend further research using machine learning approaches, specifically clustering approaches, such as K-means (Hartigan & Wong, 1979) or SOM (Kohonen, 1982), to identify more suitable patient populations based on the data under study. Predicting the need for advanced care using decision models for other, more clinically relevant populations, may yield better model performance using a smaller subset of features. Groupings of the most important features, as presented in Appendices H and G, also suggest the possibility of developing more clinically relevant patient subsets.

As discussed previously, we selected the Random Forest algorithm for decision-model building based on the need to develop high performance models that were easily interpretable to our clinical audience (Qi, 2012). Other, more advanced decision-modeling approaches, such as neural networks (Haykin & Network, 2004), have shown potential to improve marching learning performance across other healthcare-related challenges. However, neural networks are more cost intensive and yield results that are not easily interpretable (Bengio et al., 2013). Additionally, it is unclear if they can contribute to our study, given the significant levels of accuracy already achieved using Random Forest models. We recommend that neural networks be considered in a scenario where sequence/temporality of clinical events are being evaluated or where performance of Random Forest models are unsatisfactory.

In aim 3, we found that the inclusion of patient and aggregate-level SDH did not lead to improvements in decision model performance. However, the informatics community should not conclude that there is no value in including SDH for improving healthcare decision making based on these results. We believe that SDH will continue to play an integral role in influencing healthcare, as the collection and use of patient and aggregate-level SDH increases. We applaud the Institute of Medicine (IOM)'s decision to recommend that SDH be captured in Electronic Health Records (Institute of Medicine. Committee on the Recommended Social Behavioral Domains Measures for Electronic Health Records, 2014), as well as the International Classification of Disease (ICD) for enabling SDH collection by introducing additional SDH codes to ICD-10 (LaBrec, 2016). The Logical Observation Identifiers Names and Codes (LOINC) terminology is also engaged in incorporating SDH codes. however, we did not consider LOINC codes due to limited availability in the INPC dataset. We anticipate that the availability of patient-level SDH will improve as the adoption of these codes increase. Additionally, other patient-level

SDH related to an individual's family/social support may be inferred from their family medical history.

At the present time, the scarcity of SDH data availability across the U.S. presents a significant barrier for the implementation of our efforts outside the current settings. Despite recent interest in the use of social determinants for health (Alley, Asomugha, Conway, & Sanghavi, 2016; Solar & Irwin, 2010), there has been little groundwork done in the efficient collection of SDH at scale. The availability of aggregate SDH data may vary across regions. Additionally, the fragmented nature of SDH collection and lack of proper standardization of such data would also prevent its effective use in other large-scale systems. A potential alternative may be to rely on SDH elements captured by entities outside of the healthcare domain. For example, the online real estate marketplace, Zillow, presents a public API where certain social determinants covering the entire U.S. may be obtained (programmableweb.com, 2017). Other SDH on crime, environment, and education may be obtained from other commercial and non-commercial entities, albeit at different aggregate-levels. Additionally, there is a strong need for standards that support the collection and sharing of aggregate-level SDH. Despite of our inability to demonstrate value in leveraging SDH for predictive modeling, these efforts, together with previous initiatives (Kasthurirathne, Vest, et al., 2017), contribute to the meager body of existing knowledge on the use of SDH for healthcare decision making. As suggested previously, there is significant need for continued research that provides best practice guidelines on the use of SDH for use in machine learning for healthcare.

We identified a number of limitations that impact the internal and external validity of our efforts. In current clinical practice, healthcare teams that encounter patients with extremely serious cases of depression, which may lead to suicide or harmful behavior, generally take immediate action, directing these patients to the emergency department or elsewhere. This means that there will be no referral data on high-risk situations available to train the decision models—only data on situations that were considered non-life threatening. We do not anticipate this to be a problem, because patients with serious health needs would be easily detectable; thus, there is no need for an automated approach to identify such patients. The lack of data on serious cases of depression, where urgent care is needed, limits our ability to develop conventional risk scores that would be more effective for implementation purposes. As alternate approach for developing risk scores based on this scenario would be to assess the level of risk based on prediction scores and use these as a measure to identify urgency of treatment desired. A similar approach was implemented via our efforts to predict the need for various wraparound health services to patients presenting at Eskenazi Health, Indianapolis, Indiana (Kasthurirathne, Vest, et al., 2017).

Our work is limited by the accuracy and availability of aggregate socioeconomic, community, and public health data for Marion County, Indiana. For our current effort, SDH data was unavailable for patients who do not reside within Marion County. The accuracy of our models are also dependent upon the accuracy and completeness of clinical and behavioral data, as obtained from IHIE. We assume that many of the patients presenting at Eskenazi are return patients, and that IHIE is able to produce reasonable quantities of clinical and behavioral data from their health records. However, in the case of out-of-state patients or first-timers, there is very little or no clinical data at hand to provide appropriate judgment. In the case of out-of-state patients, SDH data would also be unavailable. The proper action for such a scenario is still unclear and depends on the performance of decision models built using only aggregate data. If such models perform adequately well, we may consider using them to predict outcomes for patients with no clinical data. However, this remains to be seen, based on the work to be performed.

Our aim is to identify patients in need of advanced treatment for depression. But how will we isolate these individuals from cases of historical need? Given that Eskenazi can only provide us with clinical data from 2010 forward, we anticipate that a significant majority of the data will indicate current and ongoing need for depression care. However, we are unable to facilitate other scenarios involving past need for depression care and acknowledge it as a limitation in our study.

We acknowledge numerous studies that indicate an association between depression and an individual's genetics (Cohen-Woods, Craig, & McGuffin, 2013; Flint & Kendler, 2014; Hettema, 2010; Schwartz & Petersen, 2016). Unfortunately, IHIE does not currently support or persist with any genomic data. Furthermore, while cases of depression among immediate family members may indicate predisposition, existing IHIE data does not adequately support the ability to identify patient relationships without considerable effort involving Natural Language Processing (NLP)-based efforts. Additionally, the cost of genetic testing and processing has posed a significant barrier to enabling the widespread availability of such data for study. Thus, our study did not leverage any genetic data for analysis.

Finally, the decision whether to treat patients or to refer them to mental health lies with the provider. In some cases, providers may disagree on what constitutes an advanced care need and what they can treat on their own. Thus, our gold standard of referrals used in our analysis may be influenced by differences in providers' personal preferences.

# 8 CONCLUSIONS

We leveraged a host of clinical, behavioral, and patient and population-centric social determinants of health informed by the fundamentals of precision medicine and a conceptual framework of social determinants of health to develop decision models capable of identifying patients in need of advanced care for depression. Models predicting need of advanced care across various high-risk populations and the overall patient population yielded significantly high results suitable for implementation at clinical settings. Additionally, we were able to achieve high performance with decision models that used a relatively small feature set. However, the use of SDH did not yield statistically significant performance improvements in any of the decision models that were tested. Irrespective of these results, our methods can be applied to a multitude of other healthcare questions, where the inclusion of SDH may yield improved results. Our research presents significant potential to automate the identification of patients in need of mental health services using machine learning techniques and relatively low human intervention/expertise. However, there is an urgent need for research and implementation efforts around capturing and sharing patient and population-centric SDH, as well as better approaches to including the element of temporality in modeling SDH for machine-learning purposes.

# 9 APPENDICES

## 9.1 Breakdown of Concepts of the depression terminology across each UMLS mandated semantic type and semantic group.

| | | |
|---|---|---|
| Activities and behaviors | | 389 |
| | Activity | 113 |
| | Behavior | 7 |
| | Daily or recreational activity | 29 |
| | Event | 8 |
| | Governmental or regulatory activity | 19 |
| | Individual behavior | 71 |
| | Machine activity | 8 |
| | Occupational activity | 52 |
| | Social behavior | 82 |
| Anatomy | | 204 |
| | Anatomical structure | 5 |
| | Body location or region | 33 |
| | Body part, organ, organ component | 113 |
| | Body substance | 20 |
| | Body system | 16 |
| | Embryonic structure | 1 |
| | Tissue | 16 |
| Chemicals and drugs | | 604 |
| | Amino acid peptide or protein | 152 |
| | Antibiotic | 14 |
| | Biologically active substance | 14 |
| | Biomedical or dental material | 14 |
| | Chemical | 2 |
| | Chemical viewed functionally | 6 |
| | Chemical viewed structurally | 5 |
| | Chemical drug | 2 |
| | Element, ion or isotope | 32 |
| | Enzyme | 7 |
| | Hazardous or poisonous substance | 24 |
| | Immunologic factor | 20 |
| | Indicator reagent or diagnostic aid | 7 |
| | Inorganic chemical | 14 |
| | Nucleic acid nucleoside or nucleotide | 6 |
| | Organic chemical | 148 |
| | Pharmacological substance | 137 |
| Concepts and ideas | | 2499 |
| | Classification | 37 |
| | Conceptual entity | 112 |
| | Functional concept | 421 |

| | | |
|---|---|---|
| | Group attribute | 15 |
| | Idea or concept | 299 |
| | Intellectual product | 15 |
| | Qualitative concept | 604 |
| | Quantitative concept | 492 |
| | Regulation or law | 15 |
| | Spatial concept | 210 |
| | Temporal concept | 279 |
| Disorders | | 1250 |
| | Acquired abnormality | 10 |
| | Anatomical abnormality | 7 |
| | Cell or molecular dysfunction | 2 |
| | Disease or syndrome | 278 |
| | Experimental model or disease | 2 |
| | Finding | 492 |
| | Injury or poisoning | 78 |
| | Mental or behavioral dysfunction | 144 |
| | Neoplastic process | 48 |
| | Pathologic function | 83 |
| | Sign or symptom | 106 |
| Genes and molecular sequences | | 194 |
| | Gene or genome | 185 |
| | Nucleotide sequences | 9 |
| Geographic areas | | 162 |
| | Geographic area | 162 |
| Living beings | | 415 |
| | Age group | 24 |
| | Family group | 26 |
| | Group | 11 |
| | Human | 11 |
| | Patient or disabled group | 26 |
| | Population group | 143 |
| | Professional or occupational group | 174 |
| Objects | | 213 |
| | Entity | 2 |
| | Food | 23 |
| | Manufactured object | 166 |
| | Substance | 22 |
| Occupations | | 108 |
| | Biomedical occupation or discipline | 84 |
| | Occupation or discipline | 24 |
| Organizations | | 171 |
| | Healthcare-related organization | 132 |
| | Organization | 29 |
| | Professional society | 5 |

| | | |
|---|---|---|
| | Self-help or relief organization | 5 |
| Phenomena | | 146 |
| | Biologic function | 10 |
| | Environmental effect of humans | 3 |
| | Human caused phenomenon or process | 12 |
| | Laboratory or test results | 30 |
| | Natural phenomenon or process | 41 |
| | Phenomenon or process | 50 |
| Procedures | | 1047 |
| | Diagnostic procedure | 106 |
| | Educational activity | 32 |
| | Healthcare activity | 277 |
| | Laboratory procedure | 95 |
| | Molecular biology research technique | 1 |
| | Research activity | 149 |
| | Therapeutic or preventive procedure | 387 |

