



2019

The Prevention Of Depression: A Machine Learning Approach

Ashley Ann Fuss

University of Pennsylvania, ashleyfuss@gmail.com

Follow this and additional works at: <https://repository.upenn.edu/edissertations>



Part of the [Social Work Commons](#)

Recommended Citation

Fuss, Ashley Ann, "The Prevention Of Depression: A Machine Learning Approach" (2019). *Publicly Accessible Penn Dissertations*. 3323.
<https://repository.upenn.edu/edissertations/3323>

This paper is posted at ScholarlyCommons. <https://repository.upenn.edu/edissertations/3323>

For more information, please contact repository@pobox.upenn.edu.

The Prevention Of Depression: A Machine Learning Approach

Abstract

Behavioral health disorders, specifically depression, are a serious health concern in the United States and worldwide. The consequences of unaddressed behavioral health conditions are multifaceted and have impact at the individual, relational, communal, and societal level. Despite the number of individuals who could benefit from treatment for behavioral health concerns, their difficulties are often unidentified and unaddressed through treatment. Technology carries unrealized potential to identify people at risk for behavioral health conditions and to inform prevention and intervention strategies. Drawing upon data from the National Longitudinal Study of Adolescent Health (Add Health, n=3782), this study has two aims related to advancing understanding of technology's potential value in behavioral health: 1) to develop a forecasting procedure that can be used to identify youth who are at risk of reporting a depression diagnosis as adults based on a set of input variables; and 2) to understand the developmental trajectories of depression for youth. To address the first aim of this study, random forest methodology was used to derive the forecasting algorithm. The second aim was pursued with Generalized Additive Model analysis to estimate relationships between presence of a reported depression diagnosis as an adult and youth characteristics. Findings from this study indicate that it is feasible to use a forecasting tool to identify individuals at risk of being diagnosed with depression, which can facilitate early intervention and improved outcomes. Gender, race, and receiving counseling as a youth were the most important predictors of having a reported depression diagnosis as an adult. This dissertation addresses the role of health disparities, specifically gender and race, related to depression and mental health treatment. In sum, this dissertation highlights how a machine learning forecasting tool could be used to inform prevention strategies and understanding of factors associated with receiving a depression diagnosis. This study presents and discusses these findings in addition to offering important implications for future research and practice to identify and prevent behavioral health conditions such as depression.

Degree Type

Dissertation

Degree Name

Doctor of Philosophy (PhD)

Graduate Group

Social Welfare

First Advisor

Malitta Engstrom

Keywords

Behavioral Health, Depression, Machine Learning, Prevention

Subject Categories

Social Work

THE PREVENTION OF DEPRESSION: A MACHINE LEARNING APPROACH

Ashley Ann Fuss

A DISSERTATION

in

Social Welfare

Presented to the Faculties of the University of Pennsylvania

in

Partial Fulfillment of the Requirements for the

Degree of Doctor of Philosophy

2019

Supervisor of Dissertation

Malitta Engstrom

Associate Professor, School of Social Policy & Practice

Graduate Group Chairperson

Femida Handy, Professor, School of Social Policy & Practice

Dissertation Committee

Malitta Engstrom, PhD, Associate Professor, University of Pennsylvania

Dennis P. Culhane, PhD, Professor, University of Pennsylvania

Richard Berk, PhD, Professor, University of Pennsylvania

Mary M. McKay, PhD, Dean, Washington University in St. Louis

THE PREVENTION OF DEPRESSION: A MACHINE LEARNING APPROACH

COPYRIGHT

2019

Ashley Ann Fuss

This work is licensed under the
Creative Commons Attribution-
NonCommercial-ShareAlike 3.0
License

To view a copy of this license, visit

<https://creativecommons.org/licenses/by-nc-sa/3.0/us/>

DEDICATION

“Behind every successful woman is herself.”
~Bart Jackson

Love you, mom!

ACKNOWLEDGMENT

I would first like to thank my advisor, dissertation chair, and mentor, Malitta Engstrom. Thank you so much for your continued support, patience, dedication, encouragement, and knowledge throughout this PhD journey. I am forever grateful for everything. Richard Berk, thank you, thank you, thank you, for your mentorship and for sharing your expertise with me. I have learned so much from you and could not have done this without you. I would like to thank my committee members, Dennis Culhane and Mary McKay. Your support and wisdom have meant so much to me.

I would also like to thank my amazing family, friends, and colleagues. I am so lucky to have so many wonderful people by my side who have supported me endlessly and cheered me on when I needed it the most. I appreciate each of you more than you will ever know.

ABSTRACT

THE PREVENTION OF DEPRESSION: A MACHINE LEARNING APPROACH

Ashley Ann Fuss

Malitta Engstrom, Ph.D.

Behavioral health disorders, specifically depression, are a serious health concern in the United States and worldwide. The consequences of unaddressed behavioral health conditions are multifaceted and have impact at the individual, relational, communal, and societal level. Despite the number of individuals who could benefit from treatment for behavioral health concerns, their difficulties are often unidentified and unaddressed through treatment. Technology carries unrealized potential to identify people at risk for behavioral health conditions and to inform prevention and intervention strategies. Drawing upon data from the National Longitudinal Study of Adolescent Health (Add Health, n=3782), this study has two aims related to advancing understanding of technology's potential value in behavioral health: 1) to develop a forecasting procedure that can be used to identify youth who are at risk of reporting a depression diagnosis as adults based on a set of input variables; and 2) to understand the developmental trajectories of depression for youth. To address the first aim of this study, random forest methodology was used to derive the forecasting algorithm. The second aim was pursued with Generalized Additive Model analysis to estimate relationships between presence of a reported depression diagnosis as an adult and youth characteristics. Findings from this study indicate that it is feasible to use a forecasting tool to identify individuals at risk of being diagnosed with depression, which can facilitate early intervention and improved outcomes. Gender, race, and receiving counseling as a youth were the most important predictors of having a reported depression diagnosis as an adult. This dissertation

addresses the role of health disparities, specifically gender and race, related to depression and mental health treatment. In sum, this dissertation highlights how a machine learning forecasting tool could be used to inform prevention strategies and understanding of factors associated with receiving a depression diagnosis. This study presents and discusses these findings in addition to offering important implications for future research and practice to identify and prevent behavioral health conditions such as depression.

TABLE OF CONTENTS

ABSTRACT	V
LIST OF TABLES	VIII
LIST OF ILLUSTRATIONS	IX
CHAPTER 1: INTRODUCTION	1
Background and Significance	4
Present Study	26
CHAPTER 2: METHOD	29
Research Design	29
Data and Sample	30
Measures	31
Data Analytic Strategy	38
CHAPTER 3: RESULTS	43
Random Forests Results	51
Generalized Additive Model Results	59
CHAPTER 4: DISCUSSION	68
Depression Forecasting Tool	68
Depression Trajectories	72
Youth Characteristics	80
Limitations of the Research	81
Implications	82
Conclusion	87
BIBLIOGRAPHY	88

LIST OF TABLES

Table 1: Inputs for Forecasting	32
Table 2: Characteristics of the Study Sample	43
Table 3: Adults with a Reported Depression Diagnosis Compared to Adults Without a Reported Depression Diagnosis	49
Table 4: Confusion Table for Forecasting a Reported Depression Diagnosis	52
Table 5: GAM Summary Statistics	60
Table 6: GAM Results	62

LIST OF ILLUSTRATIONS

Figure 1: Variable Importance Plot for a Reported Depression Diagnosis	53
Figure 2: Partial Dependence Plots for a Reported Depression Diagnosis on Binary Inputs	57
Figure 3: Partial Dependence Plots for a Reported Depression Diagnosis on Quantitative Inputs	58
Figure 4: Correlates of a Reported Depression Diagnosis and Predictor Variables	66
Figure 5: Reported Depression Diagnosis Fitted Values from the GAM Procedure	67

CHAPTER 1: INTRODUCTION

Behavioral health disorders are a pervasive problem among children and young adults in the United States. Each year about 20%, or one in five children, experience a mental health condition (Perou et al., 2013). Approximately one in eight children suffers from a psychiatric disorder that is serious enough to cause functional impairment (Costello, Egger, & Angold, 2005). Depression is one of the most common mental health disorders, and by 2030 is predicted to be the top contributor to the Global Burden of Disease (GBD) (Mathers & Loncar, 2006). While there is no single cause of behavioral health disorders, genetic, family, and environmental factors have been associated with developing a condition (NAMI, 2016). The stress associated with poverty and early traumatic experiences have also been linked with the development of mental health difficulties (Carr et al., 2013; Gopalan et al., 2010). Symptoms of mental health conditions can be reduced with services; however, if left untreated, mental health challenges can have serious, multifaceted, consequences.

The impact of behavioral health disorders are costly at the individual and societal level. Behavioral health conditions cost our healthcare system an estimated \$57 billion a year, which is 2.5 to 3.5 times more than individuals without a behavioral health condition (Klein & Hostetter, 2014). Depression drives the largest percentage of health care costs and accounts for more years lost to disability than any other disease (Smith, 2014; Watson Health, 2016). The consequences of behavioral health conditions on individuals' physical health, education, employment, lifetime earnings, and relationships are also well-documented (Alonso et al., 2011; Kessler et al., 2008; Porche, Costello, & Rosen-Reynoso, 2016; Parks, Svendsen, Singer, Foti, & Mauer, 2006).

Despite the number of individuals experiencing mental health difficulties, a significant gap exists between the number of individuals who need treatment and

individuals who have access to treatment. Most children with behavioral health conditions do not receive treatment. A report by the National Institute of Mental Health found that 75% of children with mental health problems did not receive any treatment (McKay et al., 2004). Additionally, the disparity between need and use is high for children who reside in poverty-impacted communities and have serious problems with stressful home environments. (McKay et al., 2004).

Identifying children with behavioral health challenges can also be difficult. While parents are usually the first to recognize difficulties their children may be experiencing, internalized behavioral health symptoms, such as depression, may be hidden from parents or hard to pick up on (Levitt et al., 2007; Logan & King, 2002). In the primary care setting, mental health challenges are also under-identified, and research suggests that only 20% of children in need of treatment are identified by physicians (Pidano et al., 2011; Sayal, 2006; Simonian, 2006).

Health care reform and the passage of the Affordable Care Act (ACA) have helped millions of Americans gain coverage. The three goals of health reform have been coined the “triple aim” and seek to improve population health, improve the patient experience, and reduce per capita costs (Berwick, Nolan, & Whittington, 2008). To achieve these goals, payers and providers are turning toward population health approaches that focus on health promotion and well-being (Rawal & McCabe, 2016). Health promotion is a component of prevention, and behavioral health promotion is characterized by a focus on well-being rather than prevention of an illness (O’Connell, Boat, & Warner, 2009). This shift to health promotion has highlighted the critical role that behavioral health plays in a person’s overall health and wellness by acknowledging the connection between physical and behavioral health.

There is also growing interest in being able to identify populations with high needs, and predictive analytics have been identified as one approach to help achieve the Triple Aim (Amarasingham et al., 2014; Rawal & McCabe, 2016). Predictive analytics, a type of machine learning, is often standard practice in the private sector and is used to inform decision-making related to sales, trading stocks, and giving out loans. Recently, machine learning methods have been applied in the public sector in areas such as criminal justice, domestic violence, child welfare, education, and physical health and are being used to inform policy and practice decisions (Berk, Sorenson, & Barnes, 2016; Berk, Sherman, Barnes, Kurtz, & Ahlman, 2009; Gillingham, 2015; Russell, 2015; Shams, Ajorlou, & Yang, 2014).

Technology that can predict patient outcomes such as hospital readmission is one approach that can be used to inform prevention strategies. Technology has been called the future of mental health treatment by the National Institute of Mental Health, and between FY 2009 and FY 2015 404 grants totaling \$445 million focused on technology (e.g. mobile apps) and mental health were awarded (NIMH, 2017). Machine learning in health care offers promising solutions to improve diagnosis and treatment, and some of the most exciting new advances in health care today can be attributed to machine learning technology (Marr, 2017).

While machine learning approaches in the context of behavioral health are still new, they offer tremendous opportunity and potential to be a tool for behavioral health prevention and health promotion. Being able to identify individuals at risk of developing a behavioral health disorder and being able to intervene before functioning is impaired is important for individuals, their families, and healthcare systems.

The first aim of this study is to develop a forecasting procedure that can be used to identify youth who are at risk of developing a depressive disorder as an adult and

could benefit from prevention or support services. This procedure seeks to predict whether a youth will have a diagnosed behavioral health condition, specifically depression as an adult, based on a set of input variables. The second aim of this study is to understand the developmental trajectories of depression for youth. This study addresses the following research questions drawing upon data from the National Longitudinal Study of Adolescent Health (Add Health): How well do random forests forecasts perform in terms of predicting which youth will report a depression diagnosis as an adult? What features distinguish youth with depressive symptoms who report a depression diagnosis as an adult from youth with depressive symptoms who do not report a depression diagnosis as an adult?

Background and Significance

In the United States, the estimated lifetime prevalence of having at least one mental health condition is 47.4% (Kessler et al., 2007). According to the Report of the Surgeon General's Conference on Children's Mental Health, one in five children has a mental health disorder (US Department of Health and Human Services, 2000). Other research estimates that between 20% and 40% of children experience an existing mental health problem (Costello, Copeland & Angold, 2011).

Depressive disorders are one of the most common mental health conditions in the United States and worldwide (Kessler et al., 2007). *The Diagnostic and Statistical Manual of Mental Disorders*, Fifth Edition (DSM-5), which is used by clinicians to diagnose and classify mental health disorders, defines features of depressive disorders as "the presence of sad, empty, or irritable mood, accompanied by somatic and cognitive changes that significantly affect the individual's capacity to function. What differs among them are issues of duration, timing, or presumed etiology" (American Psychiatric

Association, 2013, Chapter 5). Common depressive disorders include Major Depressive Disorder and Dysthymia. Adults with depressive disorders often experience sadness or the inability to feel pleasure while children with depression often demonstrate extreme irritability (American Psychiatric Association, 2013). In children, behavioral problems and attention-deficit hyperactivity disorder are also among the most prevalent mental health conditions and account for about 50% of referrals to mental health clinics (Kazdin, 1995). Children are often brought by concerned parents to mental health clinics for exhibiting disruptive behaviors, and children with disruptive behavior problems are at increased risk of a co-occurring psychiatric disorder (Fossum, Morch, Handegard & Drugli, 2007; Petittclerc et al., 2009). Internalizing child mental health difficulties are also problematic for youth, with estimates ranging from 35% to 65% of youth experiencing symptoms related to depression and/or anxiety (Horowitz, McKay, & Marshall, 2005). While symptoms associated with depression, anxiety, and behavioral problems can be improved with mental health services, if left untreated, mental health challenges can have serious multifaceted consequences.

Consequences of Behavioral Health Disorders. The costs of mental health disorders are significant at the individual and societal level. Health care costs among children with mental health conditions are estimated to be two times that of healthy children because of their mental health conditions, and a 2013 study estimated that child mental health conditions cost \$247 billion dollars annually (Perou et al., 2013). Depression drives the largest percentage of health care costs and accounts for more years lost to disability than any other disease globally (Smith, 2014; Watson Health, 2016).

Individuals with mental health conditions die on average 25 years sooner than individuals without mental health conditions due primarily to preventable co-occurring

physical health problems (Parks, Svendsen, Singer, Foti, & Mauer, 2006). Individuals with a serious physical health problem often have a co-occurring behavioral health condition, and an estimated 70% of primary care visits can be attributed to psychosocial problems (Collins, Hewson, Munger, & Wade, 2010). For individuals with behavioral health conditions and a chronic medical condition, the annual medical costs were 46% higher compared to individuals with a chronic medical condition (Patient-Centered Primary Care Collaborative, 2014). Depression impacts daily functioning and is linked with poor quality of life, worsening of co-existing physical health conditions, and increased risk of developing other non-communicable diseases and communicable diseases such as HIV and other sexually transmitted infections (STIs; Brown et al., 2006; Charlson et al., 2013; Elkington, Bauermeister, & Zimmerman, 2010; Indu et al., 2017).

Educational attainment and employment are also impacted by mental health disorders. Porche, Costello, and Rosen-Reynoso (2016) found that children with mental health conditions were less engaged in school and were more likely to have an Individualized Education Program (IEP) and be retained in a grade. Behavioral health difficulties also put youth at risk for high school incompleteness and academic failure (Moore, Redd, Burkhauser, Mbwana, & Collins, 2009; Vander Stoep, Weiss, Kuo, Cheney, & Cohen, 2003).

Mental health conditions are associated with decreased work productivity and more days out of work (Alonso et al., 2011). Previous research has estimated that annually in the United States, 3.6 billion days of work are missed due to health-related problems (Merikangas et al., 2007). Alonso et al.'s 2011 study found that on average, individuals with a mental health condition missed 31 more days of work a year, relative to individuals with no mental health condition. Over the course of a year, individuals with

a mental health problem earned on average \$16,306 less than individuals without a mental health condition (Kessler et al., 2008). Furthermore, in 2002, mental illness was associated with a \$193.2 billion reduction in personal earnings in the United States (Kessler et al., 2008).

Prevention and Health Promotion. Research suggests that the earlier behavioral health problems are addressed, the better the outcomes are, and that intervening early can reduce the likelihood of long-term impairment (Koppelman, 2004). Health promotion is a component of prevention; behavioral health promotion's emphasis is on well-being, rather than prevention of an illness (National Research Council, 2009). Mental health promotion focuses on building individual and community level strengths and resources that can be used to improve mental health functioning (Barry & Jenkins, 2007). Prevention and health promotion are similar in that "both focus on changing common influences on the development of children and adolescents to aid them in functioning well in meeting life's tasks and challenges and remaining free of cognitive, emotional, and behavioral problems that would impair their functioning" (National Research Council, 2009, p. 59).

Reducing psychosocial risk factors and building protective factors can prevent the development of behavioral health conditions (U.S. Department of Health & Human Services, 1999). As described by Sussman and Ames (2008), it is important to consider the interplay between factors at multiple levels including individual (e.g., genetics), microsocial (e.g., parenting), and macrosocial (e.g., school) systems when thinking about the prevention of behavioral health problems.

Mental health promotion operates under the assumption that mental health is a positive concept and is a fundamental element of the broader public health and health promotion agenda (Herrman, 2012). When mental health is presented in a positive way

to children, it provides them with critical skills, resources, and supports they can draw upon when faced with stress and adversity (Barry & Jenkins, 2007; Patel et al., 2007). Behavioral health promotion and prevention interventions for youth in collaboration with families have been found to not only improve mental health functioning, but also improve academic performance, interpersonal and communication skills, self-esteem, coping strategies, and overall health behavior functioning (Barry et al., 2013; Durlak et al., 2011; Tennant et al., 2007; Weare & Nind, 2011; Wilson & Lipsey, 2007). Youth mentoring programs have also been shown to promote behavioral health and are also associated with improved psychosocial and behavioral outcomes (DeWit et al., 2016; DuBois, Portillo, Rhodes, Silverhorn, & Valentine 2011; Tolan, Henry, Schoeny, Lovegrove, & Nichols, 2014).