## 9.2 Distribution of linkages between concepts of the depression terminology to other concepts across the breadth of the UMLS Metathesaurus.

| Terminology identification code | Terminology name | Number of directly linked UMLS concepts |
|---|---|---|
| MTH | Metathesaurus Names | 6174 |
| CHV | CHV | 4843 |
| NCI | NCI Thesaurus | 4487 |
| SNOMEDCT_US | SNOMED Clinical Terms US Edition | 4086 |
| MSH | MeSH | 2676 |
| SNMI | SNOMED Intl 1998 | 2028 |
| LNC | LOINC | 2018 |
| AOD | Alcohol and Other Drug Thesaurus | 1884 |
| CSP | CRISP Thesaurus | 1453 |
| LCH_NW | Library of Congress Subject Headings, Northwestern University subset | 1146 |
| SNM | SNOMED 1982 | 947 |
| LCH | Library of Congress Subject Headings (Division, Policy, Policy, & Office, 2012) | 929 |
| NCI_NCI-GLOSS | NCI Dictionary of Cancer Terms | 832 |
| NCI_CDISC | Clinical Data Interchange Standards Consortium | 816 |

| | | |
|---|---|---|
| HL7V3.0 | HL7 Version 3.0 | 708 |
| NDFRT | National Drug File - Reference Terminology | 642 |
| NCI_FDA | U.S. Food and Drug Administration | 564 |
| OMIM | OMIM | 486 |
| MEDLINEPLUS | MedlinePlus | 433 |
| DXP | DXplain | 368 |
| NCI_NICHD | National Institute of Child Health and Human Development | 363 |
| HL7V2.5 | HL7 Version 2.5 | 348 |
| ICD10CM | ICD-10-CM | 338 |
| COSTAR | COSTAR | 322 |
| CST | COSTART | 298 |
| ICD9CM | ICD-9-CM | 271 |
| HPO | HPO | 264 |
| MTHICD9 | ICD-9-CM Entry Terms | 227 |
| ATC | Anatomical Therapeutic Chemical (ATC) classification system | 206 |
| PDQ | PDQ | 203 |
| RXNORM | RXNORM | 193 |
| FMA | Foundational Model of Anatomy | 178 |
| VANDF | National Drug File | 178 |
| UWDA | Digital Anatomist | 157 |
| NCI_CTCAE | Common Terminology Criteria for Adverse Events | 145 |
| MTHSPL | FDA Structured Product Labels | 136 |
| CCS | Clinical Classifications Software | 116 |
| ICF-CY | International Classification of Functioning, Disability and Health for Children and Youth (Simeonsson et al., 2003) | 107 |
| HGNC | HUGO Gene Nomenclature Committee | 107 |
| ICF | International Classification of Functioning, Disability and Health | 101 |
| USPMG | USP Model Guidelines | 95 |
| NCI_UCUM | Unified Code for Units of Measure (Schadow & McDonald, 2005) | 79 |
| NCI_BRIDG | Biomedical Research Integrated Domain Group Model | 74 |
| GO | Gene Ontology | 66 |
| ICPC | ICPC | 64 |

| | | |
|---|---|---|
| QMR | Quick Medical Reference | 62 |
| NCBI | NCBI Taxonomy | 44 |
| NCI_NCPDP | National Council for Prescription Drug Programs | 43 |
| NCI_CareLex | Content Archive Resource Exchange Lexicon | 35 |
| NCI_DCP | NCI Division of Cancer Prevention Program | 33 |
| AIR | AI/RHEUM | 30 |
| NCI_CDC | U.S. Centers for Disease Control and Prevention | 30 |
| MCM | Glossary of Clinical Epidemiologic Terms | 28 |
| NCI_DTP | (NCI) Developmental Therapeutics Program | 26 |
| NCI_CRCH | Cancer Research Center of Hawaii Nutrition Terminology | 25 |
| NCI_CTEP-SDC | Cancer Therapy Evaluation Program - Simple Disease Classification | 24 |
| NCI_NCI-HL7 | NCI Health Level 7 | 23 |
| MTHMST | Minimal Standard Terminology (UMLS) | 22 |
| AOT | Authorized Osteopathic Thesaurus | 19 |
| NCI_ICH | International Conference on Harmonization | 17 |
| MTHHH | HCPCS Hierarchical Terms (UMLS) | 13 |
| HCPCS | HCPCS | 12 |
| NCI_KEGG | KEGG Pathway Database | 12 |
| SPN | Standard Product Nomenclature | 12 |
| NCI_DICOM | Digital Imaging Communications in Medicine | 11 |
| ICD10PCS | ICD-10-PCS | 11 |
| SOP | Source of Payment Typology | 6 |
| RAM | Clinical Concepts by R A Miller | 5 |
| SRC | Source Terminology Names (UMLS) | 1 |
| CVX | Vaccines Administered | 1 |

### 9.3 List of KFF categories, together with indicators used for measurement.

| KFF category types | Conceptual indicators | Indicators |
|---|---|---|
| Neighborhood and Physical Environment | Housing | Households with telephone<br>Households with plumbing<br>Total housing units<br>Median year structure was built<br>Monthly housing cost<br>Homeless population |

| | | |
|---|---|---|
| | | Income spent on housing |
| | Transportation | Population with private transportation<br>Population dependent on public transportation |
| | Safety | Juvenile offense charges<br>All crimes<br>Feelings of safety<br>Firearm ownership<br>Lighted walkways |
| | Parks | Access to parks |
| Economic Stability | Employment | Employed labor force<br>Unemployed labor force<br>Population employed in construction<br>Population employed in manufacturing<br>Families with two working parents |
| | Income | Median family income<br>Median household income<br>Population with poverty status determined<br>Population living below 125% poverty<br>Population living below 185% poverty |
| | Supportive assistance | Households with cash public assistance |
| Education | Higher education | Population with college or graduate school,<br>population with a bachelor's degree or higher<br>population with a high school diploma only |
| | Language | Adults in workforce speaking English<br>Adults in workforce speaking only languages other than English<br>Seniors speaking English<br>Seniors speaking only languages other than English<br>Households speaking English<br>Households speaking only languages other than English |
| Food | Hunger | Household food affordability |
| | Access to healthy options | Access to full-service grocery store<br>Fast food consumption |
| | Social integration | Access to community center/library |
| | Support systems | Neighbors are willing to help |

| | | |
|---|---|---|
| Community and social context | Community engagement | Registered to vote |
| | Discrimination | Acceptance by healthcare provider<br>Feelings of discrimination |
| Healthcare system | Health coverage | Population without health insurance<br>No visit to doctor due to high cost<br>No Rx med due to high cost |
| | Provider availability | Have single primary care provider? |
| | Quality of care | Individual's general health<br># Days poor physical health<br># Days poor mental health<br>EVER diagnosed with heart event<br>EVER diagnosed with depression<br>EVER diagnosed with high cholesterol<br>EVER diagnosed with asthma<br>EVER diagnosed with diabetes<br>EVER diagnosed with HBP<br>Life expectancy |

## 9.4 Categories of clinical data used in decision-model building

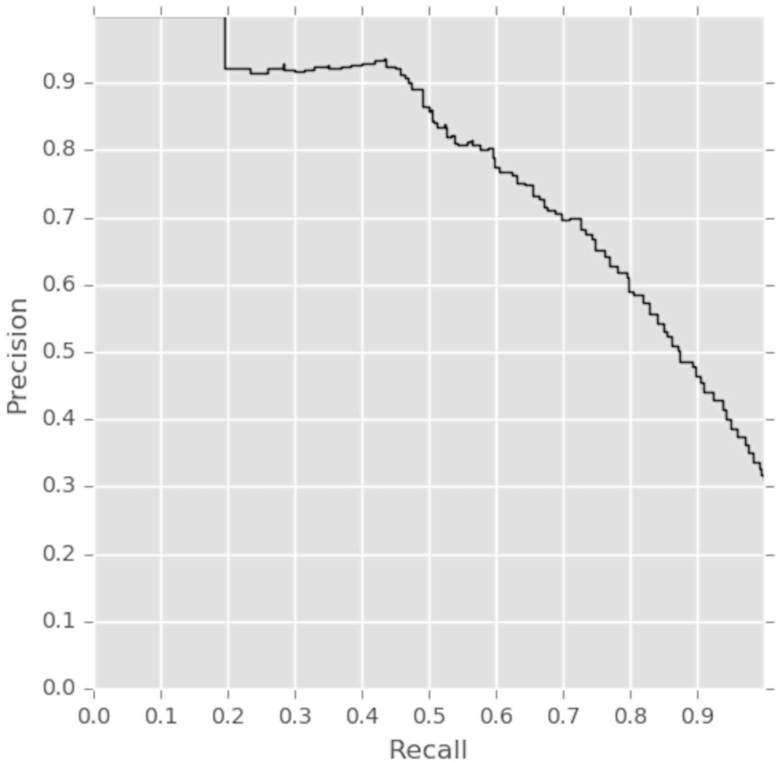
Patient demographics:
Gender, race, age, race/ethnicity

Chronic conditions:
Arthritis, asthma, coronary artery disease, cardiac arrhythmias, hypertension, hyperlipidemia, stroke, autism spectrum disorder, cancer, chronic kidney disease, chronic obstructive pulmonary disease, dementia, depression, diabetes, hepatitis, HIV AIDS, osteoporosis, schizophrenia

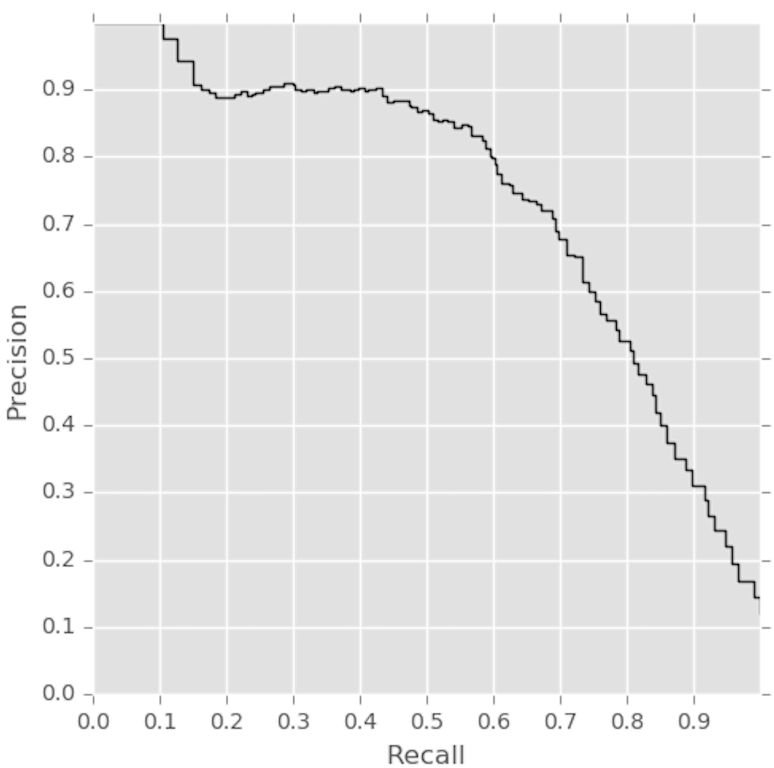
Acute conditions:
A total of 1,116 different acute diagnoses

Previous visit history:
Emergency visits, impatient visits, outpatient visits

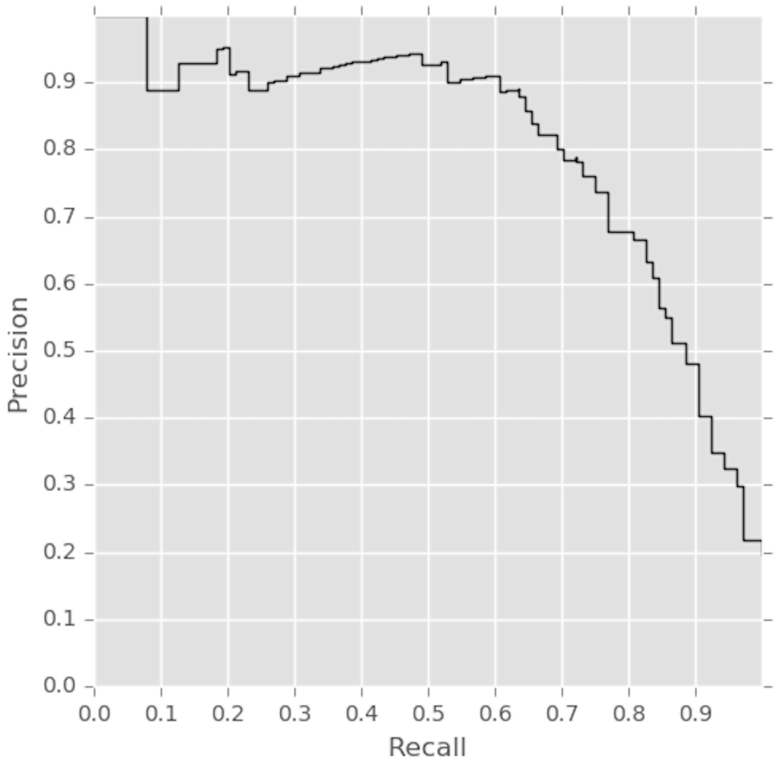
### 9.5 Precision-Recall curve for each decision model under study



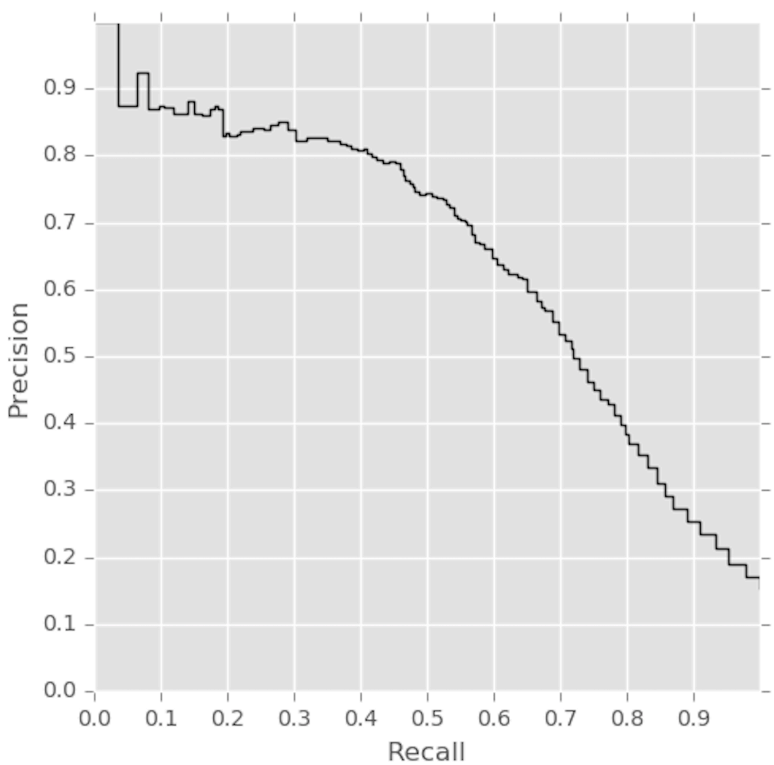

Patients with a prior diagnosis of depression (Group A)




Patients with a Charlson Index of <=1 (Group B)



Patients with a Charlson Index of <=2 (Group C)



All unique patients in groups A-C (Group D)

## 9.6 List of 20 top features (Ranked in order of best to worst) for each decision model, together with their LASSO scores

| # | Master patient population | Group A | Group B | Group C | Group D |
|---|---|---|---|---|---|
| 1 | Essential (primary) hypertension (1.0) | Number of outpatient visits > last 30 days (1.0) | Number of outpatient visits > last 30 days (1.0) | Number of outpatient visits > last 30 days (1.0) | Number of outpatient visits > 30 days (1.0) |
| 2 | Depressive disorder (0.765) | Schizophrenia (1.0) | Depressive disorder (1.0) | Number of inpatient visits prior to last 30 days (1.0) | Depressive disorder (1.0) |
| 3 | Dorsalgia (0.5) | Hyperlipidemia (1.0) | Charlson Index (1.0) | Number of inpatient visits during last 30 days (1.0) | Gender (1.0) |
| 4 | Nicotine dependence (0.5) | External injury (1.0) | Gender (1.0) | Hyperlipidemia (1.0) | Essential (primary) hypertension (1.0) |
| 5 | Joint disorder (0.48) | Depressive disorder (1.0) | Essential (primary) hypertension (1.0) | External injury (1.0) | Diabetes mellitus (1.0) |
| 6 | Gender (0.48) | Arthritis (1.0) | Episodic mood disorders (1.0) | Depressive disorder (1.0) | Charlson Index (1.0) |
| 7 | Encounter for contraceptive management (0.475) | Nonspecific findings on examination of blood (1.0) | Diabetes mellitus (1.0) | Charlson Index (1.0) | Asthma (1.0) |
| 8 | Routine general medical examination (0.475) | Other cerebral degenerations (1.0) | Hyperlipidemia (0.87) | Arthritis (1.0) | Disorders of lipoid metabolism (0.95) |
| 9 | Examination of eyes and vision | Cancer (0.995) | Cancer (0.77) | Age (1.0) | Other cerebral degenerations (0.9) |

| | | | | | |
|---|---|---|---|---|---|
| | (0.465) | | | | |
| 10 | Abdominal and pelvic pain (0.465) | Vitamin D deficiency (0.995) | Other cerebral degenerations (0.715) | Gender (1.0) | Symptoms of the respiratory system and other chest symptoms (0.85) |
| 11 | Encounter for screening for malignant neoplasms (0.46) | Charlson Index (0.985) | Asthma (0.6) | Chronic kidney disease (1.0) | Bronchitis (0.8) |
| 12 | Soft tissue disorders (0.46) | Episodic mood disorders (0.805) | Number of inpatient visits prior to last 30 days (0.555) | Chronic airway obstruction (1.0) | Number of emergency department visits prior to last 30 days (0.77) |
| 13 | Long-term (current) drug therapy (0.455) | Cataract (0.8) | Bronchitis (0.555) | Other and ill-defined cerebrovascular disease (1.0) | Number of inpatient visits prior to last 30 days (0.723) |
| 14 | Number of outpatient visits prior to last 30 days (0.43) | Asthma (0.59) | External injury (0.55) | Heart failure (1.0) | Hyperlipidemia (0.66) |
| 15 | Episodic mood disorders (0.22) | Heart failure (0.545) | Arthritis (0.545) | Acute myocardial infarction (0.995) | Disorders of fluid electrolyte and acid-base balance (0.635) |
| 16 | Pain in throat and chest (0.175) | Number of emergency department visits during last 30 days (0.54) | Disorders of fluid electrolyte and acid-base balance (0.515) | Essential (primary) hypertension (0.95) | Persistent mental disorders (0.6) |
| 17 | Encounter for screening for infectious | Chronic kidney disease (0.525) | Nondependent abuse of drugs (0.51) | Asthma (0.925) | Chronic obstructive pulmonary disease (0.612) |

| | | | | | |
|---|---|---|---|---|---|
| | and parasitic diseases (0.175) | | | | |
| 18 | Encounter for other special examination without complaint (0.15) | Sprain of neck (0.52) | Number of inpatient visits during last 30 days (0.49) | Nondependent abuse of drugs (0.9) | Arthritis (0.6) |
| 19 | Anxiety, dissociative and somatoform disorders (0.125) | Gender (0.515) | Number of emergency department visits prior to last 30 days (0.49) | Diabetes mellitus (0.85) | Acute myocardial infarction (0.55) |
| 20 | Diabetes mellitus (0.05) | Attention deficit disorder without mention of hyperactivity (0.515) | Persistent mental disorders (0.48) | Other peripheral vascular disease (0.85) | Dementia (0.51) |