Conley, Durlak, and Dickson (2013) highlight that skill-oriented prevention programs, for example, those that contain mindfulness or supervised practice, are effective at improving youth outcomes whereas psychoeducational programs are rarely effective at changing youth behavior.

Behavioral health is key to good health and impacts outcomes across an individual's lifespan (Herrman, 2012; Jenkins et al., 2011). While there is an upfront cost associated with delivering prevention and behavioral health promotion interventions, there is evidence that long-term, offering these services is financially beneficial. Wellander, Wells, and Feldman (2016) conducted a cost-offset analysis of two evidence-based preventive interventions offered to children in the school setting. Findings from their study demonstrated that the reduction in mental health problems led to cost savings over the course of the school year and that there would be a return on investment 1.5 years after implementing the prevention interventions.

Family Influences on Behavioral Health. While individual risk factors, such as genetics or temperament, may be linked to the development of behavioral health disorders, family and level of parental functioning can also play a role in the development of mental health conditions (Fossum et al., 2007). In terms of influencing the development of children's mental health disorders, it is well established that there is a relationship between caregiver or parent mental health and child mental health. Specifically, children who have parents with a mental health condition are at an increased risk of experiencing a mental health condition themselves (Lindsey et al., 2008). Studies have consistently found that children with parents who have a mental health disorder experience greater challenges than children with parents without a mental health disorder (Beardslee, Versage, & Gladstone, 1998; Lindsey et al., 2008). One study found that children of parents experiencing depression are four times more likely to develop a mental health condition relative to children with parents who do not experience depression (Lavoie & Hodgins, 1994). Other research has also shown that the best predictor of child well-being is the functioning of their caregiver (Lester et al., 2010).

Children whose parents experience high stress levels are also at an increased risk for difficulties (English, Marshall, & Stewart, 2003). A recent study found that parent strain predicted improvement in children's mental health symptoms. Specifically, high levels of parental strain were associated with less improvement in mental health symptoms over time (Accurso et al., 2015). In addition, poverty has been consistently linked with mental health difficulties, and children who reside in low-income, urban communities who are exposed to community violence and psychosocial stress are especially likely to experience a behavioral health disorder (Gopalan et al., 2010). The link between poverty and mental health difficulties can often be attributed to the high

levels of stress experienced by individuals and families living in poverty (Santiago, Kaltman, & Miranda, 2013). Finances and employment are cited as the most common reasons for stress, and financial troubles, unemployment, and other negative life events are all risk factors for developing a condition such as depression (American Psychological Association, 2016; Bonde, 2008). A 2014 review of 181 studies also found that low levels of parental monitoring, low levels of youth autonomy, and low levels of parental warmth were associated with depression in youth (Yap, Pilkington, Ryan, & Jorm, 2014).

On the contrary, supportive family relationships during childhood are associated with long-term psychosocial functioning (Paradis et al., 2011). Parent social support has also been linked with child mental health outcomes (Bussing et al., 2003). Hoagwood (2005) found that when parents reported having social support, their children had better outcomes on numerous child health outcomes. Previous research has found that positive family relationships are linked to a range of positive outcomes, including increased self-esteem and academic achievement, as well as reduction in risk of negative outcomes such as poor mental health and physical health (Paradis et al., 2011; Milevsky, 2005; Shaw et al., 2004). Paradis and colleagues (2011) found that having a family member to confide in as a child reduced the risk of mental health concerns at age 30. Work by Smokowski and colleagues (2014) also found that positive parenting—defined by parent support, parent-child future orientation, and parent education support was linked with lower levels of depression, increased levels of self-esteem, and optimism about the future. Past work has defined positive parenting as, “an umbrella term used to refer to dimensions of parenting such as warmth, sensitivity, limit setting, appropriate scaffolding, and contingency-based reinforcement” (Waller et al., 2015).

Parenting is also important in the development of children's mental health disorders. While parenting can involve challenges for all parents, parenting a child with a mental health condition can involve additional challenges and stress. Parenting practices, such as negative parent-child interactions and harsh discipline techniques, have been linked with depression and self-esteem problems in youth (Smokowski et al., 2014). Families with children who suffer from mental health conditions often have unstable and ineffective parent-child interactions in comparison to children without these conditions (George, Herman & Ostrander, 2006). High levels of family conflict and low levels of family cohesion often create coercive family environments (George, Herman & Ostrander, 2006). Parent-child conflict is related to increased anxiety, depression, aggression, and behavior problems in children (Smokowski et al., 2014). A systematic review by Wood and colleagues (2003) reported that controlling parental behavior during parent-child interactions was consistently linked with behavioral health problems in children.

Child Adversity and Behavioral Health. The relationship between childhood adversity and adult illness is well documented in prior research and childhood adversity has been linked to various adult behavioral health conditions (Curran et al., 2016). In the United States, an estimated two out of three children will experience a traumatic event before turning 16 (Copeland, Keeler, Angold, & Costello, 2007). A 2010 study by Green and colleagues found that 53% of individuals had experienced at least one adverse childhood event. The Adverse Childhood Experiences Survey (ACES) studies have also demonstrated the long term, and often negative outcomes associated with experiencing traumatic events. Early life stress, which includes emotional abuse, physical abuse, physical assault, sexual abuse, emotional neglect, physical neglect, as well as things such as parental loss or childhood illness, have also been associated with development

and mental health conditions as an adult (Bernstein et al., 2003; Carr et al., 2013). Furthermore, emotional abuse, physical abuse, and neglect are also associated with increased rates of depression (Norman et al., 2012). When trauma is experienced during a developmental period, it is possible to yield negative effects that remain with the individual throughout their life, and early life stress can trigger behavioral health conditions in adulthood (Carr et al., 2013). While exposure to trauma and experiencing early life stress is a risk factor for later mental health conditions, it is not the sole factor, and individual vulnerability also needs to be considered (Carr et al., 2013).

Since children rely on their caregiver to help them make sense of difficult and traumatic experiences, parent support and family functioning are critical factors in how children respond to exposure to trauma (Lieberman & Knorr, 2007). Porche, Costello, and Rosen-Reynoso (2016) found that children with higher numbers of adverse family experiences were more likely to have a mental health condition and that poor caregiver mental health was positively associated with child mental health diagnoses.

Challenges in family functioning are linked with childhood adversity, which is predictive of the development of mental health difficulties (Green et al., 2010). Green et al. (2010) defined childhood adversity by 12 dichotomized events and classified the childhood adversities into four categories including interpersonal loss (e.g., parent divorce), family maladaptive experiences (e.g., family violence), abuse and neglect (e.g., physical abuse), and other (e.g., economic adversity). Maladaptive family functioning was defined by parental mental illness, parental substance abuse, parental criminality, family violence, and abuse and neglect, and in the study, were most strongly correlated with the onset of a mental health condition. Research suggests that when trauma is experienced in the family home, and if it is caused by an attachment figure, the risk of mental health difficulties is increased (Carlson & Dalenburg, 2000). A study by Engstrom

and colleagues (2012) found that women in methadone treatment who experienced trauma, in the form of childhood sexual abuse, were at an increased risk for mental health difficulties and posttraumatic stress disorder (PTSD) when the childhood sexual abuse involved a family member.

Protective Factors. In their book *Children of Katrina*, Fothergill and Peek (2015) report on a longitudinal, ethnographic study of the impact Hurricane Katrina on children's health and well-being, family life, education, living circumstances, and relationships. Specifically, they present case studies of children and their developmental trajectories post-disaster and explain personal and structural factors that differentiate children who remained on their developmental trajectories after Katrina from children whose life trajectories declined after Katrina. Ultimately, Fothergill and Peek (2015) explain that financial, social, cultural, and educational resources that were available to children both pre-and-post disaster were critical in determining how children would respond to the experience. Additionally, children whose families could mobilize resources, draw upon institutions for help when needed, and access supportive adults were able to establish equilibrium in their lives and experience positive outcomes after the trauma of Katrina.

When children experience adversity, supportive and available adults are important for children to overcome these challenges and cope with stress (Easterbrooks, Ginsberg, & Lerner, 2013). Children who demonstrate high levels of mental health functioning despite their exposure to traumatic experiences are often referred to as being resilient (Masten & Narayan, 2012). Resilience can be defined as the ability to function typically or even thrive after experiencing severe trauma or adverse living conditions (Masten, 2007; Werner & Smith, 1982). Ecological models that consider a child's individual characteristics, their family, and environment have been used to help understand what contributes to resilience (Diab et al., 2015). A primary way to support

children's positive behavioral health development is to help strengthen and promote resilience by drawing upon protective factors.

While it is well-documented that exposure to childhood adversity is associated with negative health outcomes, less is known about protective factors that promote health and well-being despite experiencing adversity (Banyard, Hamby, & Grych, 2017). Parenting self-efficacy, daily parent-child interaction, parent relationship satisfaction with their partner, parent's having good quality social relationships, social support, and frequent exercise have been found to be protective factors for children's healthy development (Collishaw et al., 2016; McDonald et al., 2016). A 2014 review identified three protective factors for depression related to parenting including parental warmth, autonomy, and monitoring (Yap, Pilkington, Ryan, & Jorm, 2014).

Additionally, it is a protective factor when families have an understanding about mental health and communicate about mental health challenges (Beardslee, Gladstone, Wright, and Cooper, 2003; Greeff, Vansteenweggen, & Ide, 2006). Children's behavior and their emotional responses to adverse experiences have been linked with maternal warmth, defined by supportiveness, acceptance, and having a positive affect (Kim-Cohen, Moffitt, Caspi, & Taylor, 2004). Family stability, supportive relationships, and community engagement are also factors that can protect children and adolescents from developing behavioral health problems (Substance Abuse and Mental Health Services Administration, 2011).

A recent study by Banyard, Hamby, and Grych, (2017) examined protective factors associated with health for youth who had been exposed to high levels of adversity. Emotional regulation, being able to make meaning out of situations, practicing forgiveness in relationships, and social support at both the community and friend level were all factors related to positive health (Banyard, Hamby, & Grych, 2017). Happiness

is also related to the prevention of mental health problems, and on average, happier people experience fewer mental health problems (Diener & Seligman, 2002; Layous, Chancellor, & Lyumbomirsky, 2014).

Layous, Chancellor, and Lyumbomirsky (2014) propose ways positive activities or happiness-increasing strategies can serve as protective factors against the development of mental health conditions. “Positive activities are typically brief, simple, accessible and require little or no financial resources” and can include activities such as acts of kindness, thinking optimistically, focusing on strengths, and being grateful for what we have in life (Layous, Chancellor, & Lyumbomirsky, 2014, p. 5). Layous, Chancellor, and Lyumbomirsky (2014) state that positive activities can act as proximal protective factors that explain why individuals with similar risk factors go on to have different trajectories. Specifically, they suggest that positive activities can directly decrease proximal risk factors for behavioral health conditions, reduce the likelihood that early risk factors lead to proximal risk factors, and decrease conditions that interact with risk factors that lead to the development of conditions (Layous, Chancellor, & Lyumbomirsky, 2014). Additionally, they suggest that positive activities can facilitate adaptive coping when faced with negative or stressful experiences.

Identification of Behavioral Health Disorders. The first step to addressing behavioral health conditions is to recognize that an individual is struggling with a condition. However, identifying individuals with a behavioral health conditions can be challenging. Identifying behavioral health problems early may decrease long-term disability (Williams, Klinepeter, Palmes, Pulley, & Foy, 2004), and early identification is key to prevention. The benefits of intervening early when someone is struggling with mental health symptoms are well-established (Ginsburg et al., 2014; Wolk, Kendall, & Beidas, 2015). Research suggests that most adult behavioral health disorders begin in

childhood and that half of all mental health conditions begin by age 14 (Kessler et al., 2005; Kessler et al., 2007). Symptoms of mental health conditions are often present two to four years before developing into a condition that meets diagnostic criteria, and, on average, most conditions are not diagnosed until 10 years after the first symptoms (U.S. Department of Health & Human Services, 1999).

Children's caregivers are usually the first to recognize problems their children are experiencing, but studies have shown that caregivers often have difficulty identifying mental health conditions, such as depression, because symptoms may not be completely apparent or may be consciously concealed by the youth (Levitt et al., 2007; Logan & King, 2002). Behavioral health conditions are also under-identified in primary care settings, and primary care physicians (PCPs) have been shown to only identify one in five children who need mental health services (Pidano et al., 2011; Sayal, 2006; Simonian, 2006). Furthermore, it is known that individuals who die by suicide, a consequence of depression, often have a recent visit with a PCP before their death (Indu et al., 2017). A study by Ozer and colleagues (2009) which assessed primary care physician's rates of talking about mental health concerns with teen patients found that only about one-third (34%) of youth reported that their doctors asked them about their mental health.

PCPs encounter numerous barriers to screening for behavioral health conditions despite their interest in it. Specifically, lack of time, referral resources, reimbursement constraints, and lack of training and comfort with adequately addressing behavioral health concerns are often cited as reasons patients are not screened in primary care settings (Badger, Robinson, & Farley, 1999; Hogan, 2003; Murphy et al., 1996; Samet, Friedman, & Saitz, 2001; Trude & Stoddard, 2003).

Engagement in Behavioral Health Services. Despite the number of children in need of behavioral health services, engaging families in treatment services is challenging, and many obstacles exist for families. A significant gap exists between the number of children who need treatment and children who have access to treatment. Some research suggests that only 25% of children with mental health problems receive treatment (Hoagwood, & Olin, 2002).

Engagement is often separated into initial attendance in services and retention or ongoing engagement in services (McKay, Stoewe, McCadam, & Gonzales, 1998). McKay and Bannon (2004) define initial engagement as the phase that begins with the identification of a child's mental health problem until the child is brought to a clinic for services. Attendance after the first contact and attendance over the course of treatment are typical measures of ongoing engagement in child mental health services. However, in the literature, engagement has not been consistently defined, which makes the interpretation of the services research difficult (Becker et al. 2015).

Lindsey et al. (2014) highlight that adherence is potentially another important outcome to consider when measuring engagement. As cited in Nock and Ferriter (2005), adherence is defined as "active, voluntary, collaborative involvement of the patient in a mutually acceptable course of behavior to produce a desired preventative or therapeutic result" (Meichenbaum & Turk, 1987, p. 20). In the child mental health context, treatment adherence is often measured by the quantity and quality of therapeutic actions completed by the parent in the therapy session (e.g., participation in role plays) or between treatment sessions (e.g., homework; Nock & Ferriter, 2005).

Research suggests that as many as 50% of children in need of mental health services never receive treatment (Merikangas et al., 2010). Parents and families directly influence whether a child receives treatment (Raviv, Sharit, Raviv, & Rosenblat-Stein,

2009). Engaging parents in their child's treatment is critical given the role they play in facilitating treatment attendance (Haine-Schlagel & Walsh, 2015).

A variety of factors at the individual, family, environmental, and service system are associated with treatment engagement. Research suggests that a parent's decision to seek services for their child is related to past service experience, beliefs about their child's problem, and perceived barriers (Kerkorian, McKay, & Bannon, 2006). Kerkorian and colleagues (2006) also found that parents who felt disrespected by their child's therapist were six times more likely to doubt the usefulness of future treatment.

Prior research has examined child and family characteristics associated with engagement in child mental health services primarily using administrative data (McKay & Bannon, 2004). Families who experience poverty and identify as racial minorities often under-use treatment services (Hoberman, 1992; Kazdin, Holland, & Crowley, 1997). Demographic and clinical characteristics such as child's age, gender, and level of mental health impairment have also been considered, although the directionality of relationships between these characteristics and service engagement are often unclear (McKay & Bannon, 2004). The research findings regarding the severity of a child's mental health condition and service use are mixed. Buckner and Bassuk's study (1997) found that among children whose family was experiencing poverty, and had severe mental health symptoms, the less likely the children were to receive mental health treatment. A study by Dore, Wilkinson, and Sonis (1992), however, found a positive relationship between mental health severity and service use.

Harrison, McKay and Bannon's (2004) study examined factors associated with child mental health service use and reasons families chose not to use services once making the initial appointment. All families in the study had identified a mental health concern for their child and made an appointment at the clinic. However, 32% of families

did attend treatment services after scheduling the appointment. Parental discipline efficacy and social support were associated with service use for families. For families who made the appointment, but did not attend services, the most common reasons for non-attendance were that the therapist did not call and that the parent was too busy or overwhelmed.

Not only is getting families linked with treatment challenging, keeping them engaged in treatment can also be a challenge. Of families who enroll in treatment, more than 50% of families discontinue treatment early (Kazdin & Mazurick, 1994; Pellerin, Costa, Weems, & Dalton, 2010). One study found that more than two-thirds of families discontinue treatment by the seventh session in community-based settings (Miller, Southam-Gerow, & Allin, 2008). Similarly, McKay, Lynn, and Bannon (2005) found that only 9% of children referred for mental health services were still receiving services at 12 weeks. In poverty-impacted communities, average length of treatment is between three and four sessions (McKay, Harrison, Gonzales, Kim, & Quintana, 2002). Prior research suggests that participation in at least eight sessions is associated with more positive treatment outcomes for children and families (Koegl, Farrington, Augimeri, & Day, 2008).

Family stress, perception of treatment need, the relationship between the parent and therapist, economic disadvantage, single parent status, and identifying as a member of a racial minority group are associated with families ending treatment prematurely (Angold et al., 2002; Armbruster & Fallon, 1994; Kazdin, Holland, & Crowley, 1997). Recent work by Kim and colleagues (2015) also suggests organizational culture and therapist characteristics influenced service engagement. Specifically, high therapist caseloads were associated with low engagement while professional support and trust were associated with higher engagement (Kim et al., 2015). Families who have collaborative relationships with their therapist are also more likely to remain engaged in

treatment (Thompson et al., 2007). Johnson, Mellor, and Brann (2008) found that families with high levels of psychosocial difficulties were the most likely to end services prematurely. In addition to these factors, quality of care can also be considered a barrier to treatment engagement. There is evidence that a substantial number of individuals in mental health treatment do not receive minimally adequate care (Katz, Kessler, Lin, & Wells, 1998; Wang, Berglund, & Kessler, 2000; Wang, Demler, & Kessler, 2002). One study found that of participants currently receiving treatment, only 39% received care that was considered adequate (Wang, Demler, & Kessler, 2002).

One strategy that policymakers have used to address service quality is the implementation of evidence based practices (EBPs), which are practices that have been shown to be effective through research (Stagman & Cooper, 2010). However, many barriers exist in the adoption of these EBPs such as low fidelity, infrastructure problems, provider concerns, and the ability to implement the EBPs in their setting (Cooper & Aratani, 2009; Schoenwald et al., 2008; Tanenbaum, 2005). In addition to these engagement and quality of care challenges, children and families often experience numerous challenges accessing mental health treatment.

Barriers to Treatment. Researchers often divide barriers to treatment into two broad categories: structural/practical and perceptual. Structural/practical barriers include things like child care problems, transportation issues, not having insurance, quality of services, and lack of time. Clinic hours and wait-lists are also other examples of structural barriers that have been associated with families not accessing services (Harrison, McKay, & Bannon, 2004). Perceptual barriers include parents' beliefs about their child's need for treatment, stigma related to seeking help, and prior negative experiences with mental health providers (Owens et al., 2002). Owens et al. (2002) proposed separating out the perceptual barriers into barriers related to perceptions

about mental health problems, and barriers related to perceptions about mental health services.

Research suggests that although practical/structural barriers may interfere with families' ability to access treatment, families will often work around these barriers as long as the treatment is perceived by the parent as matching their preference (Bannon & McKay, 2005). Kazdin and colleagues have published a series of articles that examined perceived barriers to treatment experienced by families (Kazdin, Holland, & Crowley, 1997; Kazdin & Wassell, 1999; Kazdin & Wassell, 2000; Kazdin, 2000). While logistical/practical barriers were linked to treatment engagement, perceptual barriers were much better at predicting engagement (Kazdin, Holland, & Crowley, 1997). The role of stigma and how discrimination impacts service engagement have also been examined in a recent study (Clement et al., 2015). Clement and team (2015) found that experiencing discrimination because of mental health difficulties was related to low service engagement, and this relationship was mediated by mistrust in mental health services and the therapeutic relationship. Parent participation is especially critical for child and family services given the critical role parents play in ensuring their child attends treatment services (Haine-Schlagel & Walsh, 2015).

Evidence suggests that children with behavioral problems, residing in poverty-impacted communities with their families, often experience multiple multilevel stressors and barriers that impact their participation and engagement in mental health services (Franco, Pottick, & Huang, 2010). A recent study found that higher levels of parental stress were associated with lower rates of treatment attendance (Jackson, 2015).

Health Care Reform and Population Health. An estimated 63 million Americans gained health insurance, including coverage for behavioral health care, and access to services through the Affordable Care Act (ACA) and the Mental Health Parity

and Addictions Equity Act (Beronio, Glied, & Frank, 2014; Frank, Beronio, & Glied, 2014). Health care reform in the United States has led to three goals coined “The Triple Aim” which is to improve health, improve care, and reduce costs (Berwick, Nolan, & Whittington, 2008).

In addition to expanding health care coverage, the Affordable Care Act also encouraged the shift from Fee-For-Service (FFS) payments to value-based payment models. As part of this shift, there is an interest in identifying populations with high-need from a provider and payer perspective, which facilitates the inclusion of behavioral health care into the health care conversation (Rawal & McCabe, 2016). To be successful under these new payment models, population health strategies that focus not only on efficient and effective care, but also on the promotion of health and well-being, are now being considered (Rawal & McCabe, 2016).

Population health is defined as “the health outcomes of a group of individuals, including the distribution of such outcomes within the group” (Kindig & Stoddart, 2003). Population health initiatives seek to improve the health of populations by focusing on prevention and wellness instead of illness (Frieden, 2010). Behavioral health is a critical component to population health. To achieve the best possible outcomes for patients, the whole person, which includes both physical health and behavioral health, needs to be addressed.

Despite the progress that has been made, the U.S. healthcare system remains fragmented, inefficient, highly-regulated, and expensive. Technology and digital health solutions have addressed some of the most challenging aspects of health care. Technology that can predict poor patient outcomes and inform prevention approaches is one strategy that could help our health care system (Amarasingham et al., 2014).

Predictive analytics for clinical decision making have been acknowledged as one approach to achieve the Triple Aim (Amarasingham et al., 2014).

Statistical/Machine Learning. Predictive analytics includes a broad set of statistical tools that identify trends, relationships, and patterns within data that can be used to predict a future event or behavior (Eckerson, 2007). The approaches and techniques used to conduct predictive analytics can broadly be grouped into regression techniques and machine learning techniques (Eckerson, 2007). Murphy (2012) defines machine learning as “a set of methods that can automatically detect patterns in data, and then use the uncovered patterns to predict future data, or to perform other kinds of decision making under uncertainty” (p. 1). Machine learning techniques are typically grouped into two main types: 1) predictive or supervised learning approach, where the goal is to make predictions based on historical data, and 2) descriptive or unsupervised learning approach, where the goal is knowledge discovery (Murphy, 2012).

While traditional regression methods are useful if the primary goal is to understand the relationship between the predictors and dependent variable, if the primary goal is to make decisions based on the data, statistical/machine learning methods often outperform traditional regression procedures (Berk, 2016). Unlike conventional regression methods, which focus on model building, statistical/machine learning uses an algorithmic method. Berk (2016) offers a baking metaphor to describe the “black-box” nature of algorithms and explains how just like when baking bread, the baker knows and can modify the ingredients based on preference, but the baker does not know much about the physics and chemistry that go on in the oven to transform the ingredients into bread. The goal of the black-box procedures is accurate forecasting; understanding the relationship between inputs and outputs are secondary and often unknowable (Berk, Sorenson, & Barnes, 2016; Berk & Bleich, 2013)

In the private sector, companies have been successful at using big data and machine learning strategies to improve their efficiencies and operations (Davenport & Harris, 2017). The financial service and insurance industries, as well as companies like Netflix and Amazon, heavily rely on machine learning techniques to make predictions and offer recommendations. Also, more recently, machine learning techniques have begun being used in the public sector in areas such as physical health, child welfare, domestic violence, and criminal justice. Machine learning approaches are becoming more common as they often outperform classical regression techniques in dealing with prediction and classification decisions (Orrù et al., 2012; Singal et al., 2013; Yoo, Ference, Cote, & Schwartz, 2012).

In physical health care, machine learning has been used to identify patients at risk of being unnecessarily readmitted to the hospital for conditions such as heart failure and pneumonia (Shams, Ajourlou, & Yang, 2014). Billings and colleagues (2006, 2012) have also used similar strategies to identify patients at high risk for hospital readmissions. In the child welfare setting, predictive analytics have primarily been used to predict entry into and time spent in foster care (Russell, 2015) and to identify children at risk of maltreatment so that supportive services can be targeted for prevention (Gillingham, 2015). A recent study by Berk, Sorenson, and Barnes (2016) applied machine learning to forecast future domestic violence incidents that could be used to inform arraignment decisions. The use of algorithms and development of forecasting procedures have also been used in criminal justice settings (Berk & Hyatt, 2014). For example, these methods have been used to provide decision support around supervision and service needs for individuals on parole or probation (Berk, Sherman, Barnes, Kurtz, & Ahlman, 2009), to help prison officials assign incarcerated individuals to the

appropriate security level (Berk, Kriegler, & Baek, 2006), and to inform judges' sentencing decisions (Berk & Bleich, 2014).

In the context of behavioral health, application of machine learning strategies is still relatively new and underutilized. Passos et al. (2016) state that "machine learning in psychiatric research is an emerging field with great potential for innovation and paradigm shift as these algorithms facilitate integration of multiple measurements as well as allows objective predictions of previously 'unseen' observations" (p. 110). Recent studies have applied machine learning to identify behavioral markers as predictors of cocaine dependence (Ahn, Ramesh, Moeller, & Vassileva, 2016), to identify which EEG features distinguish healthy individuals from individuals with schizophrenia (Johannesen et al., 2016), and to predict anorexia nervosa (Guo, Wei, & Keating, 2016). Passos and colleagues (2016) conducted a study that investigated the use of machine learning algorithms to identify individuals with mood disorders who were at risk for suicide. Another study used machine learning algorithms to predict individuals who would develop major depressive disorder using earlier self-report data (Kessler et al., 2016). Furthermore, Galatzer-Levy and team (2014) forecasted individuals who would develop post-traumatic stress disorder (PTSD) after experiencing a traumatic event, and Carpenter et al. (2016) used machine learning to predict risk scores for anxiety disorders in children.

Despite advances in knowledge about behavioral health conditions, the translation of original research to practice takes at least ten to twenty years (Fishbein, Ridenour, Stahl, & Sussman, 2016). However, machine learning forecasts can be generated within real time (Berk, Sorenson, & Barnes, 2016) and offer tremendous opportunity to identify individuals who are at risk of developing a behavioral health condition and connecting them with prevention or treatment services to mitigate this risk.

Despite the tremendous opportunity that these machine learning algorithms offer, criticism and ethical considerations should be noted. First, there is some concern about using data for purposes other than the original purpose the data were collected (Culhane, 2016). Critics have also had concerns about the care-free view of big data and predictive analytics (Marcus & Davis, 2014). There is also some concern about bias that may exist in the data and algorithms. For example, U.S. Attorney General Eric Holder's criticism is that even when factors such as race are not included directly in the algorithm, race is embedded within other variables, which calls into question the neutrality of these algorithms (Barry-Jester et al., 2015). Bone and colleagues (2015) also highlight that if machine learning techniques are used in the absence of clinical or content expertise, it could lead to misinformed and inaccurate conclusions. Additionally, some have expressed feelings of general discomfort with the idea of allowing machines and computers to make decisions about human behavior. While these concerns about the use of technology and the role of potential bias in decision making exist, in standard criminal justice and child welfare practice, decisions about humans are often made by decision-makers who use their discretion and have their own implicit biases (Berk & Hyatt, 2014; Gillingham, 2015).

Present Study

Behavioral health disorders, specifically depression, are a serious problem in the United States and worldwide. The consequences of unaddressed behavioral health conditions are multifaceted and have impact at the individual, relational, communal, and societal level. Despite the number of individuals who could benefit from treatment for behavioral health difficulties, their difficulties are often unidentified and unaddressed through treatment. Reducing psychosocial risk factors and building protective factors can

prevent the development of behavioral health conditions, and support children's healthy development.

Through the Affordable Care Act and health care reform efforts, there has been a growing interest in being able to identify populations with high need. Technology that can predict patient outcomes and inform prevention approaches is one strategy that could help our health care system (Amarasingham et al., 2014). This study is one of the first to use machine learning strategies to inform the prevention of mental health difficulties, specifically depression. This study provides important information about the feasibility and practicality of using a forecasting tool for preventing depression in adulthood and acts a demonstration/proof of concept of how a forecasting tool could be used in a real-world setting to identify youth who are at risk of having a depression diagnosis as an adult and to inform early intervention strategies.

While we know that exposure to childhood adversity is associated with negative health outcomes, less is known about the protective factors that promote good health and well-being despite experiencing adversity (Banyard, Hamby, & Grych, 2017). This study addresses this gap by examining factors associated with the development of depression and contributes to the field's understanding of children's developmental trajectories.

This study has two aims: 1) to develop a forecasting procedure that can be used to identify youth who are at risk of reporting a depression diagnosis as adults based on a set of input variables; and 2) to understand the developmental trajectories of depression for youth. Drawing upon representative data from youth and young adults in the United States, this study addresses the following research questions:

1. How well do random forests forecasts perform in terms of predicting which youth will report a depression diagnoses as an adult?

2. What features distinguish youth with depressive symptoms who report a depression diagnosis as an adult from youth with depressive symptoms who do not report a depression diagnosis as an adult?

CHAPTER 2: METHOD

Research Design

This study is a secondary data analysis of publicly available data from Wave I and Wave IV of the National Longitudinal Study of Adolescent (Add Health). Add Health is a longitudinal study of a nationally representative sample of youth in the United States (Chen & Chantala, 2014). Data were collected at four different waves over the course of 14 years. Wave I was conducted during the 1994-1995 academic year when youth were enrolled in grades 7 through 12. The Wave IV study was a follow-up of youth from Wave I and was completed in 2008 when participants were 24 to 32 years old (17 participants in the public use sample were 33-34 during the interview). The goal of the Add Health study was to collect data on the health of American youth to help explain health and health behaviors as they transition to adulthood while accounting for multiple contexts of life (Harris, et al., 2009; Harris, 2013). The Add Health study was approved by the Institutional Review Board (IRB) at the University of North Carolina School of Public Health and is in compliance with federal regulations on the protection of human subjects (Chen, Corvo, Lee, & Hahm, 2017). Given that data from this study are de-identified and publicly available in nature, verification was received from the Institutional Review Board (IRB) at the University of Pennsylvania that IRB approval was not required for this study.

As described by Chen and Chatala (2014), the Add Health study used a school-based research design and their primary sampling frame came from the Quality Education Database (QED), which is made up of 26,666 high schools across the United States. A stratified sample of 80 high schools (defined as schools with more than 30 students and an 11th grade) was chosen. High schools were then stratified by school type (public, private, parochial), region, race/ethnicity of students, urbanicity, and size.

For every high school that was selected, a feeder school, usually a middle school that had a high proportion of its students attend the selected high school, was also recruited, totaling one school pair in 80 different communities. Some schools had grades 7 through 12, to comprise a total of 132 schools in the sample. Overall, 79% of schools contacted agreed to participate in the Add Health study.

Youth were selected from the identified schools using unequal probability sampling methods. Youth were stratified by grade and gender and then 17 students were randomly selected from each stratum for about 200 youth from each school pair. Supplemental samples based on ethnicity, genetic relatedness to siblings, adoption status, and disability were also drawn. African American/Black youth with highly educated parents were also oversampled for this study.

Data and Sample

This study used the Add Health public-use dataset, which is a random subset of half of the core sample and half of the oversample of African American/Black youth who have a parent with a college degree, totaling about one-third of the full sample. These data are available at no-cost from the Inter-University Consortium for Political and Social Research (ICPSR) and can be downloaded from the Data Sharing for Demographic Research (DSDR) website. The predictor or input variables for this study came from Wave I and the response or outcome variables came from Wave IV. At Wave I, data were collected from schools, youth, and a parent, and at Wave IV, data were collected from the original youth. Total number of respondents included in Wave IV of the public-use data is 5,114. Since this study drew upon responses from both youth and parent questions only those respondents with both youth and parent interviews at Wave I were included in this study (N=4,489)

A total of 92.5% of respondents from Wave I were located at Wave IV and 80.3% were interviewed (Harris, 2013). This response rate exceeds other national longitudinal studies, as their response rates typically range from 55% to 77% (Harris, 2013). Attrition in the study varied by gender, race, and immigrant status as women, individuals who were White, and native born individuals had higher response rates at Wave IV. Response rates were also higher for respondents with increased socioeconomic resources and parental education at Wave I. The effect of non-responsiveness was examined by using demographic, behavioral, health, and attitudinal variables from Wave I to see if any bias was introduced at Wave IV because of differences between people who did and did not respond (Harris, 2013). Results indicated that non-response bias was negligible and it was concluded that the Wave IV sample adequately represented the population interviewed at Wave I (Harris, 2013).

Measures

For this study, the outcome being forecasted is a reported depression diagnosis as an adult. This binary outcome variable was constructed from Wave IV of the Add Health data. Specifically, a new variable was computed. Participants were asked “Has a doctor, nurse or other health care provider ever told you that you have or had: depression?” and the respondents could answer yes or no. For this study, if a participant responded yes to the question, they were coded with a one (DepressionDiagnosis1) to indicate presence of a reported depression diagnosis or if they responded no they were coded with a zero (DepressionDiagnosis0), to indicate absence of a reported depression diagnosis.

While typically machine learning algorithms are atheoretical, Berk et al. (2009) highlight that for these algorithms to be accepted in practice, the inputs or predictor variables need to make conceptual sense. Therefore, for this study, the input variables are ones that could in practice be obtained from youth and could be justified empirically. Input variables include youth and parent demographic characteristics, symptoms associated with depression, health service utilization, and various risk and protective factors related to developing a behavioral health condition. Table 1 provides the inputs that were used for the forecasting procedure. Most of the inputs are binary responses (yes=1, no=0) and self-explanatory by their name. Input variables that were constructed are described below and denoted in the table with an asterisk. All variables and measures came from Wave I of the Add Health study. Variables are reported by the youth except for caregiver demographics and family financial information, which are specified.

Table 1: Inputs for Forecasting

Female
Asian
Black/African American
Hispanic/Latino/a
White
Good, very good, or excellent health
Is happy
Ever missed a social or recreational activity because of health or emotional problem
Ever missed school because of health or emotional problem
Poor appetite
Trouble falling or staying asleep
Trouble relaxing
Moodiness
Frequent crying
Feeling fearful
Ever received counseling
Received yearly physical examination
Learned about where to go for help with a health problem
Learned about suicide in school
Learned about stress in school
Repeated a grade in school
Received an out-of-school suspension
Expelled from school

Depression score on CES-D*
 Significant symptoms of depression (CES-D score 10 or higher)
 Maternal warmth
 Family connection/support*
 Times a week have dinner with least one of your parents in the same room
 Self-esteem*
 Suicidal ideation
 Ever attempt suicide
 Have friends that have tried to kill themselves
 Have friends who died by suicide
 Have family members that have tried to kill themselves
 Have family members who have died by suicide
 Adult social support*
 Have friends that care
 Participate in sports
 Exercise in the past week
 Maternal attachment*
 Maternal involvement*
 Autonomy from parents*
 Primary caregiver is married
 Primary caregiver has a college degree or higher
 Primary caregiver is employed
 Primary caregiver is happy
 Primary caregiver in good or excellent health
 Primary caregiver talks about school with youth
 Primary caregiver talks about grades with youth
 Family has trouble paying bills
 Family receives public assistance
 Family receives food stamps
 Family receives Aid to Families with Dependent Children (AFDC)
 Saw someone shoot or stab another person
 Had someone pull a knife or gun on them
 Been shot or stabbed
 Been cut or stabbed
 Been jumped
 Neighborhood connection*

Self-Esteem. Self-esteem was measured using six-items from the Rosenberg Self-Esteem Scale (Rosenberg, 1989). Participants were asked to indicate the level to which they agreed or disagreed with each statement such as, “You have a lot of good qualities” and “You like yourself just the way you are.” Responses were coded on a five-point scale ranging from *strongly agree (1)* to *strongly disagree (5)* and the six items were summed for a total score that could range from 6-30. Lower scores indicate higher levels of self-esteem (present study $\alpha=.85$). This measure has been used in previous

studies using the Add Health data with alpha values demonstrating high internal reliability (Driscoll, Russell, & Crocket, 2008; Shahar & Henrich, 2010).