## 9.7 Co-occurrence of top 20 features across each of the patient populations under test (1 = most important, 20 == least important)

| Feature | Master patient population | Group A | Group B | Group C | Group D |
|---|---|---|---|---|---|
| Essential (primary) hypertension | 1 | | 5 | 16 | 4 |
| Depressive disorder | 2 | 5 | 2 | 6 | 2 |
| Dorsalgia | 3 | | | | |
| Nicotine dependence | 4 | | | | |
| Joint disorder | 5 | | | | |
| Gender | 6 | 19 | 4 | 10 | 3 |
| Encounter for contraceptive management | 7 | | | | |
| Routine general medical examination | 8 | | | | |
| Examination of eyes and vision | 9 | | | | |

| | | | | | |
|---|---|---|---|---|---|
| Abdominal and pelvic pain | 10 | | | | |
| Encounter for screening for malignant neoplasms | 11 | | | | |
| Soft tissue disorders | 12 | | | | |
| Long-term (current) drug therapy | 13 | | | | |
| Number of outpatient visits prior to last 30 days | 14 | 1 | 1 | 1<br>53.18 | 1 |
| Episodic mood disorders | 15 | 12 | 6 | | |
| Pain in throat and chest | 16 | | | | |
| Encounter for screening for infectious and parasitic diseases | 17 | | | | |
| Encounter for other special examination without complaint | 18 | | | | |
| Anxiety, dissociative, and somatoform disorders | 19 | | | | |
| Diabetes mellitus | 20 | | 7 | 19 | 5 |
| Schizophrenia | | 2 | | | |
| Hyperlipidemia | | 3 | 8 | 4 | 14 |
| External injury | | 4 | 14 | 5 | |
| Arthritis | | 6 | 15 | 8 | 18 |
| Nonspecific findings on examination of blood | | 7 | | | |
| Other cerebral degenerations | | 8 | 10 | | 9 |
| Cancer | | 9 | 9 | | |
| Vitamin D deficiency | | 10 | | | |
| Charlson Index | | 11 | 3 | 7 | 6 |
| Cataract | | 13 | | | |
| Asthma | | 14 | 11 | 17 | 7 |
| Heart failure | | 15 | | 14 | |
| Number of emergency department visits during last 30 days | | 16 | | | |
| Chronic kidney disease | | 17 | | 11 | |
| Sprain of neck | | 18 | | | |
| Attention deficit disorder without mention of hyperactivity | | 20 | | | |

| | | | | | |
|---|---|---|---|---|---|
| Number of inpatient visits prior to last 30 days | | | 12 | 2 | 13 |
| Bronchitis | | | 13 | | 11 |
| Disorders of fluid electrolyte and acid-base balance | | | 16 | | 15 |
| Nondependent abuse of drugs | | | 17 | 18 | |
| Number of inpatient visits during last 30 days | | | 18 | 3 | |
| Number of emergency department visits prior last 30 days | | | 19 | | 12 |
| Persistent mental disorders | | | 20 | | 16 |
| Age | | | | 9 | |
| Chronic airway obstruction | | | | 12 | |
| Other and ill-defined cerebrovascular disease | | | | 13 | |
| Acute myocardial infarction | | | | 15 | 19 |
| Other peripheral vascular disease | | | | 20 | |
| Disorders of lipoid metabolism | | | | | 8 |
| Symptoms involving respiratory system and other chest symptoms | | | | | 10 |
| Chronic obstructive pulmonary disease | | | | | 17 |
| Dementia | | | | | 20 |

## 9.8 List of patient-centric SDH and number of patients reporting, as extracted from the master patient population

| Id | Description | ICD-9 and ICD-10 codes | # of patients reporting |
|---|---|---|---|
| 1 | Problems related to education and literacy | Z55.*<br>V62.3 | 127 |
| 2 | Problems related to employment and unemployment | Z56.*<br>V62.0, V62.1, V62.2 | 113 |

| | | | |
|---|---|---|---|
| 3 | Occupational exposure to risk factors | Z57.*<br>V15.85 | 29 |
| 4 | Exposure to noise, air, water, soil, pollution, radiation, etc. | Z58.* | 0 |
| 5 | Problems related to housing and economic circumstances | Z59.*<br>V60.0, V60.1, V60.2, V60.6, V60.89,V60.9 | 631 |
| 6 | Problems related to social environment | Z60.*<br>V62.4, V60.3 | 30 |
| 7 | Problems related to negative life events in childhood | Z61.* | 0 |
| 8 | Problems related to upbringing | Z62.*<br>V60.81, V61.29<br>V61.8, | 51 |
| 9 | Other problems related to primary support group, including family circumstances | Z63.*<br>V61.3, V61.41<br>V61.42, V61.49<br>V61.8, V61.9<br>V62.82, V61.01, V61.02, V61.03, V61.04, V61.05, V61.06, V61.07, V61.08, V61.09 | 446 |
| 10 | Problems related to certain psychosocial circumstances | Z64.*<br>V61.5, V61.6, V61.7 | 42 |
| 11 | Problems related to other psychosocial circumstances | Z65.*<br>V62.21, V62.22, V62.29,<br>V62.5, V62.81, V62.89, V62.9<br>V11.0, V11.1, V11.2, V11.3, V11.9 | 410 |
| 12 | Problems related to medical facilities and other health care | Z75.*<br>V60.5,<br>V63.0,V63.1,V63.2,V63.8,V63.9 | 4 |

## 9.9 List of Aggregate-level SDH extracted from the Polis Center

| Column name | Description |
|---|---|
| COLLEGEN1 | Percent of Population Age 3 and Over in College or Graduate School |
| BACHMOREN1 | Percent of Population Age 25 and Over with a Bachelor’s Degree or Higher |
| POPWDIPN1 | Percent Population (Age 25 and Over) With High School Diploma Only (Includes Equivalency) |
| CMB30PMN2 | Percent of Occupied Housing units (combination of rental and owner) Whose Occupants Pay 30% or More of Income for Housing Costs |
| TOTOWNOCN1 | Percent of All Occupied Units that are Owner Occupied |

| | |
|---|---|
| HHLDFSN1 | Percent of Households with Cash Public Assistance or Food Stamps/SNAP |
| MEDFMLYINC | Median Family Income |
| MEDHHLDINC | Median Household Income |
| TOTPOPNON1 | Percent of Population All Ages Without Health Insurance |
| TOTUNEMPN2 | Percent of Labor Force Age 16 and Over Who Are Unemployed |
| LANGENGLN1 | Percent of All Households with Household Language of English |
| PERINPOVN1 | Percent of Population in Poverty For Whom Poverty Status is Determined |
| POVB125N1 | Percent of Population Living below 125% Poverty |
| POVB185N1 | Percent of Population Living below 185% Poverty |
| NONCARN1 | Percent of Workers Age 16 and Over Who did not Drive a Car, Truck or Van as their Means of Transportation to Work |
| Life Expectancy | Life Expectancy |
| nbh_Safe | I feel safe in my neighborhood (Mean 1 Strong agree - 5 Strong Disagree) |
| nbh_Help | Neighbors are willing to help (Mean 1 Strong agree - 5 Strong Disagree) |
| nbh_Blight | Neighborhood has blight (Mean 1 Strong agree - 5 Strong Disagree) |
| nbh_Pave | Neighborhood has paved walkways (% reporting yes) |
| nbh_Wheelch | Neighborhood has walkways with wheeled access (% reporting yes) |
| nbh_Light | Neighborhood has lighted walkways (% reporting yes) |
| nbh_Conn | Neighborhood walkways connect to other major streets/neighborhoods (% reporting yes) |
| nbh_Groc | Full service grocery store within 10-min walk from home (% reporting yes) |
| nbh_Commctr | Community center/library within 10-min walk from home (% reporting yes) |
| nbh_Park | Park/greenway/playground within 10-min walk from home (% reporting yes) |
| nbh_PubTrans | Bus/public transportation within 10-min walk from home (% reporting yes) |
| smokcurr | Current smoker (% reporting yes) |
| pacalc | Engages in moderate physical activity for 150 min or more per week (% reporting yes) |
| cd_drinkbinge | # occasions in past 30 days: 4/more drinks (F) or 5/more drinks (M) (Mean of 0 to 30 days) |
| cd_violc | Member of household was a victim of violence in past 12 months (% reporting yes) |
| cd_gun | Firearm in home (% reporting yes) |
| cd_addict | Member of household ever addicted to meds (% reporting yes) |

| | |
|---|---|
| hth_gen | Describe own general health (Mean of 1 Excellent 2 Good 3 Fair 4 Poor 5 Very poor) |
| hth_daysphys | # Days poor physical health (Mean of 0 to 30 days) |
| hth_daysment | # Days poor mental health (Mean of 0 to 30 days) |
| cd_heart | EVER diagnosed with Heart Event (% reporting yes) |
| cd_depr | EVER diagnosed with Depression (% reporting yes) |
| cd_chol | EVER diagnosed with Hi Chol (% reporting yes) |
| cd_asth | EVER diagnosed with Asthma (% reporting yes) |
| cd_diab | EVER diagnosed with Diab (% reporting yes) |
| cd_hbp | EVER diagnosed with HBP (% reporting yes) |
| fd_fastfood | Number of times ate fast food last 7 days (Mean) |
| fd_foodsecure | Describe household food affordability last 12 months (Mean of 1 “Could always afford enough food to eat,” 2 “Sometimes couldn’t afford enough food to eat,” 3 “Often couldn’t afford enough food to eat” |
| hth_hrstv | Hours TV per day (Mean) |
| hc_1provdr | Have single primary care provider? (% reporting yes) |
| hc_costdr | No visit to Dr. in last 12 months due to hi $ (% reporting yes) |
| hc_costmed | No Rx med in last 12 months due to hi $ (% reporting yes) |
| hc_respect | Feel accepted/respected by current healthcare provider (% reporting yes) |

### 9.10 Distribution of patient populations across demographic areas

| | Group A | Group B | Group C | Group D |
|---|---|---|---|---|
| Patient set size | 11,349 | 29,840 | 7,415 | 84,317 |
| Distribution across HPA | 19/19 | 19/19 | 19/19 | 19/19 |
| Distribution across Zip codes | 37/38 | 37/38 | 37/38 | 38/38 |
| Distribution across census tracts | 224/224 | 224/224 | 221/224 | 224/224 |

# 10 REFERENCES


Agency for Healthcare Research and Quality. (2010). 2010 National Healthcare Disparities Report. Retrieved from http://archive.ahrq.gov/research/findings/nhqrdr/nhdr10/index.html

Alexopoulos, G. S., Abrams, R. C., Young, R. C., & Shamoian, C. A. (1988). Cornell scale for depression in dementia. *Biological Psychiatry, 23*(3), 271-284.

Allen, J., Balfour, R., Bell, R., & Marmot, M. (2014). Social determinants of mental health. *International Review of Psychiatry, 26*(4), 392-407.

Alley, D. E., Asomugha, C. N., Conway, P. H., & Sanghavi, D. M. (2016). Accountable health communities—addressing social needs through Medicare and Medicaid. *New England Journal of Medicine, 374*(1), 8-11.

Almalki, M., Gray, K., & Sanchez, F. M. (2015). The use of self-quantification systems for personal health information: big data management activities and prospects. *Health Information Science and Systems, 3*(S1), S1.

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders (DSM-5®)*: American Psychiatric Pub.

American Psychiatric Association. (2016). What is Depression? Retrieved from http://www.psychiatry.org/patients-families/depression/what-is-depression

Anda, R. F., Whitfield, C. L., Felitti, V. J., Chapman, D., Edwards, V. J., Dube, S. R., & Williamson, D. F. (2002). Adverse childhood experiences, alcoholic parents, and later risk of alcoholism and depression. *Psychiatric Services*.

Anderson, R. J., Freedland, K. E., Clouse, R. E., & Lustman, P. J. (2001). The prevalence of comorbid depression in adults with diabetes a meta-analysis. *Diabetes Care, 24*(6), 1069-1078.

Aneshensel, C. S., & Sucoff, C. A. (1996). The neighborhood context of adolescent mental health. *Journal of Health and Social Behavior*, 293-310.

Annis, R., Racher, F., & Beattie, M. (2004). *Rural community health and well-being: A guide to action*: Rural Development Institute.

Aronson, A. R., & Lang, F.-M. (2010). An overview of MetaMap: historical perspective and recent advances. *Journal of the American Medical Informatics Association, 17*(3), 229-236.

Arroll, B., Goodyear-Smith, F., Crengle, S., Gunn, J., Kerse, N., Fishman, T., . . . Hatcher, S. (2010). Validation of PHQ-2 and PHQ-9 to screen for major depression in the primary care population. *The Annals of Family Medicine, 8*(4), 348-353.

Aubin, H., Tilikete, S., & Barrucand, D. (1995). [Depression and smoking]. *L'Encephale, 22*(1), 17-22.

Bagby, R. M., Ryder, A. G., Schuller, D. R., & Marshall, M. B. (2004). The Hamilton Depression Rating Scale: has the gold standard become a lead weight? *American Journal of Psychiatry, 161*(12), 2163-2177.

Bates, D. W., Saria, S., Ohno-Machado, L., Shah, A., & Escobar, G. (2014). Big data in health care: using analytics to identify and manage highrisk and high-cost patients. *Health Affairs, 33*(7), 1123-1131.

Beck, A. T., Steer, R. A., & Brown, G. K. (1996). BDI-II. Beck Depression Inventory Second Edition. Manual. *San Antonio: Psychological Corporation*.

Bengio, Y., Courville, A., & Vincent, P. (2013). Representation learning: A review and new perspectives. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 35*(8), 1798-1828.

Berman, J. J. (2004). Tumor taxonomy for the developmental lineage classification of neoplasms. *BMC Cancer, 4*(1), 88.

Berwick, D. M., Nolan, T. W., & Whittington, J. (2008). The triple aim: care, health, and cost. *Health Affairs, 27*(3), 759-769.

Blazer, D. G., Kessler, R. C., & McGonagle, K. A. (1994). The prevalence and distribution of major depression in a national community sample: the National Comorbidity Survey. *Age (years), 15*(24), 24-27.

Blumenthal, J. A., Lett, H. S., Babyak, M. A., White, W., Smith, P. K., Mark, D. B. (2003). Depression as a risk factor for mortality after coronary artery bypass surgery. *The Lancet, 362*(9384), 604-609.

Bodenhamer, D. J., Colbert, J. T., Comer, K. F., & Kandris, S. M. (2011). Developing and sustaining a community information system for central Indiana: SAVI as a case study *Community quality-of-life indicators: Best cases V* (pp. 21-46): Springer.

Bodenreider, O. (2004). The unified medical language system (UMLS): integrating biomedical terminology. *Nucleic Acids Research, 32*(suppl 1), D267-D270.

Brank, J., Grobelnik, M., & Mladenić, D. (2005). A survey of ontology evaluation techniques.

Braveman, P., Egerter, S., & Williams, D. R. (2011). The social determinants of health: coming of age. *Annual Review of Public Health, 32*, 381-398.

Breiman, L. (2001). Random forests. *Machine Learning, 45*(1), 5-32.

Bressert, S. (2017). Persistent Depressive Disorder (Dysthymia) Symptoms. Retrieved from https://psychcentral.com/disorders/dysthymic-disorder-symptoms/

Brewster, C., Alani, H., Dasmahapatra, S., & Wilks, Y. (2004). Data driven ontology evaluation.

Brighton, H., & Mellish, C. (2002). Advances in instance selection for instance-based learning algorithms. *Data Mining and Knowledge Discovery, 6*(2), 153-172.

Brink, T. L., Yesavage, J., & Lum, O. (1983). *Geriatric Depression Scale*.

Brownson, R. C., Colditz, G. A., & Proctor, E. K. (2017). *Dissemination and implementation research in health: translating science to practice*: Oxford University Press.

Butler, R. R. (2007). Icd-10 general equivalence mappings: Bridging the translation gap from icd-9. *Journal of AHIMA, 78*(9), 84-86.

Cape, J., Whittington, C., Buszewicz, M., Wallace, P., & Underwood, L. (2010). Brief psychological therapies for anxiety and depression in primary care: meta-analysis and meta-regression. *BMC Medicine, 8*(1), 38.

Carneiro, G., Nascimento, J., & Bradley, A. P. (2015). Unregistered multiview mammogram analysis with pre-trained deep learning models. Paper presented at the International Conference on Medical Image Computing and Computer-Assisted Intervention.

Casalino, L., Gillies, R. R., Shortell, S. M., Schmittdiel, J. A., Bodenheimer, T., Robinson, J. C., . . . Wang, M. C. (2003). External incentives, information technology, and organized processes to improve health care quality for patients

with chronic diseases. *Journal of the American Medical Association, 289*(4), 434-441.
Charlson, M. E., Pompei, P., Ales, K. L., & MacKenzie, C. R. (1987). A new method of classifying prognostic comorbidity in longitudinal studies: development and validation. *Journal of Chronic Diseases, 40*(5), 373-383.
Chetty, R., Hendren, N., Kline, P., & Saez, E. (2014). Where is the land of opportunity? The geography of intergenerational mobility in the United States. *The Quarterly Journal of Economics, 129*(4), 1553-1623.
Cohen-Woods, S., Craig, I., & McGuffin, P. (2013). The current state of play on the molecular genetics of depression. *Psychological Medicine, 43*(4), 673-687.
Coltman, T., Devinney, T. M., Midgley, D. F., & Venaik, S. (2008). Formative versus reflective measurement models: Two applications of formative measurement. *Journal of Business Research, 61*(12), 1250-1262.
Comer, K. F., Grannis, S., Dixon, B. E., Bodenhamer, D. J., & Wiehe, S. E. (2011). Incorporating geospatial capacity within clinical data systems to address social determinants of health. *Public Health Reports, 126*(3_suppl), 54-61.
Conner, K. O., Copeland, V. C., Grote, N. K., Koeske, G., Rosen, D., Reynolds, C. F., & Brown, C. (2010). Mental health treatment seeking among older adults with depression: the impact of stigma and race. *The American Journal of Geriatric Psychiatry, 18*(6), 531-543.
Corchado, J. M., De Paz, J. F., Rodríguez, S., & Bajo, J. (2009). Model of experts for decision support in the diagnosis of leukemia patients. *Artificial Intelligence in Medicine, 46*(3), 179-200.
Corrigan, P. (2004). How stigma interferes with mental health care. *American Psychologist, 59*(7), 614.
Cuijpers, P., van Straten, A., Andersson, G., & van Oppen, P. (2008a). Psychotherapy for depression in adults: a meta-analysis of comparative outcome studies: American Psychological Association.
Cuijpers, P., van Straten, A., Andersson, G., & van Oppen, P. (2008b). Psychotherapy for depression in adults: a meta-analysis of comparative outcome studies. *Journal of Consulting and Clinical Psychology, 76*(6), 909.
Cullen, M. R., Cummins, C., & Fuchs, V. R. (2012). Geographic and racial variation in premature mortality in the US: analyzing the disparities. *PLoS One, 7*(4), e32930.
Danaher, A. (2011). *Reducing disparities and improving population health: The role of a vibrant community sector*: Wellesley Institute.
Di Monte, D. A. (2003). The environment and Parkinson’s disease: is the nigrostriatal system preferentially targeted by neurotoxins? *The Lancet Neurology, 2*(9), 531-538.
Division, L. o. C. S. C., Policy, L. o. C. O. f. S. C., Policy, L. o. C. C., & Office, S. (2012). *Library of Congress Subject Headings*.
Dowrick, C., & Frances, A. (2013). Medicalising unhappiness: new classification of depression risks more patients being put on drug treatment from which they will not benefit. *BMJ, 347*(7), f7140.
Duke, J. D., Morea, J., Mamlin, B., Martin, D. K., Simonaitis, L., Takesue, B. Y., Dexter, P. R. (2014). Regenstrief Institute's Medical Gopher: A next-generation