Family Connection/Support. Family support was measured by creating a four-item scale using the questions: “How much do you feel that people in your family understand you?”, “How much do you want to leave home?”, “How much do you feel that you and your family have fun together?”, and “How much do you feel your family pays attention to you?” Other research using Add Health data have used a similar scale to measure family support (Zhu, 2018). Responses were coded on a five-point scale ranging from *not at all* (1) to *very much* (5). The item about leaving home was reverse coded so that a higher rating reflected the youth indicating not wanting to leave home. The four items were summed for a total score that could range from 4-20 with higher scores indicating higher levels of family support (present study $\alpha=.76$). The alpha from a previous study with the Add Health data had a similar value of .75 (Zhu, 2018).

Maternal Attachment. Maternal attachment was measured using a two-item scale that has been used in prior research using the Add Health data (Beaver et al., 2015, Schreck, Fisher, & Miller, 2004; Wright, Beaver, Delisi, & Vaughn, 2008). Participants were asked, “How close do you feel to your mother?” and “How much do you think that she cares about you?” Responses were coded on a five-point scale ranging from *not at all* (1) to *very much* (5). The two items were summed for a total score that could range from 2-10 with higher scores indicating higher levels of maternal attachment (present study $\alpha=.64$). Previous studies using this scale have had alpha values ranging from .64 to .70 (Beaver et al., 2015, Schreck, Fisher, & Miller, 2004; Wright, Beaver, Delisi, & Vaughn, 2008).

Maternal Involvement. Maternal involvement was measured using a ten-item scale that has been used in prior research using the Add Health data (Beaver et al., 2015; Beaver, 2008; Cheng & Lo, 2017; Crosnoe & Elder, 2004). Participants were asked to indicate which activities such as “Gone shopping” or “Had a talk about a personal problem you were having” had occurred with their mothers in the past month. Responses were coded *yes (1) and no (0)* and summed for a total score that could range from 0-10, with higher scores indicating higher levels of maternal involvement (present study $\alpha=.53$). Previous studies using this scale have had alpha values ranging from .52 to .61 (Beaver et al., 2015; Beaver, 2008; Cheng, Tyrone, & Lo, 2017; Crosnoe & Elder, 2004).

Autonomy from Parents. Youth autonomy from their parents was measured using a seven-item scale where participants were asked questions about if their parents allowed them to make their own decisions such as, “Do your parents let you make your own decisions about what time you go to bed on a weeknight?” and “Do your parents let you make your own decisions about the people you hang around with?” This measure has been used in prior studies with Add Health data, and alpha values ranged from .57 to .64 (Barnes & Morris, 2012; Beaver et al., 2015; Cheng, Tyrone, & Lo, 2017; Wright, Beaver, Delisi, & Vaughn, 2008). Responses were coded *yes (1) and no (0)* and summed for a total score that could range from 0-7, with higher scores indicating higher levels of autonomy (present study $\alpha=.61$).

Neighborhood Connection. Neighborhood connection was measured by creating a three-item scale with the questions: “You know most of the people in your neighborhood”, “In the past month, you have stopped on the street to talk with someone

who lives in your neighborhood”, and “People in this neighborhood look out for each other.” Responses were coded *true (1) and false (2)*, and the three items were summed for a total score that could range from 3-6, with lower scores indicating higher levels of neighborhood connection (present study $\alpha=.57$). Previous research with the Add Health data has measured neighborhood connection in a similar way with alphas ranging from .55 to .63 (Bazaco et al., 2016; Cheng & Lo, 2017; Li et al., 2018).

Adult Social Support. Social support from adults was measured using a three-item scale where youth were asked “How much do you feel that adults/your parents/your teachers care about you?” Responses were coded on a five-point scale ranging from *not at all (1) to very much (5)*. The three items were summed with total scores that could range from 3-15, with higher scores indicating higher levels of perceived social support from adults (present study $\alpha=.57$). Prior studies with the Add Health data have measured social support using these items in a similar way (Wight, Botticello, & Aneshensel, 2006).

For the second research question, a binary response variable was constructed using data from both Wave I and Wave IV. At Wave I, participants completed a nine-item derivative of the Center for Epidemiological Studies-Depression (CES-D), a measure that assesses symptoms associated with depression (Radloff, 1977). Total scores on the shortened CES-D could range from 0 to 27, with higher scores indicating more symptoms of depression. Each item for this measure is coded on a four-point scale from *never or rarely (0) to most of the time or all of the time (3)*. The CES-D is a frequently used, self-report measure that is well-validated and has been used to identify individuals at risk for depression (Radloff, 1977). Consistent with previous research using Add Health data, a cutoff score of 10 was used to indicate that an individual was

experiencing significant symptoms of depression (Boardman & Alexander, 2011; Esposito et al., 2017; Fletcher, 2009; Holway, Umberson, & Thomeer, 2017). For this study, if a participant scored a 10 or higher on the CES-D, they were considered to have symptoms of depression at Wave I (present study $\alpha = .78$). Previous research with the Add Health data have had similar alpha values ranging from .80 to .81, demonstrating adequate levels of internal reliability for this measure.

If a respondent answered yes to the question “Has a doctor, nurse or other health care provider ever told you that you have or had: depression?” at Wave IV they were considered to have a reported depression diagnosis. Based on the Wave I symptomology data and Wave IV diagnostic data a new binary depression variable was constructed with two possible categories.

The two possible outcome categories are:

- Depression0: symptoms of depression at Wave I, but no reported depression diagnosis at Wave IV.
- Depression1: symptoms of depression at Wave I and a reported depression diagnosis at Wave IV.

Self-esteem (as defined above), family connection (as defined above), maternal involvement (as defined above), maternal attachment, (as defined above), times a week eat dinner with family, neighborhood connection (as defined above), gender, race/ethnicity, presence of caring adults, presence of caring friends, whether the youth received physical in last year, suicidal ideation, ever received counseling, and whether the youth exercised in past week were used as predictors of the response variable, reported depression diagnosis. These variables have been linked empirically as

protective and risk factors for depression (Collishaw et al., 2016; Easterbrooks, Ginsberg, & Lerner, 2013; Fossum et al., 2007; McDonald et al., 2016; Milevsky, 2005; Paradis et al., 2011; Shaw et al., 2004; Yap, Pilkington, Ryan, & Jorm, 2014).

Data Analytic Strategy

Data analysis was conducted in three phases: 1) data cleaning 2) descriptive statistics 3) analysis. Data cleaning, was done in Stata 15 (StataCorp, 2017). Wave I and Wave IV data were merged and matched on participant ID. Frequencies and maximum and minimum values for each study variable were obtained and data quality were determined high. Responses that were “refused,” “don’t know,” or “does not apply” for any question were set to missing. Missing data was assessed and was low across study variables (ranging from 0% to 2%).

Given the low rates of missing data on each variable, listwise deletion was used. Listwise deletion omits cases from the data with missing data on any variable and is appropriate when the number of missing values for each variable is low. When the data are assumed to be missing completely at random, listwise deletion does not introduce bias since under the MCAR assumption the cases with complete data are thought to be equivalent to those cases without any missing data (Allison, 2001). Additionally, even if violations of MCAR or even missing at random exist for predictor variables, listwise deletion is robust and is often considered an “honest” method for handling missing data (Allison, 2001). Past studies have demonstrated trustworthy results using listwise deletion when missing data is low (Bennet, 2001; Dong & Peng, 2013). This strategy reduced the total number of cases in the sample from 4,489 to 3,782. Descriptive statistics were conducted for each variable to gain a good understanding and be able to

describe the study sample. Stata files were then converted to R files, which were used for analyses.

Random Forest. To address the first aim of this study, to develop a forecasting procedure that can be used to identify youth who are at risk of reporting a depression diagnosis as an adult, random forest (Brieman, 2001) was used to derive the forecasting algorithm. Analysis was conducted in R using the *randomForest* library. The random forest algorithm produces hundreds of classification and regression trees by taking a random sample of cases and predictors to determine the best splits of the data and creates forecasts by aggregating the results of all of the trees (Berk, 2016). Random forest builds on Classification and Regression Trees (CART) and the bagging algorithm, but is different in that random forest computes averages over hundreds of trees to address the instability of trees produced by CART for more stable estimates, the sampling of predictors, and use of out-of-bag (OOB) data for fitted values (Berk, 2016). Each classification tree is grown with a random sample of observations from the training data, about two-thirds of the whole sample, and the observations that are not chosen are used as the OOB test data (Berk et al., 2009). The random sampling is done with replacement, meaning that the same observation can be used more than once when the classification tree is being created (Berk & Hyatt, 2014). Inherently, random forests construct test data, and the analysis does not need to begin with the data being separated into a training and test data set (Berk & Hyatt, 2014). Fitted values are displayed in a confusion table, a key output from random forest, and is constructed from OOB data to represent out-of-sample performance so that the confusion tables are “honest” estimates (Berk, 2016). Forecasting performance is evaluated with the OOB test data using the same input variables and outcome as the training data, but with observations not used to build the forecasting procedure (Berk & Hyatt, 2014).

Random forest allows for many predictors and can even handle more input variables than observations. In practice, a large number of weak predictors on the aggregate can greatly improve forecasting accuracy (Berk, 2016; Berk, Sorenson, & Barnes, 2016). Consistent with the goal of machine learning, the main goal of random forest is to use all available input variables to achieve accurate forecasts and not to determine which input variables are most useful (Berk, Sorenson, & Barnes, 2016). Unlike conventional regression analysis, machine learning algorithms can handle correlated input variables (Berk, Sorenson, & Barnes, 2016).

Random forest also allows for the relative cost of false negative and false positive forecasting errors to be built directly into the algorithm (Berk et al., 2009). Berk and Hyatt (2014) highlight that the consequences of the two kinds of forecasting errors, a false positive and false negative, are different and that there are always tradeoffs to consider. For this study, a 10 to 1 target cost ratio was set, meaning that the consequence of predicting that an individual will be classified as not having a reported depression diagnosis as adult, but does have a reported depression diagnosis as an adult (false negative) is ten times worse than predicting that an individual will be classified as having a reported depression diagnosis, but does not have reported depression (false positive). The sample size for the less common outcome (having a reported depression diagnosis) was set to 400 (two-thirds of the data) and then the sample of the other group was tuned until a satisfactory cost-ratio close to the target of 10:1 was achieved. The rationale for this cost-ratio is that failing to identify someone who is at risk of having a reported depression diagnoses and not offering them supportive resources or services that can mitigate this risk is worse than offering someone extra support that may not be needed.

The degree to which the forecasting procedure is accepted depends partially on if the results make sense to stakeholders, and, therefore, it can be important to look at which predictors contribute most to the forecasts (Berk et al., 2009). Variable importance plots are constructed to demonstrate the individual contribution among the input variables and show the reduction in prediction accuracy when a predictor is shuffled (Berk, 2016). Partial dependence plots were also created to show the average relationship between each predictor and the response variable. Finally, empirical margins were computed to characterize the reliability of the forecasting results. While the first aim of this study focused on the ability to accurately forecast which youth are at risk of reporting a diagnosed depression disorder as an adult, the second aim focused on explanation and advancing understanding of why some youth with depression symptoms go on to report depression as an adult and other youth with depression symptoms do not go on to report a depression diagnoses as an adult.

Generalized Additive Model. To address the second aim of this study, to understand the developmental trajectories of depression for youth, Generalized Additive Model (GAM) was used. Analysis was conducted in R using *gam()* from the *mgcv* library. GAM with a binary response variable is an extension of binomial regression from the Generalized Linear Model (GLM), but does not assume that the predictor variables are linearly related to the response variable (Berk, 2016). Instead, each predictor can have its own functional relationship to the response variable with several link functions and disturbance distributions (Berk, 2016). Additionally, the nature of the relationship between the predictors and the response variable does not have to be specified and the data dictate the nature of the relationships (Austin, 2007). In healthcare, GAMs have been used to describe general cancer rates, lung cancer rates, and HIV occurrence

(Boyle et al., 2003; Clements, Armstrong, & Moolgavkar, 2005; Marschner & Bosch, 1998).

Output from GAM is like that of conventional regression and includes a deviance value and coefficients for non-smoothed factor or categorical variables that can be interpreted as if it were logistic regression. Fitted values are also generated for each smoothed predictor to show its relationship to the response variable and when the response variable is binary, as in this study, the GAM output includes fitted values in logit units and fitted proportions.

For this study, self-esteem, family connection, maternal involvement, maternal attachment, times a week eat dinner with family, neighborhood connection, gender, race/ethnicity, presence of caring adults, presence of caring friends, received physical in last year, suicidal ideation, ever received counseling, and whether exercised in past week served as predictor variables. While there are other potentially important predictors, such as genetic factors and exposure to traumatic events, to explain why a youth with behavioral health symptoms goes on to either report or not report a diagnosed depression condition as an adult, it should be noted that this model is misspecified and the analysis is operating under the wrong model perspective. Data are assumed to be generated randomly and independently from a joint probability distribution when working under the wrong model perspective (Berk, Brown, Buja, Geogre, & Zhao, 2018). It should also be noted that data for the proposed study are observational and for this study causal relationships cannot be established.

CHAPTER 3: RESULTS

Univariate Results

Descriptive Statistics. Descriptive statistics are reported for all study variables in Table 2. For continuous variables, the range, mean, and standard deviation are included in Table 2, and for binary categorical variables, sample size and percentages are reported in Table 2.

Table 2. Characteristics of the Study Sample, N=3,782

Characteristic	N (%)/ Mean (S.D)	Minimum	Maximum
<i>Youth Demographics</i>			
Female	2,035 (53.8%)	-	-
Asian	123 (3.3%)	-	-
Black/African American	882 (23.3%)	-	-
Hispanic/Latino/a	346 (9.2%)	-	-
White	2,651 (70.1%)	-	-
Good, very good, or excellent health	3,539 (93.6%)	-	-
Is happy	3,032 (80.2%)	-	-
Ever missed a social or recreational activity because of health or emotional problem	69 (1.8%)	-	-
Ever missed school because of health or emotional problem	158 (4.2%)	-	-
Poor appetite	509 (13.5%)	-	-
Trouble falling or staying asleep	883 (23.4%)	-	-
Trouble relaxing	510 (13.5%)	-	-
Moodiness	1,388 (36.7%)	-	-
Frequent crying	251 (6.6%)	-	-
Feeling fearful	237 (6.3%)	-	-
Ever received counseling	461 (12.2%)	-	-
Received yearly physical examination	2,601 (68.8%)	-	-
Learned about where to go for help with a health problem	3,142 (83.1%)	-	-
Learned about suicide in school	2,575 (68.1%)	-	-
Learned about stress in school	2,433 (64.3%)	-	-
Repeated a grade in school	707 (18.7%)	-	-
Received an out-of-school suspension	952 (25.2%)	-	-
Expelled from school	142 (3.8%)	-	-

Depression score on CES-D	5.6 (4.1)	0	25
Significant symptoms of depression (CES-D score 10 or higher)	612 (16.2%)	-	-
Depression diagnosis (Wave IV)	591 (15.6%)	-	-
Mom perceived as warm	3,453 (91.3%)		
Family connection/support	15.2 (3.2)	4	20
Times having dinner with at least one of your parents in the same room in the last week	4.7 (2.4)	0	7
Self-esteem	11.2 (3.5)	6	30
Suicidal ideation	505 (13.4%)	-	-
Ever attempt suicide	140 (3.7%)	-	-
Have friends that have tried to kill themselves	680 (18.0%)	-	-
Have friends who died by suicide	107 (2.8%)	-	-
Have family members who have tried to kill themselves	180 (4.8%)	-	-
Have family members who have died by suicide	39 (1.0%)	-	-
Have friends who care	3,238 (85.6%)	-	-
Participate in sports	2,746 (72.6%)	-	-
Exercise in the past week	3,171 (83.8%)	-	-
Autonomy from parents	5.1 (1.5)	0	7
Maternal attachment	9.4 (1.1)	2	10
Maternal involvement	4.1 (2.0)	0	10
Adult social support	12.8 (1.7)	3	15
Neighborhood connection	4.8 (1.0)	3	8
<i>Youth Exposure to Violence</i>			
Saw someone shoot or stab another person	409 (10.8%)	-	-
Had someone pull a knife or gun on them	429 (11.3%)	-	-
Been shot or stabbed	41 (1.1%)	-	-
Been cut or stabbed	157 (4.2%)	-	-
Been jumped	365 (9.7%)	-	-
<i>Parent Characteristics</i>			
Married	2,783 (73.6%)	-	-
Has a college degree or higher	1,001 (26.5%)	-	-
Employed	2,814 (74.4%)	-	-
Is happy	3,659 (96.8%)	-	-
Good, very good or excellent health	3,305 (87.4%)	-	-
Talks about school with youth	1,704 (45.1%)	-	-
Talks about grades with youth	1,968 (52.0%)	-	-
<i>Family Financial Demographics</i>			
Family has trouble paying bills	645 (17.1%)	-	-

Family receives public assistance	296 (7.8%)	-	-
Family receives food stamps	432 (11.4%)	-	-
Family receives Aid to Families with Dependent Children (AFDC)	249 (6.6%)	-	-

Youth Demographic Characteristics. During Wave I of the Add Health study, youth were enrolled in grades 7 through 12 and between 12 and 20 years old. Almost twenty percent of youth ($n = 707$, 19%) reported repeating a grade in school. About half ($n = 2,035$, 54%) of youth identified as female. A little over two-thirds of youth identified as White ($n = 2,651$, 70%), about a quarter of youth ($n = 882$, 23%) identified as Black or African American, 9% of youth ($n = 346$) identified as Hispanic or Latino/a, and 3% of youth ($n = 123$) identified as Asian (youth could identify with more than one race/ethnicity).

Self-Reported Health and Service Use. The majority of youth (94%, $n = 3,539$) reported being in either good, very good, or excellent health, and 80% ($n = 3,032$) reported being happy. Most of the youth reported that they never missed a social/recreational activity ($n = 3,713$, 98%) or school ($n = 3,624$, 96%) because of a health or emotional problem. About two-thirds ($n = 2,601$, 69%) of youth received an annual physical exam, and 12% of youth ($n = 461$) reported receiving psychological or emotional counseling in the last year. Average depression score on the CES-D was 5.6 ($SD = 4.1$) on a scale from 0-25. Depression score was skewed to the right, which indicates that most youth had low depression scores. However, 16% ($n = 612$) of youth reported having significant symptoms of depression (defined by having a CES-D score of 10 or higher). This percentage is consistent with recent statistics, which estimate that about 13% of youth between the ages of 12 to 17 have experienced at least one depressive episode over the last year (NSDUH, 2017). Thirteen percent of youth ($n =$

505) reported having seriously thought about committing suicide at some point over the last year, and 4% of youth ($n = 140$) have attempted suicide in the last year. These results are similar to recent estimates of youth suicidal thoughts and attempts (CDC, 2016).