homegrown electronic medical record system. *International Journal of Medical Informatics, 83*(3), 170-179.
eClinicalWorks. (2016). Population Health (CCMR). Retrieved from https://www.eclinicalworks.com/products-services/population-health-ccmr/
Eisner, M. D., Katz, P. P., Lactao, G., & Iribarren, C. (2005). Impact of depressive symptoms on adult asthma outcomes. *Annals of Allergy, Asthma & Immunology, 94*(5), 566-574.
Eskenazi Health. (2016). About Us. Retrieved from http://www.eskenazihealth.edu/about-us
Eskenazi Health. (2016). History. Retrieved from http://www.eskenazihealth.edu/about/history
Evenson, K. R., Goto, M. M., & Furberg, R. D. (2015). Systematic review of the validity and reliability of consumer-wearable activity trackers. *International Journal of Behavioral Nutrition and Physical Activity, 12*(1), 159.
Fakoor, R., Ladhak, F., Nazi, A., & Huber, M. (2013). Using deep learning to enhance cancer diagnosis and classification. Paper presented at the Proceedings of the International Conference on Machine Learning.
Ferrari, A. J., Charlson, F. J., Norman, R. E., Patten, S. B., Freedman, G., Murray, C. J., . . . Whiteford, H. A. (2013). Burden of depressive disorders by country, sex, age, and year: findings from the global burden of disease study 2010. *PLoS Med, 10*(11), e1001547.
Field, T. (2010). Postpartum depression effects on early interactions, parenting, and safety practices: a review. *Infant Behavior and Development, 33*(1), 1-6.
Flint, J., & Kendler, K. S. (2014). The genetics of major depression. *Neuron, 81*(3), 484-503.
Fonti, V., & Belitser, E. (2017). Feature Selection using LASSO.
Force, U. P. S. T. (2009). Screening for depression in adults: US preventive services task force recommendation statement. *Annals of Internal Medicine, 151*(11), 784.
Fryers, T., Melzer, D., Jenkins, R., & Brugha, T. (2005). The distribution of the common mental disorders: social inequalities in Europe. *Clinical Practice and Epidemiology in Mental Health, 1*(1), 14.
Garbi, M. d. O. S. S., Hortense, P., Gomez, R. R. F., Silva, T. d. C. R. d., Castanho, A. C. F., & Sousa, F. A. E. F. (2014). Pain intensity, disability and depression in individuals with chronic back pain. *Revista Latino-Americana de Enfermagem, 22*(4), 569-575.
General, O. o. t. S., & Services, C. f. M. H. (2001). Mental health: Culture, race, and ethnicity: A supplement to mental health: A report of the surgeon general.
Giese-Davis, J., Collie, K., Rancourt, K. M., Neri, E., Kraemer, H. C., & Spiegel, D. (2011). Decrease in depression symptoms is associated with longer survival in patients with metastatic breast cancer: a secondary analysis. *Journal of Clinical Oncology, 29*(4), 413-420.
Gilbody, S., Richards, D., Brealey, S., & Hewitt, C. (2007). Screening for depression in medical settings with the Patient Health Questionnaire (PHQ): a diagnostic meta-analysis. *Journal of General Internal Medicine, 22*(11), 1596-1602.
Gilbody, S., Sheldon, T., & Wessely, S. (2006). Should we screen for depression? *BMJ: British Medical Journal, 332*(7548), 1027.

Glassman, A. H., Helzer, J. E., Covey, L. S., Cottler, L. B., Stetner, F., Tipp, J. E., & Johnson, J. (1990). Smoking, smoking cessation, and major depression. *Journal of the American Medical Association, 264*(12), 1546-1549.

Gligorijević, V., Malod-Dognin, N., & Pržulj, N. (2016). Integrative methods for analyzing big data in precision medicine. *Proteomics, 16*(5), 741-758.

Goldberg, D. (2011). The heterogeneity of "major depression." *World Psychiatry, 10*(3), 226-228.

Goodell, S., Druss, B. G., Walker, E. R., & Mat, M. (2011). Mental disorders and medical comorbidity. *Robert Wood Johnson Foundation*, 2.

Greenberg, P. E., Fournier, A.-A., Sisitsky, T., Pike, C. T., & Kessler, R. C. (2015). The economic burden of adults with major depressive disorder in the United States (2005 and 2010). *The Journal of Clinical Psychiatry, 76*(2), 155-162.

Guyon, I., & Elisseeff, A. (2003). An introduction to variable and feature selection. *Journal of Machine Learning Research, 3*(Mar), 1157-1182.

Hamilton, M. (1960). A rating scale for depression. *Journal of Neurology, Neurosurgery, and Psychiatry, 23*(1), 56.

Hartigan, J. A., & Wong, M. A. (1979). Algorithm AS 136: A k-means clustering algorithm. *Journal of the Royal Statistical Society. Series C (Applied Statistics), 28*(1), 100-108.

Hauskrecht, M., & Fraser, H. (2000). Planning treatment of ischemic heart disease with partially observable Markov decision processes. *Artificial Intelligence in Medicine, 18*(3), 221-244.

Hawkins, D. M. (2004). The problem of overfitting. *Journal of Chemical Information and Computer Sciences, 44*(1), 1-12.

Hawton, K., i Comabella, C. C., Haw, C., & Saunders, K. (2013). Risk factors for suicide in individuals with depression: a systematic review. *Journal of Affective Disorders, 147*(1), 17-28.

Hawton, K., Rodham, K., Evans, E., & Weatherall, R. (2002). Deliberate self harm in adolescents: self report survey in schools in England. *BMJ, 325*(7374), 1207-1211.

Haykin, S., & Network, N. (2004). A comprehensive foundation. *Neural Networks, 2*(2004), 41.

Heiman, H. J., & Artiga, S. (2015). Beyond health care: the role of social determinants in promoting health and health equity. *Health, 20*(10), 1-10.

Heslin, K., Elixhauser, A., & Steiner, C. (2012). Hospitalizations involving mental and substance use disorders among adults, 2012. *HCUP Statistical Brief, 191*.

Hettema, J. M. (2010). Genetics of depression. *Focus, 8*(3), 316-322.

Hinz, A., Krauss, O., Hauss, J., Höckel, M., Kortmann, R., Stolzenburg, J., & Schwarz, R. (2010). Anxiety and depression in cancer patients compared with the general population. *European Journal of Cancer Care, 19*(4), 522-529.

Huang, S. H., LePendu, P., Iyer, S. V., Tai-Seale, M., Carrell, D., & Shah, N. H. (2014). Toward personalizing treatment for depression: predicting diagnosis and severity. *Journal of the American Medical Informatics Association, 21*(6), 1069-1075.

Hyman, S. (2014). Depression needs large human-genetics studies. *Nature, 515*(7526), 189.

Institute of Medicine. Committee on the Recommended Social Behavioral Domains Measures for Electronic Health Records. (2014). *Capturing Social and Behavioral Domains and Measures in Electronic Health Records: Phase 2*: National Academies Press.

Jackson, G. L., Powers, B. J., Chatterjee, R., Bettger, J. P., Kemper, A. R., Hasselblad, V., . . . Kendrick, A. S. (2013). The patient-centered medical home: a systematic review. *Annals of Internal Medicine, 158*(3), 169-178.

Jain, A., & Zongker, D. (1997). Feature selection: Evaluation, application, and small sample performance. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 19*(2), 153-158.

Jolliffe, I. (2002). *Principal component analysis*: Wiley Online Library.

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 Associatoin, 306*(15), 1688-1698.

Kasthurirathne, S., Biondich, P., Mamlin, B., Cullen, T., & Grannis, S. (2017). *An Incremental Adoption Pathway for Developing Precision Medicine Based Healthcare Infrastructure for Underserved Settings*.

Kasthurirathne, S. N., Dixon, B. E., Gichoya, J., Xu, H., Xia, Y., Mamlin, B., & Grannis, S. J. (2016). Toward better public health reporting using existing off the shelf approaches: A comparison of alternative cancer detection approaches using plaintext medical data and non-dictionary based feature selection. *Journal of Biomedical Informatics, 60*, 145-152.

Kasthurirathne, S. N., Dixon, B. E., Gichoya, J., Xu, H., Xia, Y., Mamlin, B., & Grannis, S. J. (2017a). Toward better public health reporting using existing off the shelf approaches: The value of medical dictionaries in automated cancer detection using plaintext medical data. *Journal of Biomedical Informatics, 69*, 160-176.

Kasthurirathne, S. N., Dixon, B. E., Gichoya, J., Xu, H., Xia, Y., Mamlin, B., & Grannis, S. J. (2017b). Toward better public health reporting using existing off the shelf approaches: The value of medical dictionaries in automated cancer detection using plaintext medical data. *Journal of Biomedical Informatics*.

Kasthurirathne, S. N., Vest, J. R., Menachemi, N., Halverson, P. K., & Grannis, S. J. (2017). Assessing the capacity of social determinants of health data to augment predictive models identifying patients in need of wraparound social services. *Journal of the American Medical Informatics Association, 25*(1), 47-53.

Katon, W., Von Korff, M., Ciechanowski, P., Russo, J., Lin, E., Simon, G., Young, B. (2004). Behavioral and clinical factors associated with depression among individuals with diabetes. *Diabetes Care, 27*(4), 914-920.

Kerr, L. K., & Kerr Jr, L. D. (2001). Screening tools for depression in primary care. *Western Journal of Medicine, 175*(5), 349.

Kesler, R., Berglund, P., Dernler, O., Jin, R., Koretz, D., Merikangas, K., Wang, P. (2003). National Comorbidity Survey Replication: The epidemiology of major depressive disorder: results from the National Comorbidity Survey Replication (NCS-R). *Journal of the American Medical Association, 289*, 3095-3105.

Kessler, R. C. (2003). Epidemiology of women and depression. *Journal of Affective Disorders, 74*(1), 5-13.

Kessler, R. C., Chiu, W. T., Demler, O., & Walters, E. E. (2005). Prevalence, severity, and comorbidity of 12-month DSM-IV disorders in the National Comorbidity Survey Replication. *Archives of General Psychiatry, 62*(6), 617-627.

Khalilia, M., Chakraborty, S., & Popescu, M. (2011). Predicting disease risks from highly imbalanced data using random forest. *BMC Medical Informatics and Decision Making, 11*(1), 51.

Kodner, D. L., & Spreeuwenberg, C. (2002). Integrated care: meaning, logic, applications, and implications–a discussion paper. *International Journal of Integrated Care, 2*(4).

Kohonen, T. (1982). Self-organized formation of topologically correct feature maps. *Biological Cybernetics, 43*(1), 59-69.

Kung, S., Alarcon, R. D., Williams, M. D., Poppe, K. A., Moore, M. J., & Frye, M. A. (2013). Comparing the Beck Depression Inventory-II (BDI-II) and Patient Health Questionnaire (PHQ-9) depression measures in an integrated mood disorders practice. *Journal of Affective Disorders, 145*(3), 341-343.

LaBrec, P. (2016). ICD-10 adds more detail on the social determinants of health. Retrieved from https://www.3mhisinsideangle.com/blog-post/icd-10-adds-more-detail-on-the-social-determinants-of-health/

Lehmann, L. P., Hubbard, J. R., & Martin, P. R. (2001). Substance abuse and depression. *Substance Aabuse in the Mentally and Physically Disabled*, 33-58.

Lépine, J.-P., & Briley, M. (2011). The increasing burden of depression. *Neuropsychiatric Disease and Treatment, 7*(Suppl 1), 3.

Liang, Z., Zhang, G., Huang, J. X., & Hu, Q. V. (2014). Deep learning for healthcare decision making with EMRs. Paper presented at the Bioinformatics and Biomedicine (BIBM), 2014 IEEE International Conference on.

Lin, E. H., Rutter, C. M., Katon, W., Heckbert, S. R., Ciechanowski, P., Oliver, M. M., McCulloch, D. K. (2010). Depression and Advanced Complications of Diabetes A prospective cohort study. *Diabetes Care, 33*(2), 264-269.

Lipton, Z. C., Kale, D. C., Elkan, C., & Wetzel, R. (2015). Learning to diagnose with LSTM recurrent neural networks. *arXiv preprint arXiv:1511.03677*.

Loeb, D. F., Binswanger, I. A., Candrian, C., & Bayliss, E. A. (2015). Primary care physician insights into a typology of the complex patient in primary care. *The Annals of Family Medicine, 13*(5), 451-455.

Lorant, V., Deliège, D., Eaton, W., Robert, A., Philippot, P., & Ansseau, M. (2003). Socioeconomic inequalities in depression: a meta-analysis. *American Journal of Epidemiology, 157*(2), 98-112.

Marmot, M., Allen, J., Bell, R., Bloomer, E., & Goldblatt, P. (2012). WHO European review of social determinants of health and the health divide. *The Lancet, 380*(9846), 1011-1029.

Marmot, M., Friel, S., Bell, R., Houweling, T. A., Taylor, S., & Health, C. o. S. D. o. (2008). Closing the gap in a generation: health equity through action on the social determinants of health. *The Lancet, 372*(9650), 1661-1669.

Martin Sanchez, F., Gray, K., Bellazzi, R., & Lopez-Campos, G. (2014). Exposome informatics: considerations for the design of future biomedical research information systems. *Journal of the American Medical Informatics Association*, 386-390.

Mathers, C. (2008). *The global burden of disease: 2004 update*: World Health Organization.

Mathers, C., Boerma, T., & Fat, D. M. (2008). *The global burden of disease: 2004 update*: World Health Organization.

Mathers, C., Fat, D. M., & Boerma, J. T. (2008). *The global burden of disease: 2004 update*: World Health Organization.

McAfee, A., Brynjolfsson, E., Davenport, T. H., Patil, D., & Barton, D. (2012). Big data: the management revolution. *Harvard Business Review, 90*(10), 60-68.

McDonald, C. J., Overhage, J. M., Barnes, M., Schadow, G., Blevins, L., Dexter, P. R. (2005). The Indiana network for patient care: a working local health information infrastructure. *Health Affairs, 24*(5), 1214-1220.

McGinnis, J. M., & Foege, W. H. (1993). Actual causes of death in the United States. *Journal of the American Medical Association, 270*(18), 2207-2212.

McPherson, S., & Armstrong, D. (2006). Social determinants of diagnostic labels in depression. *Social Science & Medicine, 62*(1), 50-58.

Miotto, R., Wang, F., Wang, S., Jiang, X., & Dudley, J. T. (2017). Deep learning for healthcare: review, opportunities and challenges. *Briefings in Bioinformatics*.

Mojtabai, R. (2013). Clinician-identified depression in community settings: concordance with structured-interview diagnoses. *Psychotherapy and Psychosomatics, 82*(3), 161-169.

Mojtabai, R., & Olfson, M. (2011). Proportion of antidepressants prescribed without a psychiatric diagnosis is growing. *Health Affairs, 30*(8), 1434-1442.

Motik, B., Patel-Schneider, P. F., Parsia, B., Bock, C., Fokoue, A., Haase, P., Sattler, U. (2009). OWL 2 web ontology language: Structural specification and functional-style syntax. *W3C Recommendation, 27*(65), 159.

Muhrer, J. C. (2010). Detecting and dealing with substance abuse disorders in primary care. *The Journal for Nurse Practitioners, 6*(8), 597-605.

National Institute of Mental Health. (2015a). Depression: What You Need To Know. Retrieved from https://www.nimh.nih.gov/health/publications/depression-what-you-need-to-know/index.shtml

National Institute of Mental Health. (2015b). Depression. What you need to know. Retrieved from http://www.nimh.nih.gov/health/publications/depression-what-you-need-to-know-12-2015/depression-what-you-need-to-know-pdf_151827.pdf

Ng, T., Niti, M., Tan, W., Cao, Z., Ong, K., & Eng, P. (2007). Depressive symptoms and chronic obstructive pulmonary disease: effect on mortality, hospital readmission, symptom burden, functional status, and quality of life. *Archives of Internal Medicine, 167*(1), 60-67.

Nielsen, T. D., & Jensen, F. V. (2009). *Bayesian networks and decision graphs*: Springer Science & Business Media.

Nouwen, A., Winkley, K., Twisk, J., Lloyd, C., Peyrot, M., Ismail, K. (2010). *Type 2 diabetes mellitus as a risk factor for the onset of depression: a systematic review and meta-analysis*: Springer.

Office of Disease Prevention and Health Promotion. (2014). Social Determinants of Health. Retrieved from https://www.healthypeople.gov/2020/topics-objectives/topic/social-determinants-of-health#five

Onyike, C. U., Crum, R. M., Lee, H. B., Lyketsos, C. G., & Eaton, W. W. (2003). Is obesity associated with major depression? Results from the Third National Health and Nutrition Examination Survey. *American Journal of Epidemiology, 158*(12), 1139-1147.

Overhage, J. M. (2013). Trends in Health Information Exchanges. Retrieved from https://innovations.ahrq.gov/perspectives/trends-health-information-exchanges

Overhage, J. M. (2016). The Indiana Health Information Exchange. In B. E. Dixon (Ed.), *Health Information Exchange: Navigating and Managing a Network of Health Information Systems* (1 ed., pp. 267-279). Waltham, MA: Academic Press.

Patel, V., Lund, C., Hatherill, S., Plagerson, S., Corrigall, J., Funk, M., & Flisher, A. J. (2010). Mental disorders: equity and social determinants. *Equity, Social Determinants and Public Health Programmes, 115*.

Pearson Education. (2016). Beck Depression Inventory. Retrieved from http://www.pearsonclinical.com/psychology/products/100000159/beck-depression-inventoryii-bdi-ii.html?Pid=015-8018-370&Mode=summary

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., . . . Dubourg, V. (2011). Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research, 12*(Oct), 2825-2830.

Peng, H., Long, F., & Ding, C. (2005). Feature selection based on mutual information criteria of max-dependency, max-relevance, and min-redundancy. *Pattern Analysis and Machine Intelligence, IEEE Transactions on, 27*(8), 1226-1238.

Pohjasvaara, T., Vataja, R., Leppävuori, A., Kaste, M., & Erkinjuntti, T. (2001). Depression is an independent predictor of poor long-term functional outcome post-stroke. *European Journal of Neurology, 8*(4), 315-319.

Polani, D. (2013). Kullback-leibler divergence *Encyclopedia of Systems Biology* (pp. 1087-1088): Springer.

Polis Center. (2018). About us. Retrieved from http://polis.iupui.edu/index.php/about-us/

Prati, R. C., Batista, G. E., & Monard, M. C. (2009). Data mining with imbalanced class distributions: concepts and methods. Paper presented at the IICAI.

programmableweb.com. (2017). Zillow API. Retrieved from https://www.programmableweb.com/api/zillow

Qi, Y. (2012). Random forest for bioinformatics *Ensemble Machine Learning* (pp. 307-323): Springer.

Quinlan, J. R. (1986). Induction of decision trees. *Machine Learning, 1*(1), 81-106.

Radloff, L. (1977). Center for Epidemiological Studies Depression Scale. *Applied Psychological Measurement, 1*, 385-401.

Radloff, L. S. (1977). The CES-D scale: A self-report depression scale for research in the general population. *Applied Psychological Measurement, 1*(3), 385-401.

Saloheimo, H. P., Markowitz, J., Saloheimo, T. H., Laitinen, J. J., Sundell, J., Huttunen, M. O., . . . Katila, H. O. (2016a). Psychotherapy effectiveness for major depression: a randomized trial in a Finnish community. *BMC Psychiatry, 16*(1), 1.

Saloheimo, H. P., Markowitz, J., Saloheimo, T. H., Laitinen, J. J., Sundell, J., Huttunen, M. O., . . . Katila, H. O. (2016b). Psychotherapy effectiveness for major depression: a randomized trial in a Finnish community. *BMC Psychiatry, 16*(1), 131.

Santos, M. J. d., Kawamura, H. C., & Kassouf, A. L. (2012). Socioeconomic conditions and risk of mental depression: An empirical analysis for Brazilian citizens. *Economics Research International, 2012*.

Saveanu, R. V., & Nemeroff, C. B. (2012). Etiology of depression: genetic and environmental factors. *Psychiatric Clinics, 35*(1), 51-71.

Schadow, G., & McDonald, C. J. (2005). The unified code for units of measure (UCUM). *Regenstrief Institute and Indiana University School of Informatics*.

Schmidhuber, J. (2015). Deep learning in neural networks: An overview. *Neural Networks, 61*, 85-117.

Schneeweiss, S. (2014). Learning from big health care data. *New England Journal of Medicine, 370*(23), 2161-2163.

Schuurman, N., Bell, N., Dunn, J. R., & Oliver, L. (2007). Deprivation indices, population health and geography: an evaluation of the spatial effectiveness of indices at multiple scales. *Journal of Urban Health, 84*(4), 591-603.

Schwartz, T. L., & Petersen, T. (2016). *Depression: Treatment Strategies and Management*: CRC Press.

Scott, D. W. (2009). Sturges' rule. *Wiley Interdisciplinary Reviews: Computational Statistics, 1*(3), 303-306.

Shafer, A. B. (2006). Meta-analysis of the factor structures of four depression questionnaires: Beck, CES-D, Hamilton, and Zung. *Journal of Clinical Psychology, 62*(1), 123-146.

Sharabiani, M. T., Aylin, P., & Bottle, A. (2012). Systematic review of comorbidity indices for administrative data. *Medical Care, 50*(12), 1109-1118.

Sharp, L. K., & Lipsky, M. S. (2002). Screening for depression across the lifespan. *American Family Physician, 66*, 1001-1008.