Support and Family Connection. The majority of youth, 86% ($n = 3,238$), reported having friends who care about them and high levels of social support from adults in their lives, with an average score of 12.8 ($SD = 1.7$) on a scale from 3-15. Social support from adults was skewed to the left meaning that most youth reported high levels of social support from adults in their life. Youth also reported high levels of maternal attachment, with an average score of 9.4 ($SD = 1.1$) out of 10. Maternal attachment scores were also skewed to the left, indicating that most youth reported high levels of maternal attachment. However, maternal involvement reporting was relatively low, with an average score of 4.1 ($SD = 2.0$) out of 10. Average family connection/support score was high at 15.2 ($SD = 3.2$) out of 20. On average, youth reported having dinner five nights a week ($SD = 2.4$) with at least one parent present over the last seven days. Many youth reported having dinner with at least one parent all seven nights. A majority of youth, 91% ($n = 3,453$), also perceived their mother as being warm and loving towards them. These variables are important to consider within the context of depression as we know that family support and family relationships play a protective role for youth in the development of behavioral health conditions.

Other Protective and Risk Factors. Overall, youth reported good levels of self-esteem with an average score of 11.2 ($SD = 3.5$) on a scale from 6 to 30, with lower scores indicating higher levels of self-esteem, and relatively high levels of autonomy from their parents ($M = 5.1$, $SD = 1.5$, Range 0-7). Most youth reported participating in a sports activity, 73% ($n = 2,746$), or exercising, 84% ($n = 3,171$), over the last week. Self-

esteem, social activities, and physical activity have been linked to depression as protective factors, and, therefore, are important variables to consider (Collishaw et al., 2016; Fiorilli, Capitello, Barni, Buonomo, & Gentile, 2019; Hilbert et al., 2019; McDonald et al., 2016; Sowislo & Orth, 2013). Some youth reported having either a friend (18%, $n = 680$) or family member (5%, $n = 180$) who has attempted suicide or a friend (3%, $n = 107$) or family member (1%, $n = 39$) who has died by suicide. While most youth reported relatively low exposure to violence in their community, 11% ($n = 429$) of youth reported having had someone pull a knife or gun on them, and 11% ($n = 409$) reported having seen someone shoot or stab another person. Youth also reported moderate levels of neighborhood connection, with an average score of 5 ($SD = 1.0$) on a scale from 3-8. Exposure to community violence and knowing someone who has either attempted or died by suicide have been shown in past research to have a negative impact on youth mental health (Abrutyn & Mueller, 2014; Chen, Corvo, Lee, & Hahm, 2017; Gopalan et al., 2010; Gould et al., 2018). Remaining youth characteristics are provided in Table 2.

Primary Caregiver and Financial Demographics. The majority of parents who completed the Parent Questionnaire Survey were married, 74% ($n = 2,783$), and employed outside of the home, 74% ($n = 2,814$). About a quarter, 27% ($n = 1,001$), had a four-year college degree or higher. Almost all parents, 97% ($n = 3,659$), reported being happy, and 87% ($n = 3,305$) rated their health as good, very good, or excellent. About a fifth, 17% ($n = 645$), of parents reported that they had trouble paying bills. Eleven percent of families ($n = 432$) reported receiving food stamps, 8% of families ($n = 296$) reported receiving public assistance, and 7% of families ($n = 249$) reported receiving Aid to Families with Dependent Children (AFDC).

Depression Diagnosis. At Wave IV, when youth were adults between the ages of 24 and 32, 16% ($n = 591$) reported having received a depression diagnosis by a

health care professional. Table 3 shows similarities and differences in youth characteristics between adults with and without a reported depression diagnosis. Differences were observed in various areas; while these differences are small they are consistent with other research related to depression. On average, adults with a reported depression diagnosis were more likely to identify as female or White, in comparison to, adults without a reported depression diagnosis. Specifically, 73% of individuals with a reported depression diagnosis identified as female compared to 50% of individuals without a reported depression diagnosis who identified as female. Furthermore, among individuals who identified as White, 86% reported a depression diagnosis, and 68% of individuals did not report a depression diagnosis. Among individuals who identified as Black or African American, 15% reported a depression diagnosis, and 25% did not report a depression diagnosis.

Overall, adults with a reported depression diagnosis experienced difficulties in several domains during their youth relative to adults without a reported depression diagnosis. For example, adults with a reported depression diagnosis reported being less happy, having trouble sleeping, being moody, crying often, feeling fearful, having thought about suicide, having attempted suicide, and knowing someone who has attempted or died by suicide as a youth compared to adults without a reported depression diagnosis. Given the symptoms associated with depression, these results are not surprising. Furthermore, a quarter (25%) of adults with a reported depression diagnosis reported receiving counseling as a youth compared to only 10% of adults without a reported depression diagnosis. Adults with a reported depression diagnosis were also more likely to experience significant symptoms of depression as a youth (28%) relative to adults without a reported depression diagnosis (14%).

Table 3: Adults with a Reported Depression Diagnosis Compared to Adults Without a Reported Depression Diagnosis, N=3,782

Characteristic	Reported Depression (n = 591) % (N)/ Mean (S.D)	No Reported Depression (n = 3,191) % (N)/ Mean (S.D)
<i>Youth Demographics</i>		
Female***	73.3% (433)	50.2% (1,602)
Asian	2.0% (12)	3.5% (111)
Black/African American***	14.6% (86)	25.0% (796)
Hispanic/Latino/a*	6.9% (41)	9.6% (305)
White***	82.6% (488)	67.8% (2,163)
Good, very good, or excellent health**	90.5% (535)	94.1% (3,004)
Is happy***	72.9% (431)	81.5% (2,601)
Ever missed a social or recreational activity because of health or emotional problem**	3.4% (20)	1.5% (49)
Ever missed school because of health or emotional*** problem	7.3% (43)	3.6% (115)
Poor appetite***	22.0% (130)	11.9% (379)
Trouble falling or staying asleep***	34.7% (205)	21.3% (678)
Trouble relaxing***	21.3% (126)	12.0% (384)
Moodiness***	50.6% (299)	34.1% (1,089)
Frequent crying***	15.9% (94)	4.9% (157)
Feeling fearful***	10.5% (62)	5.5% (175)
Ever received counseling***	24.9% (147)	9.8% (314)
Received yearly physical examination	71.7% (424)	68.2% (2,177)
Learned about where to go for help with a health problem*	79.9% (472)	83.7% (2,670)
Learned about suicide in school	66.0% (390)	68.5% (2,185)
Learned about stress in school	61.9% (366)	64.8% (2,067)
Repeated a grade in school	18.6% (110)	18.7% (597)
Received an out-of-school suspension	24.9% (147)	25.2% (805)
Expelled from school	3.1% (18)	3.9% (124)
Depression score on CES-D***	7.3 (4.8)	5.3 (3.9)
Significant symptoms of depression (CES-D score 10 or higher) ***	28.4% (168)	13.9% (444)
Mom perceived as warm***	86.8% (591)	92.1% (2,940)
Family connection/support***	14.3 (3.4)	15.4 (3.1)

Times a week have dinner with least one of your parents in the same room	4.6 (2.6)	4.8 (2.4)
Self-esteem***	12.3 (3.9)	11.0 (3.4)
Suicidal ideation***	26.9% (159)	10.9% (347)
Ever attempt suicide***	9.0% (53)	2.7% (87)
Have friends that have tried to kill themselves***	28.6% (169)	16.0% (511)
Have friends that died by suicide ***	5.1% (30)	2.4% (77)
Have family members that have tried to kill themselves***	8.8% (52)	4.0% (128)
Have family members who have died by suicide**	2.2% (13)	0.8% (26)
Have friends who care	85.1% (503)	85.7% (2,735)
Participate in sports**	67.9% (401)	73.5% (2,345)
Exercise in the past week	85.1% (503)	85.7% (2,735)
Autonomy from parents	5.1 (1.5)	5.1 (1.5)
Maternal attachment**	9.3 (1.2)	9.4 (1.0)
Maternal involvement	4.2 (2.0)	4.0 (2.0)
Adult social support**	12.6 (1.8)	12.9 (1.7)
Neighborhood connection	4.8 (1.0)	4.8 (1.0)
<i>Youth Exposure to Violence</i>		
Saw someone shoot or stab another person	11.2% (66)	10.8% (343)
Had someone pull a knife or gun on them	11.0% (65)	11.4% (364)
Been shot or stabbed	1.2% (7)	1.1% (34)
Been cut or stabbed	4.7% (28)	4.0% (129)
Been jumped	8.1% (48)	9.9% (317)
<i>Parent Characteristics</i>		
Married	73.3% (433)	73.6% (2,350)
Has a college degree or higher	26.4% (156)	26.5% (845)
Employed	72.8% (430)	74.7% (2,384)
Is happy	96.8% (572)	96.7% (3,087)
Good, very good or excellent health	86.1% (509)	87.6% (2,796)
Talks about school with youth*	41.0% (242)	45.8% (1,462)
Talks about grades with youth*	48.1% (284)	52.8% (1,684)
<i>Family Financial Demographics</i>		
Family has trouble paying bills	16.8% (99)	17.1% (546)
Family receives public assistance	8.3% (49)	7.7% (247)
Family receives food stamps	11.7% (69)	11.4% (363)
Family receives Aid to Families with Dependent Children (AFDC)	8.3% (49)	6.3% (200)

*p<.05, **p<.01, ***p<.0001 (differences between subgroups)

Random Forests Results

Random forest was used to derive a forecasting algorithm to identify youth who are at risk of having a reported depression diagnosis as an adult. Table 4 below presents the random forest confusion table. The actual cost ratio is 9.4, which is very close to the target of 10:1. Of the 3782 cases, 1554 were misclassified. The overall cost-weighted error rate is 77% $[(10 * 150) + (1404)] / 3782$. Use error, as shown in Table 4, is particularly important to pay attention to as it provides estimates of how well in a real-world practice setting the random forests algorithm will forecast. When a forecast is for no diagnosis, the assigned class is correct 92% of the time and when a forecast is for diagnosis, the assigned class is correct 24% of the time. The large difference in forecasting skill is related to the 10 to 1 cost ratio. From a policy and practice perspective, we are willing to accept a high number of false positives as a tradeoff to achieving high accuracy and a low number of false negatives. Without any predictors, using Bayes classifier, no diagnosis would always be forecasted, and we would be wrong 16% of the time. However, after this random forest application, when no diagnosis is forecasted, we are only wrong 8% of the time. This improvement is dramatic in the ability to forecast no diagnosis. If this procedure were used in practice, the error rate for forecasting no diagnosis would be reduced by half. If you look at the columns of the confusion table, only 150 cases classified as no diagnosis actually had a diagnosis (false negatives). While the 1404 cases who were classified as having a diagnosis, but did not have a diagnosis (false positives) may seem substantial, this finding is attributed to the policy choice that was made and how the cost ratio was set. For this study, it was important not to miss identifying cases as not having a diagnosis and in exchange, the price that was paid was over-classifying cases as having a diagnosis. Again though, this

decision was a policy choice that was made, and the cost ratio could be readjusted to meet the needs and expectations of its stakeholders. From the model error, we can see that random forests classifies 25% of cases with a depression diagnosis incorrectly and 44% of cases with no depression diagnosis incorrectly. The empirical margins were computed to characterize the reliability of these results. The average of the empirical margins were .25, which suggests there may be some reliability concerns with these results.

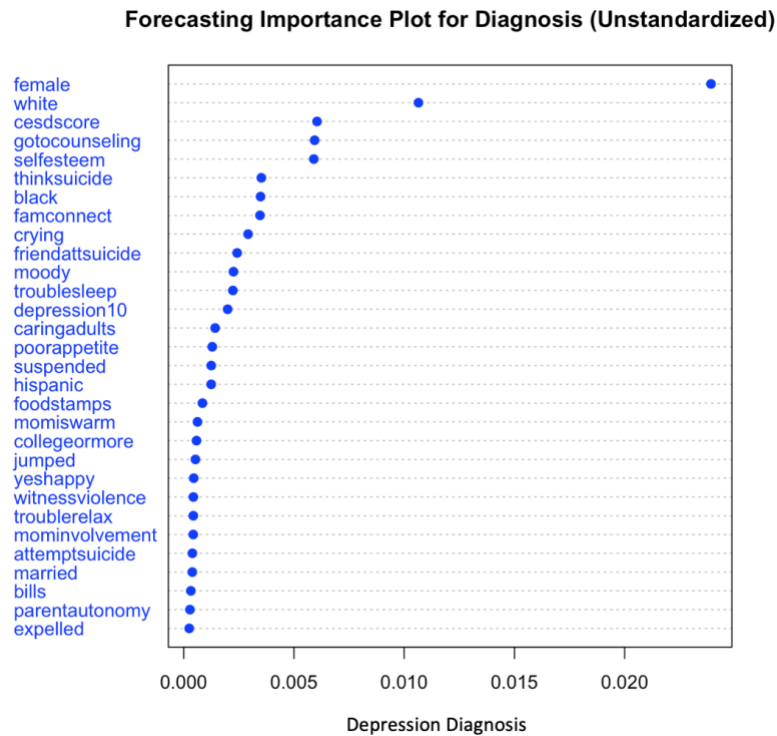
Table 4. Confusion Table for Forecasting a Reported Depression Diagnosis with a 10:1 Cost Ratio (N=3782)

	Classify as No Diagnosis	Classify as Diagnosis	Model Error
No Diagnosis	1787	1404	0.44
Diagnosis	150	441	0.25
Use Error	0.08	0.76	Overall Error= 0.41

Figure 1 presents the unstandardized variable importance plots for forecasting whether a youth will report a diagnosed depressive disorder as an adult. Variable importance plots show the reduction in classification accuracy when each predictor is randomly shuffled. Whether a youth is female and White are the two most important inputs for forecasting a reported depression diagnosis. When female is shuffled, classification accuracy decreases by 2.4 percentage points. When White is shuffled, classification accuracy decreases by 1 percentage point. Depression score on the CES-D, having ever gone to counseling as a youth, youth self-esteem, and suicidal ideation as a youth are the next most important predictors in terms of classification accuracy for an individual having a reported depression diagnosis as an adult. When CES-D depression score is shuffled, classification accuracy decreases by 0.9 of a percentage point, and when having ever gone to counseling as a youth is shuffled, accuracy

decreases by 0.7 of a percentage point. Additionally, when self-esteem score is shuffled, accuracy decreases by 0.6 of a percentage point, and when suicidal ideation is shuffled accuracy decreases by 0.5 of a percentage point. The remaining inputs are of little importance in terms of classification accuracy (contributing to less than 0.5% reduction in accuracy).

Figure 1: Variable Importance Plot for a Reported Depression Diagnosis with a 10 to 1 Target Cost Ratio (N=3782)



Based on previous research and knowledge about behavioral health conditions and depression, it is not surprising that these predictors contribute most in terms of classification accuracy. However, the fact that gender and race matter and contribute the most to forecasting accuracy over all other predictors, including score on a depression

inventory, risk factors, and protective factors, is a bit surprising. But we must keep in mind that what is being forecasted is presence of a reported depression diagnosis, not whether an individual has experienced depression. It is likely that this finding speaks more to health disparities that exist when it comes to receiving a depression diagnosis rather than experiencing depression. This finding is in detail in the next chapter.

While variable importance plots are helpful in terms of knowing the usefulness of each input, partial dependence plots are useful for describing how each predictor is related to the response variable. Figure 2 presents results for the binary variables female, White, ever received counseling, and suicidal ideation. Holding all other variables constant, for individuals who identify as female, White, ever received counseling as a youth, and experienced suicidal ideation as a youth, the chances of having a depression diagnosis as an adult are greater relative to males, individuals who do not identify as White, individuals who never received counseling as a youth, and individuals who didn't have suicidal ideation as a youth. These results are consistent with the literature in terms of the likelihood of having a depression diagnosis being greater as an adult if, as a youth, the individual ever received counseling or had thoughts of suicide (Gould et al., 2018; Maulik, Eaton, & Bradshaw, 2011). Additionally, women are more likely to be diagnosed with depression compared to men (Mayo Clinic, 2019; Whiteman, Ruggiano, & Thomlison, 2016), so this finding is also consistent with previous research. The relationships between the predictor variables female, suicidal ideation, and ever receiving counseling and the response variable are weak, though, and the chances of having a depression diagnosis do not differ that much when the logits are converted into proportions. For example, the smallest logit for females is 0.00 compared to about -0.13 for males. The proportions of having a reported depression diagnosis can be compared as .50 for females and .46 for males. Additionally, the largest logit for youth

with suicidal ideation is about 0.19 compared to 0.0 for youth without suicidal ideation. When converted to proportions, these estimates become .55 and .50, respectively. For youth who have ever received counseling, the chances of having a reported depression diagnosis as an adult are higher compared to youth who never received counseling. The largest logit for youth who received counseling is 0.15 compared to 0.0 for youth who never received counseling. When converted to proportions, these estimates become .54 and .50, respectively. For youth who identify as White, the chances of having a reported depression diagnosis as an adult are higher compared to youth who identify as Black or African American, Hispanic/Latino/a, or Asian. The smallest logit for youth who identify as White is 0.00 and -1.14 for youth who identify as Black or African American, Hispanic/Latino/a or Asian. When converted to proportions, these estimates become .50 and .24, respectively. The proportion of White youth who report a depression diagnosis indicates a strong association between identifying as White and reporting a depression diagnosis as an adult. This finding is interesting as it raises questions about depression prevalence and access to treatment. Given that the outcome is reported depression diagnosis, the role of stigma, help-seeking behaviors, and access to treatment provide possible explanations of this finding and are discussed in detail in the next chapter.

Figure 3 presents results for the depression score on the CES-D and self-esteem score. As CES-D depression score increases, chance of having a reported depression diagnosis as an adult also increases. This increase is linear, and once the CES-D score reaches 15, levels off (a score of 10 or higher on CES-D indicates symptoms of depression). The largest logit is about .17, and the smallest logit is about -0.08. As proportions, they become .58 and .46, indicating a moderate association between youth depression score and having a reported depression diagnosis as an adult. Given that the CES-D is a depression screening tool and that a higher score indicates greater

symptoms of depression, this relationship would be expected. As self-esteem score increases, the chance of having a reported depression diagnosis as an adult also increases. Given how the self-esteem measure is scored, higher scores on the scale indicate lower levels of self-esteem. This finding means that lower levels of self-esteem as a youth are associated with an increased chance of having a reported depression diagnosis as an adult. This increase is also linear, and once the score reaches 20, it levels off. The largest logit is about .14 and the smallest logit is about -0.05. As proportions, they become .57 and .48, indicating a moderate association between youth self-esteem score and reporting a depression diagnosis as an adult. This finding is consistent with previous research which has shown an association between low self-esteem and depression (Fiorilli, Capitello, Barni, Buonomo, & Gentile, 2019; Hilbert et al., 2019; Sowislo & Orth, 2013).