Simeonsson, R. J., Leonardi, M., Lollar, D., Bjorck-Akesson, E., Hollenweger, J., & Martinuzzi, A. (2003). Applying the International Classification of Functioning, Disability and Health (ICF) to measure childhood disability. *Disability and Rehabilitation, 25*(11-12), 602-610.

Skinner, M. K. (2011). Environmental epigenomics and disease susceptibility. *EMBO Reports, 12*(7), 620-622.

Solar, O., & Irwin, A. (2010). A conceptual framework for action on the social determinants of health.

Substance Abuse and Mental Health Services Administration. (2014). Behavioral Health Trends in the United States: Results from the 2014 National Survey on Drug Use and Health. Retrieved from http://www.samhsa.gov/data/sites/default/files/NSDUH-FRR1-2014/NSDUH-FRR1-2014.pdf

Tanner, E. K., Martinez, I. L., & Harris, M. (2014). Examining functional and social determinants of depression in community-dwelling older adults: Implications for practice. *Geriatric Nursing, 35*(3), 236-240.

Targum, S. D., & Rosenthal, N. (2008). Seasonal affective disorder. *Psychiatry (Edgmont), 5*(5), 31.

Thombs, B. D., Coyne, J. C., Cuijpers, P., De Jonge, P., Gilbody, S., Ioannidis, J. P., Ziegelstein, R. C. (2012). Rethinking recommendations for screening for

depression in primary care. *Canadian Medical Association Journal, 184*(4), 413-418.
Thombs, B. D., Ziegelstein, R. C., Roseman, M., Kloda, L. A., & Ioannidis, J. P. (2014). There are no randomized controlled trials that support the United States Preventive Services Task Force guideline on screening for depression in primary care: a systematic review. *BMC Medicine, 12*(1), 13.
Tibshirani, R. (1996). Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society. Series B (Methodological)*, 267-288.
Torres, O., Rothberg, M. B., Garb, J., Ogunneye, O., Onyema, J., & Higgins, T. (2015). Risk factor model to predict a missed clinic appointment in an urban, academic, and underserved setting. *Population Health Management, 18*(2), 131-136.
Trinh, N., Youn, S. J., Sousa, J., Regan, S., Bedoya, C. A., Chang, T. E., Yeung, A. (2011). Using electronic medical records to determine the diagnosis of clinical depression. *International Journal of Medical Informatics, 80*(7), 533-540.
Turner, C. R., Wolf, A. L., Fuggetta, A., & Lavazza, L. (1998). Feature engineering. Paper presented at the Proceedings of the 9th international workshop on software specification and design.
U.S. National Library of Medicine. (2013). Fact Sheet: UMLS Metathesaurus. Retrieved from https://www.nlm.nih.gov/pubs/factsheets/umlsmeta.html
U.S. National Library of Medicine. (2015). Fact Sheet: Unified Medical Language System. Retrieved from https://www.nlm.nih.gov/pubs/factsheets/umls.html
U.S. National Library of Medicine. (2016). UMLS Metathesaurus. Retrieved from https://www.nlm.nih.gov/research/umls/knowledge_sources/metathesaurus/
Valenstein, M., Vijan, S., Zeber, J. E., Boehm, K., & Buttar, A. (2001). The cost-utility of screening for depression in primary care. *Annals of Internal Medicine, 134*(5), 345-360.
Vest, J. R., & Gamm, L. D. (2010). Health information exchange: persistent challenges and new strategies. *Journal of the American Medical Informatics Association, 17*(3), 288-294.
Wade, T. D., Bulik, C. M., Neale, M., & Kendler, K. S. (2000). Anorexia nervosa and major depression: shared genetic and environmental risk factors. *American Journal of Psychiatry*.
Wagner, H. R., Burns, B., Broadhead, W., Yarnall, K., Sigmon, A., & Gaynes, B. (2000). Minor depression in family practice: functional morbidity, co-morbidity, service utilization and outcomes. *Psychological Medicine, 30*(06), 1377-1390.
Wells, K. B., Sherbourne, C., Schoenbaum, M., Duan, N., Meredith, L., Unützer, J., Rubenstein, L. V. (2000). Impact of disseminating quality improvement programs for depression in managed primary care: a randomized controlled trial. *Journal of the American Medical Association, 283*(2), 212-220.
West, D., Mangiameli, P., Rampal, R., & West, V. (2005). Ensemble strategies for a medical diagnostic decision support system: A breast cancer diagnosis application. *European Journal of Operational Research, 162*(2), 532-551.
Williams, J. B. (2001). Standardizing the Hamilton Depression Rating Scale: past, present, and future. *European Archives of Psychiatry and Clinical Neuroscience, 251*(2), 6-12.

Williams Jr, J. W., Barrett, J., Oxman, T., Frank, E., Katon, W., Sullivan, M., Sengupta, A. (2000). Treatment of dysthymia and minor depression in primary care: a randomized controlled trial in older adults. *Journal of the American Medical Association, 284*(12), 1519-1526.

Winters, N. C., & Metz, W. P. (2009). The wraparound approach in systems of care. *Psychiatric Clinics, 32*(1), 135-151.

Wong, M.-L., Dong, C., Andreev, V., Arcos-Burgos, M., & Licinio, J. (2012). Prediction of susceptibility to major depression by a model of interactions of multiple functional genetic variants and environmental factors. *Molecular Psychiatry, 17*(6), 624-633.

World Health Organization. (2010). A conceptual framework for action on the social determinants of health.

World Health Organization. (2012). Depression: A Global Crisis. World Mental Health Day, October 10 2012. *World Federation for Mental Health, Occoquan, Va, USA*.

World Health Organization. (2013a). Information sheet: Premature death among people with severe mental disorders. Retrieved from http://www.who.int/mental_health/management/info_sheet.pdf

World Health Organization. (2013b). Information sheet: Premature death among people with severe mental disorders. Retrieved from http://www.who.int/mental_health/management/info_sheet.pdf

World Health Organization. (2013c). Information sheet. Premature death among people with severe mental disorders. Retrieved from http://www.who.int/mental_health/management/info_sheet.pdf

World Health Organization. (2014). Social determinants of mental health: World Health Organization.

World Health Organization. (2016). Premature death among people with severe mental disorders. Retrieved from http://www.who.int/mental_health/management/info_sheet.pdf

Yohannes, A., Willgoss, T., Baldwin, R., & Connolly, M. (2010). Depression and anxiety in chronic heart failure and chronic obstructive pulmonary disease: prevalence, relevance, clinical implications and management principles. *International Journal of Geriatric Psychiatry, 25*(12), 1209-1221.

Youden, W. J. (1950). Index for rating diagnostic tests. *Cancer, 3*(1), 32-35.

Yu, C., Lin, J, & Lin-Tan, D. (2004). Environmental exposure to lead and progression of chronic renal diseases: a four-year prospective longitudinal study. *Journal of the American Society of Nephrology, 15*(4), 1016-1022.

Zhang, P. (1993). Model selection via multifold cross validation. *The Annals of Statistics*, 299-313.

# 11 CURICULUM VITAE

Suranga N. Kasthurirathne

Education

- Indiana University — School of Informatics and Computing
  Indianapolis, IN, USA

  Degree: Doctor of Philosophy,
  Public Health (Minor)

- University of Westminster — University of Westminster, UK

  Degree: Bachelor of Engineering in Software Engineering

Research interests

- Use of standardized data exchange, data analytics, Natural Language Processing, and machine learning to address a multitude of healthcare challenges.

Training

- Fall 2013 to present: Research Assistantship at the Regenstrief Institute
  Participated in various research projects at the Regenstrief Institute.

- Summer 2016: Epidemiology and Population Health Summer Institute at Colombia University (EPIC).
  Scholarship to attend the Transforming Public Health Surveillance summer program.

Teaching experience

- HIM-M200 Database Design for Health Information Management, Fall 2015 to Summer 2018.

Peer reviewed publications

- Kasthurirathne, S. N., Vest, J. R., Menachemi, N., Halverson, P. K., & Grannis, S. J. (2017). Assessing the capacity of social determinants of health data to augment predictive models identifying patients in need of wraparound social services. *Journal of the American Medical Informatics Association*.
- Kasthurirathne, S. N., Dixon, B. E., Gichoya, J., Xu, H., Xia, Y., Mamlin, B., & Grannis, S. J. (2017). Toward better public health reporting using existing off the shelf approaches: The value of medical dictionaries in automated cancer detection using plaintext medical data. *Journal of Biomedical Informatics*, *69*, 160-176

- Kasthurirathne, S. N., Mamlin, B. W., & Cullen, T. (2017). Leveraging the value of human relationships to improve health outcomes. *Applied Clinical Informatics*, *8*(1), 108-121
- Kasthurirathne, S. N., Dixon, B. E., Gichoya, J., Xu, H., Xia, Y., Mamlin, B., & Grannis, S. J. (2016). Toward better public health reporting using existing off the shelf approaches: A comparison of alternative cancer detection approaches using plaintext medical data and non-dictionary based feature selection. *Journal of Biomedical Informatics*, *60*, 145-152.
- Kasthurirathne, S. N., Mamlin, B., Kumara, H., Grieve, G., & Biondich, P. (2015). Enabling better interoperability for healthcare: Lessons in developing a standards based application programing interface for electronic medical record systems. *Journal of Medical Systems*, *39*(11), 182.

Posters

- Kasthurirathne, S. N., Mamlin, B., Grieve, G., & Biondich, P. (2015). Towards standardized patient data exchange: Integrating a FHIR based API for the Open Medical Record System. Studies in Health Technology and Informatics. 216: 932-932.
- Kasthurirathne, S. N., Dixon, B. E., & Grannis, S. J. (2015). Evaluating methods for identifying cancer in free-text pathology reports using various machine learning and data pre-processing approaches. Studies in Health Technology and Informatics, 216, 1070-1070.

Presentations

- Kasthurirathne, S. N., Purkayastha, S. (2016). The developmental impact of infrastructures: A case of OpenMRS interoperability. Presented at the Regenstrief Work in Progress session, IN, USA.
- Gichoya, J.W., Kasthurirathne S.N., Genereaux, B., Karunarathne, M., Purkayastha, S., Kahn, C.E. (2016). Using HL7 FHIR to integrate structured reporting reports into an Open Source Medical Records System. Proceedings of the Society of Imaging Informatics in Medicine.
- Kasthurirathne, S. N., Mamlin, B., Kumara, H., Grieve, G., & Biondich, P. (2015). The OpenMRS FHIR module: A standardized API for the OpenMRS platform. Presented at Medinfo 2015, São Paulo, Brazil.