Figure 2: Partial Response Plots for a Reported Depression Diagnosis on Binary Inputs (N=3782)

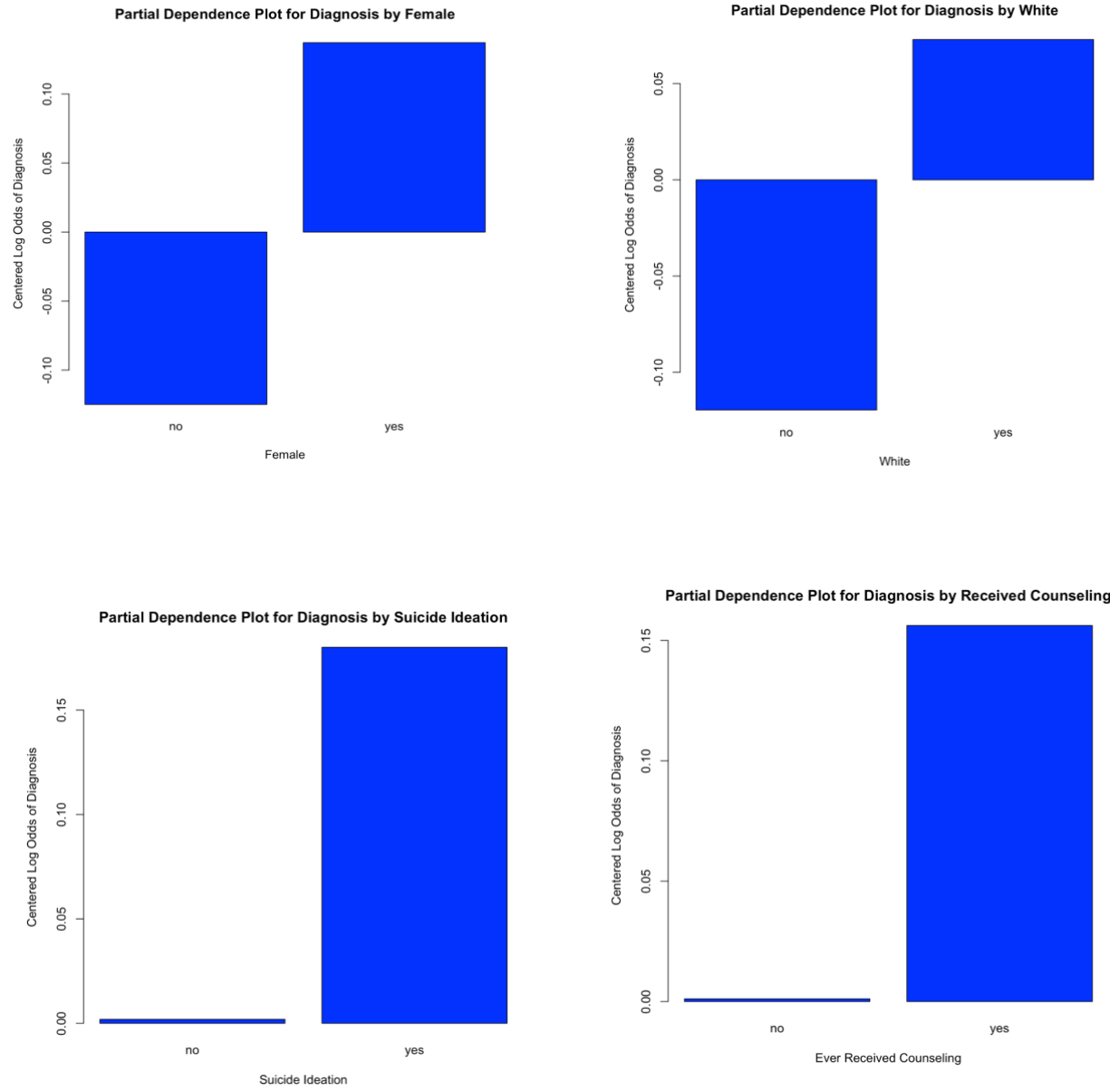
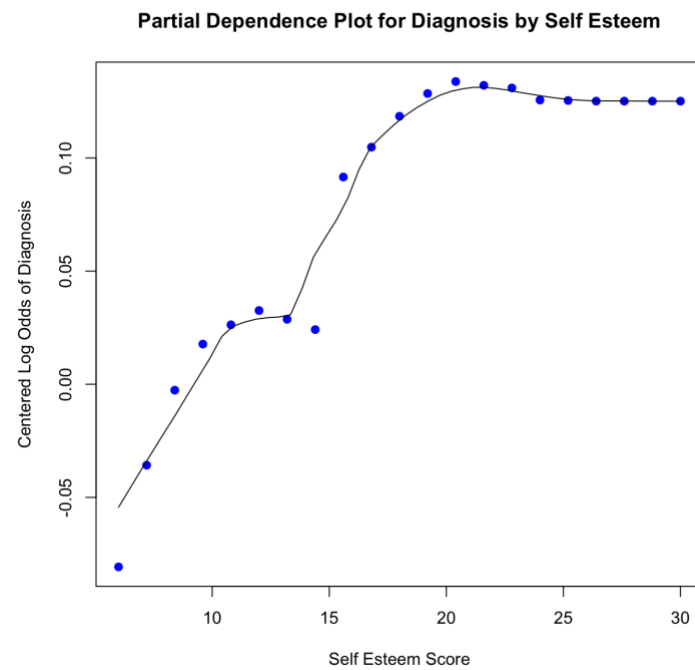
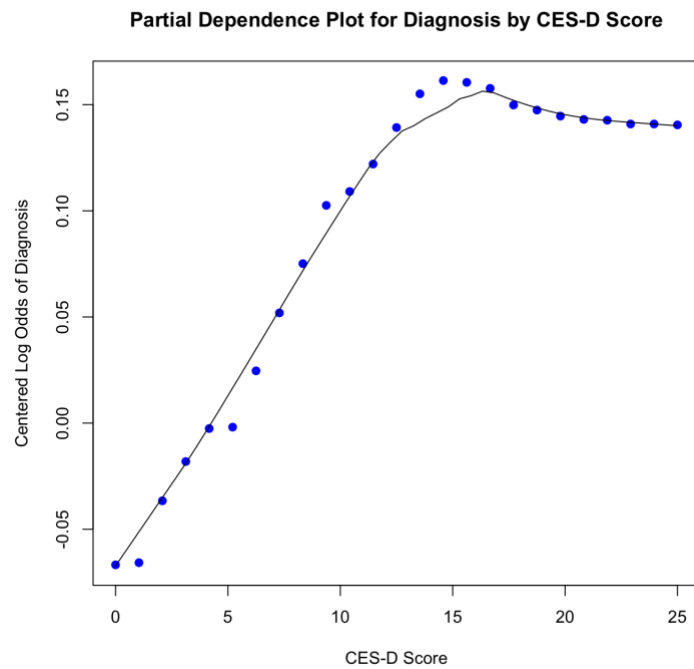


Figure 3: Partial Dependence Plots for a Reported Depression Diagnosis on Quantitative Inputs (N=3782)



Generalized Additive Model Results

Summary Statistics. Of the total original sample ($n = 4,489$), 769 (17%) had symptoms of depression as a youth at Wave I, and at Wave IV, about a quarter, 27% ($n = 207$), of the individuals with depression symptoms as youth had a reported depression diagnosis as an adult. The remaining 73% ($n = 562$) of individuals with symptoms of depression as a youth had no reported depression diagnosis as an adult. On average, adults with a reported depression diagnosis were more likely to be female and White relative to adults without a reported diagnosis. Additionally, adults with a reported depression diagnosis were more likely to have received an annual physical examination and experienced suicidal ideation as a youth. For example, 48% of adults with a reported depression diagnosis had suicidal ideation as a youth compared to 31% of adults with no reported depression diagnosis. Adults with a reported depression diagnosis reported lower levels of family connection/support and lower levels of self-esteem as a youth compared to adults without a reported depression diagnosis. Further, 38% of adults with a reported depression diagnosis reported receiving counseling as a youth whereas 18% of adults without a reported depression diagnosis reported receiving counseling as a youth. Table 5 presents summary statistics for all variables included in the Generalized Additive Model analysis.

Table 5. GAM Summary Statistics

Characteristics	Full Sample (n = 769) N (%)/ Mean (S.D)	Reported Depression Diagnosis (n = 207) N (%)/ Mean (S.D)	No Reported Depression Diagnosis (n = 562) N (%)/ Mean (S.D)
Depression Diagnosis	207 (26.9%)	-	-
Female	538 (70.0%)*	169 (81.6%)	369 (65.7%)
White	488 (63.5%)*	161 (77.8%)	327 (58.2%)
Ever received counseling	179 (23.3%)*	79 (38.2%)	100 (17.8%)
Have friends that care	595 (77.4%)	165 (79.7%)	430 (76.5%)
Have adults that care	610 (79.3%)	157 (75.9%)	453 (80.6%)
Received yearly physical examination	481 (62.6%)*	143 (69.1%)	338 (60.1%)
Suicidal ideation	272 (35.4%)*	100(48.3%)	172 (30.6%)
Exercise in the past week	629 (81.8%)	176 (85.0%)	453 (80.6%)
Family connection/support	12.9 (3.5) *	12.5 (3.5)	13.1 (3.4)
Maternal involvement	4.0 (2.1)	4.2 (2.0)	3.9 (2.1)
Self-esteem	14.0 (4.0) **	14.8 (4.3)	13.7 (3.9)
Times a week have dinner with least one of your parents in the same room in last seven days	3.9 (2.7) *	4.2 (2.7)	3.7 (2.6)
Maternal attachment	8.9 (1.5)	8.9 (1.6)	8.9 (1.5)
Neighborhood attachment	3.9 (1.0)	3.8 (1.0)	3.9 (1.0)

*p<.05, **p<.01, ***p<.0001 (differences between subgroups)

Generalized Additive Model (GAM). Generalized Additive Model analysis was used to estimate relationships between presence of a reported depression diagnosis as an adult and youth characteristics. GAM was implemented in R using the *gam()* procedure in the *mgcv* library. This procedure uses penalized regression splines as its smoothing function and the smoothing parameter estimation problem is addressed by the generalized cross-validation(GCV) statistic. Specifically, GAM was implemented to answer the question: What features distinguish youth with depressive symptoms who report a depression diagnosis as an adult from youth with depressive symptoms who do not report a depression diagnosis as an adult?

A semiparametric model was used with female, White, ever received counseling, having friends who care, having adults who care, receiving a physical in the last year, suicidal ideation, exercising in the past week, family connection, maternal involvement, self-esteem, times a week eat dinner as a family in the last week, maternal attachment, and neighborhood connection as predictors of the response variable that was coded to represent having symptoms of depression as a youth and a depression diagnosis as an adult. All predictors were from when the adult was a youth (Wave I). These predictors were selected as they include various protective and risk factors that have been deemed important and associated with depression according to previous research (Collishaw et al., 2016; Easterbrooks, Ginsberg, & Lerner, 2013; Kim-Cohen, Moffitt, Caspi, & Taylor, 2004; McDonald et al., 2016)

Overall, 13.2% of the deviance was accounted for by the 14 predictor variables. Gender, race, and receiving counseling as youth were the most important predictors in terms of having a reported depression diagnosis as an adult among individuals who experienced symptoms of depression as youth. When the individual is female, the odds of having a reported diagnosis of depression as an adult are multiplied by 2.50. When the individual is White, the odds of having a reported diagnosis of depression as an adult are multiplied by 2.40. Additionally, when the individual received counseling as a youth, the odds of having a reported diagnosis of depression as an adult are multiplied by 2.50. In practical terms, these associations are likely to be important and may contribute to understanding of who seeks care for behavioral health difficulties such as depression and understanding of health disparities. These considerations are discussed in detail in the next chapter.

The coefficients for the remaining variables: having adults who care, having friends who care, having received a physical in the last year, and having exercised in the last

week ranged from -0.30 to 0.31, which convert to small differences in odds ratios for having a depression diagnosis as an adult among individuals who experienced symptoms of depression as a youth (less than 1.5). All GAM results are displayed in Table 6.

Table 6 GAM Results

Linear Terms	Estimate (SE)	T-statistic
Intercept ***	-3.0 (0.4)	-7.1
Female***	0.9 (0.2)	4.2
White***	0.9 (0.2)	4.2
Ever received counseling***	0.9 (0.2)	4.5
Have friends who care	0.2 (0.3)	
Have adults who care	-0.3 (0.3)	-1.1
Suicidal ideation **	0.5 (0.2)	2.7
Exercise in the past week	0.3 (0.2)	1.3
Received yearly physical examination	0.3 (0.2)	1.6
Smoother Terms	EDF	X ²
Family connection/support	1.0	1.0
Maternal involvement	3.5	4.5
Self-esteem	1.0	2.3
Times a week have dinner* with at least one of your parents in the same room in the last week	1.7	8.0
Maternal attachment	1.7	4.6
Neighborhood attachment	1.0	0.1
Deviance Explained = 13.2%		

*p<.05, **p<.01, ***p<.0001

Figure 4 shows the smoothed plots and how the predictor variables of self-esteem, family connection, maternal involvement, times a week eating dinner as a family in the last week, maternal attachment, and neighborhood connection are related to the response variable of reported depression diagnosis as an adult among individuals who experienced depression symptoms as youth. The values for the effective degrees of freedom, which are noted on the vertical axis for the six smoothed predictors, were determined by an automated search over values of the generalized cross-validation

(GCV) statistic. The shaded areas in the plots are the error bands and represent plus and minus two standard deviations for the fitted values. They demonstrate uncertainty associated with the estimates and are most prevalent in areas where the data are sparse.

In general, self-esteem, times a week eating dinner as a family in the last week, and maternal attachment have a linear relationship with the logit of depression diagnosis in adulthood when depression symptoms were present as a youth. Overall, the relationships are positive and as youth self-esteem score, times a week a youth eats dinner with their family, and youth maternal attachment score increase, the odds of having a reported depression diagnosis as an adult when depression symptoms were present as a youth also increase.

While the relationships appear to be weak, when the logit units are transformed into probability units, the difference in the proportion of having a depression diagnosis as an adult when depression symptoms were present as a youth is about .27 greater when a youth has an average self-esteem score of 25 compared to youth with an average self-esteem score of 10 (lower scores indicate higher levels of self-esteem). The difference in the proportion of having a reported depression diagnosis as an adult when depression symptoms were present as a youth is about .12 greater for youth who reported having dinner seven days per week with at least one member of their family as compared to youth who reported having dinner one day per week with at least one member of their family. The difference in proportion for having a reported diagnosis of depression as an adult when depression symptoms were present as a youth is about .20 greater when a youth has an average maternal attachment score of nine compared to youth with a maternal attachment score of six. Family connection/support also appears to be linearly related with the logit of depression diagnosis as an adult when depression symptoms

were present as a youth. Overall, this relationship is negative, meaning that as family connection/support scores increase, the odds of having a reported depression diagnosis as an adult when depression symptoms were present as a youth decrease. This relationship is weak, and when logit units are transformed into probability units, the difference in proportion for having a reported depression diagnosis as an adult when depression symptoms were present as a youth is less than .10 when a youth has an average family connection/support score of five compared to youth with a family connection/support score of 15.

Given that these relationships are all linear, this finding suggests that they probably did not have to be smoothed in the first place. There is no relationship between neighborhood connection and having a depression diagnosis as an adult when depression symptoms were present as youth, given the horizontal line displayed in the graph. A horizontal line means that the slope is zero, which tells us that the value of y does not change based on the value of x . The relationship between maternal involvement and logit of a reported depression diagnosis as an adult when depression symptoms were present as a youth appears slightly non-linear. Specifically, the relationship is positive until maternal involvement score reaches a 3, then slightly dips and becomes negative from 4-5, and then becomes positive again as maternal involvement score reaches 5 or higher. While the possible range of scores on this scale is from 0-10, with higher scores indicating higher levels of maternal involvement, from a practical standpoint the dip/change in direction does not seem to have clear meaning or significance as it relates to having a reported depression diagnosis as an adult when depression symptoms were present as a youth. The relationship between maternal involvement and a reported depression diagnosis is weak, and when logit units are transformed into probability units, the difference in proportion for having a reported

depression diagnosis as an adult when depression symptoms were present as a youth is less than .05 when a youth has an average maternal involvement score of three compared to youth with a maternal involvement score of five.

When the predictors are combined in a linear way, we receive fitted values; a histogram of the fitted values in proportions is shown in Figure 5. Fitted proportions range from 0.10 to 0.90. The graph shows considerable variation among the cases and is heavily skewed to the right, with the majority of observations falling between 0.1 and 0.3. As shown in Table 5, the mean value for cases with a reported depression diagnosis as an adult when depression symptoms were present as a youth falls within this range (27%). In the histogram, 0.5 can be thought of as an arbitrary cut off, and given that most of the cases fall below this mark, it indicates that most cases are predicted not to have a reported depression diagnosis as an adult when depression symptoms were present as a youth. From a practical and clinical standpoint, it may be important to know which individuals are represented in the right of the histogram.

Figure 4. Correlates of a Reported Depression Diagnosis and Predictor Variables

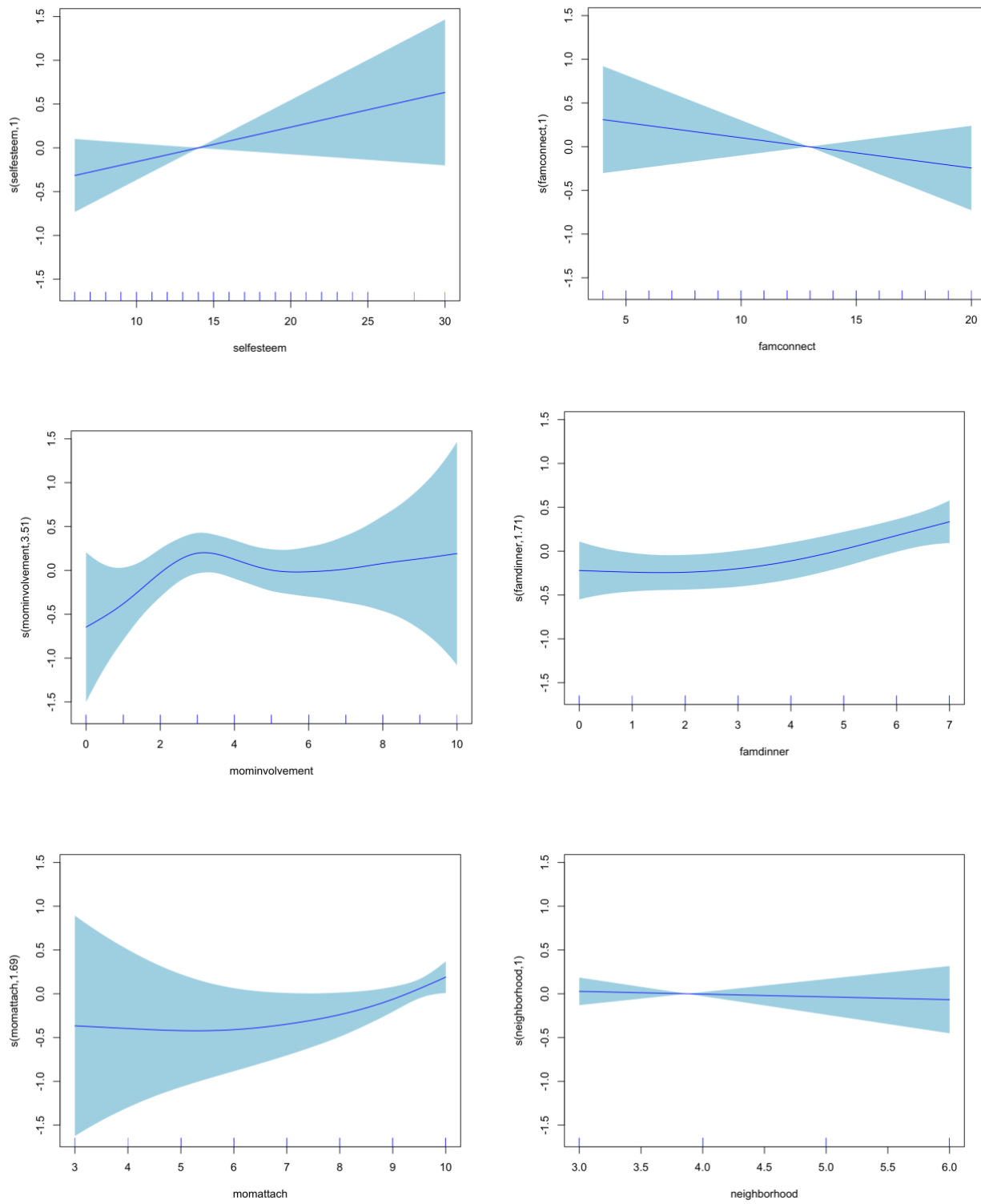
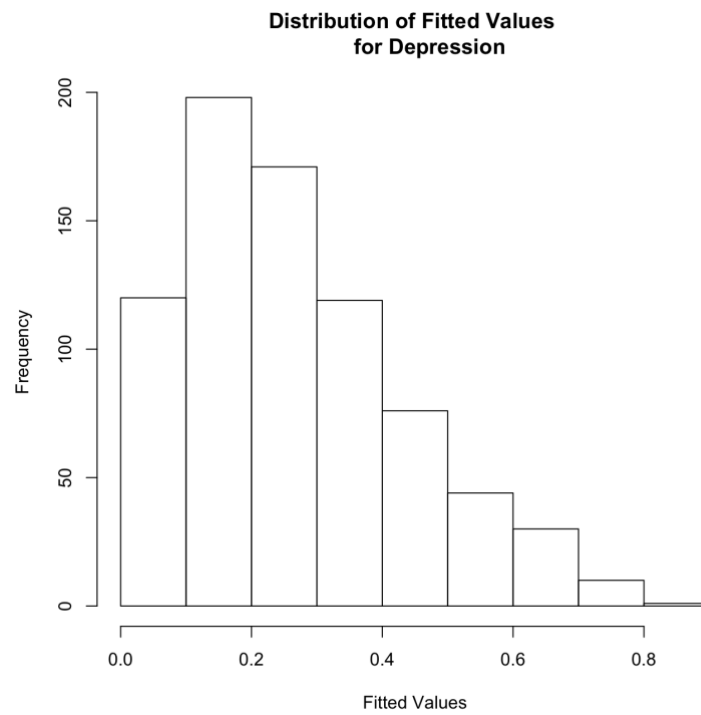


Figure 5. Reported Depression Diagnosis as an Adult Fitted Values from the GAM Procedure



CHAPTER 4: DISCUSSION

This study aimed to develop a forecasting tool that can be used to identify youth at risk of being diagnosed with depression as an adult. Additionally, this study investigated the developmental trajectories of depression for youth. The following section discusses the findings of this study within the context of prior research and real-world applicability. Implications of this study and limitations are also discussed.

Depression Forecasting Tool

This study is a preliminary step towards the integration of technology solutions into the treatment of behavioral health conditions. This study demonstrated the feasibility of developing a forecasting procedure that can be used as a tool for identifying youth who are at risk of being diagnosed with depression as an adult. Using a set of input variables collected from youth, this tool did a good job of forecasting which youth would not have a reported depression diagnosis as an adult with a 92% accuracy rate. This procedure was able to cut the error rate in half when classifying no diagnosis. Specifically, race, gender, youth depression score on the CES-D, receiving counseling as a youth, youth self-esteem, and youth suicidal ideation were the most important factors in terms of forecasting accuracy. If an algorithm like this one were replicated, these factors may be variables to consider including.

Whenever a method or idea that deviates from traditional approaches is proposed, providing a proof of concept to demonstrate practicality is an important first step. Therefore, the feasibility finding is important in demonstrating how technology-based approaches, such as machine learning algorithms, have the potential to improve the identification, assessment, and treatment of behavioral health conditions such as depression.

Despite major scientific advances in the United States, behavioral health difficulties remain a persistent problem for millions of Americans and many people never engage in treatment. According to Mental Health in America's 2019 State of Mental Health in America report, since last year there has been an increase in the percentage of individuals who report serious thoughts of suicide and an increase in the number of individuals who report experiencing at least one major depressive episode. Mainstream media has also recently drawn attention to the commonness of mental health difficulties after numerous celebrities and people in the public eye died by suicide. Now more than ever, we need to do better to ensure that people who are struggling get connected to the care that they need and deserve. While technology and machine learning strategies are not a silver bullet for behavioral health disorders, this study provides evidence for the potential of using a forecasting tool as a prevention mechanism and strategy to identify individuals who could benefit from receiving mental health services. Specifically, a tool like this one could help identify people who are likely to be diagnosed with depression in the future. Early identification is key to prevention and prior research has shown that intervening early rather than waiting for symptoms to further develop is beneficial (Ginsburg et al., 2014; Wolk, Kendall, & Beidas, 2015).

Despite knowing the importance of identifying behavioral health problems early, significant identification challenges exist, and many disorders often go undiagnosed (U.S. Department of Health & Human Services, 1999; Williams, Klinepeter, Palmes, Pulley, & Foy, 2004). A forecasting tool such as this one has the potential to help providers identify individuals at risk for depression and aligns with recent work being done by researchers at Virginia Tech where Chiu and colleagues are attempting to use Artificial Intelligence and machine learning algorithms to diagnose mental illness (Zarley, 2019). Compared to physical health conditions, where blood tests and X-rays can be

used to diagnose conditions, diagnosing behavioral health conditions is often much more subjective. The article highlights how challenging it is to quantify feelings and “measure the mind,” making it difficult to diagnosis mental illness using the DSM guidelines. This work provides hope that machine learning can positively impact our understanding and treatment of behavioral health conditions.

If we think about the process of treatment for behavioral health conditions, an individual often experiences or exhibits symptoms of a condition and either is referred to services or seeks services independently (identification). The individual is then engages in an assessment with a provider, which informs diagnosis and next steps. Next, the individual engages in services to treat their condition (treatment). While this description may be over-simplified, and in practice may not be so linear, the point is that identification is typically the first step. Early detection is key to prevention and identifying problems early reduces the chance of long-term disability associated with behavioral health problems (Williams, Klinepeter, Palmes, Pulley, & Foy, 2004). Hence, when thinking about prevention and early intervention strategies, a machine learning tool, such as this one could be used to identify people at risk of developing depression.

Interestingly, most recent work that has been done using machine learning in the behavioral health context has focused on assessment and treatment verses identification. For example, a 2018 article discusses how AI and technology are being used as solutions for individuals who do not have access to mental health services or cannot afford therapy (Garg & Glick, 2018). Specifically, technology solutions such as virtual therapy and chatbot counseling are discussed. Given how behavioral health is typically discussed and treated though, it is not surprising that there has not been much emphasis on prevention and early detection. This study provides an example of how a

forecasting tool can be used as a prevention strategy and provides preliminary evidence that can be used to inform future work and research.

In partnership with healthcare providers, a tool like this one could be used in existing healthcare settings. For example, a pediatric primary care center could be a great fit given that most youth visit a primary care doctor at least once annually. While research suggests that two out of three youth struggling with depression are not identified by primary care clinicians and do not receive care or treatment (Burns et al., 1995; Leaf et al., 1996), other research shows a positive impact when youth engage in screening. Wissow et al. (2013) found that universal mental health screening in primary care increased referral rates for evaluation and treatment and improved communication about mental health difficulties between providers, parents, and youth.

The American Academy of Pediatrics updated their clinical practice guidelines in 2018 and endorses annual universal depression screening for youth who are 12 and older during regular well-visits (Zuckerbrot et al., 2018). Additionally, Zuckerbrot et al. (2018) recommend that youth who experience high risk for depression be identified. A previous history or family history of mental health difficulties, psychosocial stressors, trauma history, and somatic symptoms are considered risk factors for future episodes of depression (Zuckerbrot et al., 2018). A recent study by Leslie and Chike-Harris (2018) found that screening youth for depression during well-visits and sick-visits led to increases in the number of youth who were identified and diagnosed with depression.

These practice guidelines, in combination with the changing payment landscape of the U.S. healthcare system, call for the exploration of new and innovative solutions to be able to identify and provide services with individuals who experience “high risk.” Despite any changes in federal healthcare policy that may emerge, the shift from fee-for-service to value-based payment systems is most likely here to stay for some time

(Langer, Antonelli, Chamberlain, Pan, & Keller, 2018). In a value-based healthcare system, providers are paid based on patient outcomes and there is an emphasis on population health management. This study's demonstration of the feasibility of using a forecasting tool to identify individuals at risk of being diagnosed with depression aligns with this emphasis.

Depression Trajectories

The second aim of this study was to advance the understanding of developmental trajectories of depression for youth. Specifically, using longitudinal data, this study explored what differentiates youth with symptoms of depression who go on to report a depression diagnosis as an adult from youth with symptoms of depression who do not go on to report a depression diagnosis as adult. The main finding from this study related to youth depression trajectories was that race and gender were the most important factors in terms of who would have a reported depression diagnosis as an adult. Longitudinal data were used to examine these trajectories from youth to adulthood. Smokowkoski and team (2014) highlight how even though developmental mental health research is about trajectories and change over time, most research in the area is cross-sectional rather than longitudinal.

This study found that the factors which most influenced whether youth would have a reported depression diagnosis as an adult the most were if the youth identified as White, female, and had ever received counseling as a youth. From a prevention standpoint, this finding is not overly useful, but is consistent with previous research about individuals who get diagnosed with depression most often. Specifically, research consistently shows that women are almost twice as likely to experience and be diagnosed with depression compared to men (Mayo Clinic, 2019; Whiteman, Ruggiano,

& Thomlison, 2016). Furthermore, behavioral health conditions such as depression are often underdiagnosed and under-treated among people who identify as Black or African American and Hispanic/Latino/a compared to people who identify as White (Stockdale, Lagomasino, Siddique, McGuire, & Miranda, 2008; Young, Klap, Sherbourne, & Wells, 2001).

While this finding is not novel and does not provide new insight as to understanding why, among a group of young people with depression symptoms, some report a depression diagnosis as an adult and others do not, it does perhaps highlight an important and larger issue of health disparities and who has access to health services. The National Institutes of Health (2014) defines health disparities as, “differences that exist among specific population groups in the United States in the attainment of full health potential that can be measured by differences in incidence, prevalence, mortality, burden of disease, and other adverse health conditions.”

Despite design or methodology, research has consistently found that individuals who identify as White are healthier than people who identify with almost all other racial groups (except individuals who identify as Asian; National Center for Health Statistics, 2016). Research also shows that while these disparities exist in various areas of health including life expectancy, heart disease, infant mortality, and obesity, behavioral health disparities also exist (Baciu et al., 2017; Safran et al., 2009). Individuals, particularly people who identify as Black/African American, are less likely to ask questions with their healthcare providers and less likely to request information about their health (Patel & Bakken, 2011; Eliacin et al., 2016).

While rates of mental health conditions are similar across different ethnic/racial groups, the consequences of these conditions are often worse for individuals who do not identify as White (APA, 2017). Additionally, individuals who do not identify as White are

less likely to receive behavioral health services. A 2015 report found that among individuals with any mental health conditions, 48% of people who identify as White received treatment, 31% of people who identify as Black or African American received treatment, 31% of people who identify as Hispanic received treatment, and 22% of people who identify as Asian received treatment (AHRQ, 2016). Recent work has also shown that the mental health of Black/African American youth needs more attention as the suicide rate for Black/African American youth is increasing compared to suicide rates for other children of the same age. A 2015 study showed that suicide rates were twice as high for Black/African American youth compared to White youth ages five to eleven (Bridge et al., 2015). While people across different racial/ethnic groups are less likely to seek treatment compared to their White counterparts, research also suggests that once they enter treatment, they are more likely to end treatment early (Cook et al., 2015; Fortuna et al., 2010).

While many factors may explain racial disparities in health care, within the context of this study, a factor that is important to consider involves differences in trust or distrust in healthcare providers. The degree to which an individual seeks out medical care and health services, retains long term relationships with healthcare providers, and adheres to treatment is greatly influenced by the level of trust and therapeutic relationship between the individual and provider (Boulware, Cooper, Ratner, LaVeist, & Powe, 2016; Hall, Dugan, Zheng, & Mishra, 2001; Peterson, 2002). Healthcare providers' cultural awareness in practice, as well as perceived racial bias and levels of empathy have also been associated with contributing factors for not using health services (Constatine, 2007; Cooper et al., 2012; Thomspson & McCable, 2012). Gender also plays a role in help-seeking behaviors. On average, women are more likely to use behavioral health care services than men (Matheson et al., 2014; SAMSHA, 2015).

In addition, when thinking about help-seeking behaviors and perceived need for mental health services, it is important to consider how a person's race/ethnicity and gender may impact decisions to acknowledge and enter behavioral health services. For example, cultural differences between groups of people may influence what someone considers a mental health difficulty that requires treatment versus usual day-to-day stress (Ault-Brutus & Alegria, 2018). Race and gender have consistently been linked with mental health service use. Women use mental health services at higher rates than men and people who identify as White use mental health services at higher rates than people who identify as African American/Black, Asian, or Hispanic/Latino/a (Kessler et al., 2005; Narendorf et al., 2018; Wang et al., 2005). Among a group of women with depression, White women were more likely than women from other racial/ethnic groups to think they needed behavioral health treatment (Nadeem, Lange, & Jeanne, 2009). Ault-Brutus and Alegria (2018) also suggest that individuals' social network may influence why perceived need may vary across different racial/ethnic groups. For example, White individuals may be more likely to perceive that they need treatment because they have been more exposed to mental health conditions and treatment through their social network (Nadeem, Lange, & Jeanne, 2009).

The role of stigma and shame associated with behavioral health conditions may also be important to consider when thinking about differences in help seeking related to gender and race/ethnicity. Prior research suggests that women consistently report more positive attitudes about seeking mental health treatment than men (Chandra & Minkovitz, 2007; Vogel, Heimerdinger-Edwards, Hammer, & Hubbard, 2011). Women may also be more likely than men to talk about mental health concerns or symptoms with friends or family. A recent study found that women often use their social connections to confide in whereas men use their social connections as a way to distract

themselves from their symptoms and struggles (Martínez-Hernández, Carceller-Maicas, DiGiacomo, & Ariste, 2016). Research also suggests that men are often hesitant to think of themselves as struggling with depression because they associate a depression diagnosis with a perceived threat to their masculinity (Seidler et al., 2016). A 2007 study found that men often feel higher levels of stigma related to seeking help for mental health problems, relative to women (Chandra & Minkovitz, 2007). Recent 2018 guidelines released by the American Psychological Association state that traditional ideologies about masculinity can have a negative effect on boys and men and the way that they express their emotions. Further, stigma associated with behavioral health difficulties and fear of how others would react to them receiving services is often greater for people identifying as African American/Black (Brown et al., 2010; Matthews, Corrigan, Smith, & Aranda, 2006).

Since Weissman's landmark article in the 1970s which noted differences in depression by gender, a significant amount of research has explored this disparity (Weissman & Klerman, 1977; Salk, Hyde, & Abramson, 2017). An array of factors and interactions of factors, including biological differences (e.g., hormonal, neurological, and genetic considerations) and psychosocial factors (e.g., socioeconomic resources, traumatic experiences, coping skills, and personality) have been found to influence gender differences in depression (Afifi, 2007; Salk, Hyde, & Abramson, 2017). Other research has proposed that the increased prevalence of depression among women is related to how women perceive and respond to stress (Kelly, Tyrka, Anderson, Price, & Carpenter, 2008). Compared to men, women are more likely to report experiencing greater anxiety and sadness from stress (Chaplin et al., 2008). Women are also more likely than men to experience trauma and experience negative consequences associated with stress (Chaplin et al., 2008; Keyes et al., 2012; Kucharska, 2017; Matud,

2004). Gender has been found to moderate the relationship between trauma and mental health symptoms, with a stronger association among women than men (Breslau & Anthony, 2007; Kucharska, 2017).

In addition to gender and race, having received counseling as a youth was also an important factor related to reporting a depression diagnosis as an adult. This finding could be attributed to the fact that if a person received counseling for a mental health concern as a youth, they may be more likely to seek services again if struggling with a mental health concern as an adult. This explanation aligns with prior research showing that prior positive experiences with mental health treatment predict future service engagement (Maulik, Eaton, & Bradshaw, 2011). However, this finding also highlights the importance of adolescent mental health treatment as a point of early intervention and prevention of depression in adulthood.

Despite scientific advancements in recent years, the quality of mental health treatment has not improved, and in some circumstances, has worsened (Hayes, Marston, Walters, King, & Osborn, 2017). This gap in treatment quality can be partially attributed to the lack of a systematic approach to measuring quality (Kilbourne et al., 2018). Further, the behavioral health field does not have an agreed upon set of quality indicators for psychosocial treatments (Pincus, Spaeth-Rublee, & Watkins, 2011).

Over 650 Evidence Based Treatments (EBTs) for various behavioral health concerns have been developed and tested in an effort to improve mental health treatment for youth (Chorpita et al., 2016). However, despite the abundance of EBTs, they are typically not delivered in community-based mental health clinics (Gyani, Shafran, Myles, & Rose, 2014; Zima et al., 2005). Research findings examining the effectiveness of youth mental health services delivered in community-based settings have also been mixed (Southam-Gerow et al, 2010; Weisz et al, 2012).

On average, less than half of individuals who report depression receive adequate treatment (Kessler et al., 2005). In fact, most people with depression receive treatment from primary care providers instead of mental health professional (Bilsker, Goldner, & Jones, 2007). Guidelines from the American Psychiatric Association suggest a person diagnosed with depression should receive treatment that includes antidepressant medication and/or psychotherapy for at least four to eight weeks. Studies have found that 30% to 79% of individuals in treatment for mood disorders such as depression receive treatment that does meet the threshold of minimally adequate care (Duhouze, Fournier, Gauvin, & Roberge, 2012; Eisenberg & Chung, 2012; Stein et al., 2013; Wang et al., 2005). Despite this large range found across studies, Puyat and colleagues (2016) note that this evidence highlights that many individuals with depression receive inadequate treatment. Additionally, their study found that men and younger adults had higher odds of receiving minimally adequate treatment relative to women and older adults (Puyat, Kazanjian, Golder, & Wong, 2016). Overall, the finding related to receiving counseling among youth and a reported depression diagnosis as an adult suggests that it would be helpful to further examine mental health counseling in adolescence, including access to evidence-supported treatments, outcomes, and implication for mental health in adulthood.

Bringing it all together, the finding that what differentiates youth with symptoms of depression who receive a reported depression diagnosis as an adult from youth with depression symptoms who do not receive a reported depression diagnosis as an adult are factors such as identifying as female, identifying as White, and having received counseling as a youth highlights the underlying issue of health disparities. Rather than helping to understand the developmental trajectories of depression, this finding may reflect the question of who is likely to seek services for a behavioral health condition

such as depression. In practice, to receive a depression diagnosis, an individual must go through a series of steps. First, a person must have a perceived or identified need; second, the person has to find and make an appointment with a provider; and finally, the person has to visit a healthcare provider for treatment. Given what we know about rates of diagnosis among different groups and differences in help-seeking behaviors, it makes sense that access to care and differences in who is likely to seek help in the first place is a plausible explanation for this finding. Past studies have found that individuals who identify as Black/African American or Latino/a are less likely to have access to quality care and treatment given the availability of providers where they live (Blanco et al. 2007; Hasnain-Wynia et al. 2007).

Connecting this finding to the first aim of the study, it demonstrates the importance of and need to identify individuals at risk of having a behavioral health condition in a universal and non-stigmatizing way. Together, these findings also highlight the importance of using a health promotion framework when talking about behavioral health. While most people would not second guess seeking care for a broken bone or another serious physical health concern, it would be ideal if this belief could also hold true for behavioral health symptoms. Health promotion which focuses on general well-being and keeping people healthy may be a helpful approach at reducing behavioral health difficulties (O'Connell, Boat, & Warner, 2009). For example, a 2019 population-based study found that regardless of demographic factors, poor mental health was consistently linked to poor diet and nutrition (Banta, Segovia-Siapco, Crocker, Montoya, & Alhusseini, 2019). The connection between physical health and mental health is important to keep in mind when thinking about prevention strategies.

Youth Characteristics

Although it was not a specific aim of the study, an important finding emerged regarding differences in youth characteristics among individuals who reported a depression diagnosis as adults that warrants some discussion. On average, adults with a reported depression diagnosis experienced several difficulties as youths than adults without a reported depression diagnosis across many domains. Specifically, adults with a reported depression diagnosis were more likely to have symptoms of depression as youth, reported being less happy and moody, had trouble sleeping, cried often, felt fearful, had thoughts of suicide, had attempted suicide, and knew someone who attempted or died by suicide as youth compared to adults without a depression diagnosis. Additionally, adults with a reported depression diagnosis reported lower levels of self-esteem, lower levels of family connection, and lower levels of support from adults as youth relative to adults without a reported depression diagnosis.

This finding highlights that differences exist between adults with a reported depression diagnosis and adults without a reported depression diagnosis when they are as young as middle schoolers and high schoolers. This finding is important because it means that it may be possible to identify these individuals sooner rather than later and intervene earlier in hopes of preventing symptoms from getting worse and increasing the chances that youth will grow up to be healthy adults. While the onset of depression often begins in adolescence, research shows that only half of youth with depression receive a diagnosis before they are adults (Kessler, Avenevoli, & Merikangas, 2001; Patel, Flisher, Getick, & McGorry, 2007). Additionally, when youth experience depression symptoms and they are not properly addressed, symptoms are likely to recur throughout their lives (Hammen, 2009).

This finding provides further support that early detection and screening for behavioral health conditions are necessary. Furthermore, as a healthcare system, more needs to be done to identify and connect people in need of care with the appropriate support and services. It is plausible that if youth who are at risk of being diagnosed with depression as adults are identified sooner and offered support, given tools, and receive prevention services, the chances of them struggling with depression as adults may be reduced and their likelihood of experiencing health in adulthood may be increased.

Limitations of the Research

Because this study is one of the first to use machine learning strategies within the context of the prevention of depression, it is exploratory by nature. Hence, there are a few limitations worth noting. First, this study relies on self-report data for all variables, and, therefore, responses may be subject to social desirability bias, or answering questions in ways seen as socially acceptable. Additionally, self-report surveys could be impacted by a respondent's mood that day and how the individual perceives and remembers past events or experiences. Second, the outcome and main variable of interest for both aims of this study was presence of a reported depression diagnosis as an adult. Specifically, the question asked, "Has a doctor, nurse or other health care provider ever told you that you have or had depression?" Therefore, the outcome variable is dependent on individuals accurately reporting whether they have received a depression diagnosis. As we know, depression can go underdiagnosed, so it is possible that some individuals may have experienced depression, but never sought help for it or were never diagnosed, and, therefore, responded no to this question. Thus, depression diagnosis in this study may be underreported. Third, as this study relies on secondary data from Wave I and Wave IV of the Add Health study, only variables collected in the

original study were available for this study. As such, some important variables related to the development of depression such as parental incarceration, parental mental health, and childhood abuse information were not available for this study. Fourth, the alpha values for some measures such as maternal attachment, maternal involvement, autonomy from parents, and neighborhood connection were low, which indicates that these measures have a questionable level of internal consistency. Finally, the results from this machine learning forecasting procedure may be different by race and gender. If a tool such as this one were to be used in practice, it would be important to explore potential differences in performance. Despite these limitations, this study contributes to research related to depression among youth and young adults and to the field of social policy and practice as it 1) provides support for the concept of using a machine learning forecasting tool to identify individuals with behavioral health conditions, such as depression, 2) offers insight into the development of a depression diagnosis for youth while emphasizing the role that health disparities and access to care play, and 3) highlights the importance of early detection and universal screening for behavioral health conditions.

Implications

The findings from this study have important implications for further research and practice. First, future research that addresses this study's limitations is needed. For example, rather than relying on individuals' self-report of a depression diagnosis, one could administer a depression assessment tool that is used to diagnose depression. It would be important to determine how forecasting skill and accuracy would compare when depression diagnosis is measured differently. Additionally, future research which replicates this study is also needed to validate the findings and accuracy of this study.

Specifically, results should be explored for differences in performance related to gender and race. Finally, while risk factors for developing behavioral health conditions have been studied in depth, less is known about protective factors that promote good health and well-being (Banyard, Hamby, & Grych, 2017). Hence, additional research is needed to address this gap in knowledge as understanding these distinguishing features is critical to informing prevention strategies and programs.

In addition to future research, the findings from this study have important practice implications. First, this study demonstrated that it is feasible to develop a forecasting tool that can be used to identify mental health difficulties. While this tool relied on a specific set of input variables and is likely not to be exactly replicated, a similar tool could be developed depending on the data and information one had available. A tool like this one could be implemented in various practice settings including a primary care clinic, behavioral health organization, or even at the behavioral healthcare system level. In the primary care setting, a tool like this one may be specifically helpful in identifying people with underlying behavioral health conditions and beginning conversations about the importance of mental health and how it also impacts our physical health. While this application may deviate from standard practice, with some training it could be feasible. For example, previous work supports the notion of screening for behavioral health conditions such as depression in the primary care setting (Leslie & Chike-Harris, 2018; Lewandowski et al., 2016). Integrating the use of a tool like this one into primary care well-visits could potentially lead to more people being diagnosed and connected to care, which is needed given that many individuals with a mental health difficulties never receive treatment. For example, every person who attends a well-visit would answer a series of questions related to depression. Using a combination of demographic and

depression screening information, people's data would be entered into the tool to see if they could benefit from being referred to behavioral health services.

Similarly, a behavioral health or social service organization could implement a similar tool using existing data that are collected from individuals as part of the standard intake process to identify individuals who are most at risk to ensure that they remain engaged in treatment or have access to services. Integrating a tool like this one in standard care and combining provider expertise/clinical judgement with data and/or a decision support tool has the potential to improve patient outcomes. Decision support tools take into account client information (e.g. demographics or clinical data) to offer personalized treatment (Graham, James, & Spertus, 2018). A recent study found that the use of decision support tools led to improvements in clinical practice and in the delivery of prevention services (Graham, James, & Spertus, 2018). For example, clinical decision support tools can provide individualized assessments and have been shown to reduce errors associated with medication, improve prescribing practices, and promote evidence-based care (Hunt, Haynes, Hanna, & Smith, 1998; Kawamoto, Houlihan, Balas, & Lobach, 2005). Additionally, a recent study conducted by Kaiser Permanente and the Mental Health Research Network found that prediction models, which included electronic health record data in addition to self-report depression data, could predict suicide risk following outpatient visits and outperform traditional suicide risk assessments (Simon et al., 2018).

At the system level, as many states are transitioning from a fee-for-service to value-based care payment models, health care providers and payers are now more than ever focused on data and being able to achieve positive patient health outcomes. An unintentional potential consequence of payment models such as these, which incentivize providers to produce positive outcomes, is the risk of providers not wanting to serve

patients experiencing high-risks and high-needs (Koning & Heinrich, 2013). Providers and payers can use forecasting tools to identify individuals at risk for chronic conditions and intervene before the condition worsens as long term health problems are often hard to treat and expensive (Bresnick, 2018). The Association of American Medical Colleges (AAMC) stated, “Across all [reimbursement] models, the identification, stratification, and management of high-risk patients is central to improving quality and cost outcomes. The use of predictive modeling to proactively identify patients who are at highest risk of poor health outcomes and will benefit most from intervention is one solution believed to improve risk management for providers transitioning to value-based payment” (p.5, p.97).

In sum, this study shows that it is possible to develop a forecasting tool that can be used to identify mental health difficulties. This finding is important as it demonstrates the feasibility and practicality of using innovative technology solutions to support prevention and intervention strategies within the context of behavioral health. Integrating tools like this one into standard practice of care has the potential to improve overall health and well-being.

Additionally, this study has important implications as it highlights that despite work that has been done to address health disparities, disparities are still prevalent and more needs to be done to ensure equal access to high quality healthcare services among all people. It also emphasizes the importance of being thoughtful about how mental health is discussed and presented and implies that mental health literacy efforts are needed as research suggests that knowledge about mental health difficulties and mental health literacy facilitates help-seeking behaviors (Eschenbeck et al., 2019). Mental health literacy, which is often a component of health promotion or prevention programs,

is defined by Jorm (2012) as “knowledge and beliefs about mental disorders which aid their recognition, management, or prevention.”

While there are numerous factors associated with barriers to care and the decision to seek mental health treatment, a systematic review found that stigma, embarrassment, and problems identifying mental health symptoms were the most influential barriers to seeking care (Gulliver, Griffiths, & Christensen, 2010). Hence, as providers and as a health system, more attention and effort need to be given to strategies that build mental health literacy and reduce the stigma associated with mental health conditions. Health promotion and prevention initiatives that focus on stress, well-being, and mental health literacy have demonstrated positive outcomes and greater level of social and emotional competencies (Fenwick-Smith, Dahlberg, & Thompson, 2018; O'Reilly, Svirydenka, Adams & Dogra, 2018). However, less is known about long-term impact of these initiatives, and more research is needed to further evaluate their effectiveness (O'Reilly, Svirydenka, Adams & Dogra, 2018).

As our healthcare system is working towards achieving the Triple Aim of improved patient experience of care, improved population health, and reduction in cost, investing in solutions to better identify people in need of behavioral health services and focusing on health promotion and prevention strategies have the potential to help achieve the Triple Aim and ensure that all individuals are equipped with the opportunity and tools to live a healthy life.

Conclusion

In sum, this dissertation highlights how a machine learning forecasting tool could be used to inform prevention strategies and factors associated with receiving a depression diagnosis. Findings from this study indicate that it is feasible and practical to use a forecasting tool to identify individuals at risk of being diagnosed with depression. Machine learning tools have the potential to improve the diagnosis and treatment of behavioral health conditions and subsequently may help individuals live healthier lives. Additionally, this dissertation emphasizes the role health disparities, specifically gender and race, may play in seeking care and having access to quality mental health treatment. Future research is needed to better understand the developmental trajectories of depression for youth and what differentiates youth with depression symptoms who are diagnosed with depression as an adult from youth with depression symptoms without a depression diagnosis as an adult. More attention and work focusing on health promotion and prevention should also be considered. This study presents and discusses these findings in addition to offering important implications for future research and practice to identify and prevent behavioral health conditions such as depression.

BIBLIOGRAPHY

- Abrutyn, S., & Mueller, A. S. (2014). Are suicidal behaviors contagious in adolescence? Using longitudinal data to examine suicide suggestion. *American Sociological Review*, 79(2), 211-227.
- Accurso, E. C., Garland, A. F., Haine-Schlagel, R., Brookman-Frazee, L., & Baker-Ericzén, M. J. (2015). Factors contributing to reduced caregiver strain in a publicly funded child mental health system. *Journal of Emotional and Behavioral Disorders*, 23(3), 131-143.
- Afifi, M. (2007). Gender differences in mental health. *Singapore Medical Journal*, 48(5), 385.
- Agency for Healthcare Research and Quality. (2016) 2015 National Healthcare Quality and Disparities Report.
- Ahn, W. Y., Ramesh, D., Moeller, F. G., & Vassileva, J. (2016). Utility of machine-learning approaches to identify behavioral markers for substance use disorders: impulsivity dimensions as predictors of current cocaine dependence. *Frontiers in Psychiatry*, 7, 34.
- Allison, P. D. (2001). *Missing data* (Vol. 136). Sage publications.
- Alonso, J., Petukhova, M., Vilagut, G., Chatterji, S., Heeringa, S., Üstün, T. B., ... & Bruffaerts, R. (2011). Days out of role due to common physical and mental conditions: results from the WHO World Mental Health surveys. *Molecular Psychiatry*, 16(12), 1234-1246.
- Amarasingham, R., Patzer, R. E., Huesch, M., Nguyen, N. Q., & Xie, B. (2014). Implementing electronic health care predictive analytics: considerations and challenges. *Health Affairs*, 33(7), 1148-1154.
- American Psychiatric Association. (2013). Diagnostic and statistical manual of mental

- disorders (5th ed.). Arlington, VA: Author.
- American Psychological Association (2016). Stress in America: The Impact of Discrimination. Stress in America™ Survey. American Psychological Association: Washington DC.
- Angold, A., Erkanli, A., Farmer, E. M., Fairbank, J. A., Burns, B. J., Keeler, G., & Costello, E. J. (2002). Psychiatric disorder, impairment, and service use in rural African American and white youth. *Archives of General Psychiatry*, 59(10), 893-901.
- Armbruster, P., & Fallon, T. (1994). Clinical, sociodemographic, and systems risk factors for attrition in a children's mental health clinic. *American Journal of Orthopsychiatry*, 64(4), 577.
- Ault-Brutus, A., & Alegria, M. (2018). Racial/ethnic differences in perceived need for mental health care and disparities in use of care among those with perceived need in 1990–1992 and 2001–2003. *Ethnicity & Health*, 23(2), 142-157.
- Austin, P. C. (2007). A comparison of regression trees, logistic regression, generalized additive models, and multivariate adaptive regression splines for predicting AMI mortality. *Statistics in Medicine*, 26(15), 2937-2957.
- Baciu, A., Negussie, Y., Geller, A., Weinstein, J. N., & National Academies of Sciences, Engineering, and Medicine. (2017). The state of health disparities in the United States. In *Communities in Action: Pathways to Health Equity*. National Academies Press (US).
- Badger, L., Robinson, H., & Farley, T. (1999). Management of mental disorders in rural primary care a proposal for integrated psychosocial services. *Journal of Family Practice*, 48(10), 813-813.
- Bannon Jr, W. M., & McKay, M. M. (2005). Are barriers to service and parental

- preference match for service related to urban child mental health service use?. *Families in Society*, 86(1), 30-34.
- Banta, J.E., Segovia-Siapco, G., Crocker, C.B., Montoya, D., & Alhusseini, N. (2019) Mental health status and dietary intake among California adults: A population-based survey. *International Journal of Food Sciences and Nutrition*.
- Banyard, V., Hamby, S., & Grych, J. (2017). Health effects of adverse childhood events: Identifying promising protective factors at the intersection of mental and physical well-being. *Child Abuse & Neglect*, 65, 88-98.
- Barnes, J. C., & Morris, R. G. (2012). Young mothers, delinquent children: Assessing mediating factors among American youth. *Youth Violence and Juvenile Justice*, 10(2), 172-189.
- Barocas, S., & Selbst, A. D. (2016). Big data's disparate impact. *Calif. L. Rev.*, 104, 671.
- Barry, M. M., & Jenkins, R. (2007). *Implementing mental health promotion*. Elsevier Health Sciences.
- Barry, M. M., Clarke, A. M., Jenkins, R., & Patel, V. (2013). A systematic review of the effectiveness of mental health promotion interventions for young people in low and middle income countries. *BMC Public Health*, 13(1), 835.
- Barry-Jester A.M., Casselman, B. and Goldstein, D. (2015) Should Prison Sentences Be Based on Crimes That Haven't Been Committed Yet? [on-line] Available from: <http://fivethirtyeight.com/features/prison-reform-risk-assessment/> [11.07.16].
- Bazaco, M. C., Pereira, M. A., Wisniewski, S. R., Zgibor, J. C., Songer, T. J., Burke, J. D., & Fabio, A. (2016). Is there a relationship between perceived neighborhood contentedness and physical activity in young men and women. *Journal of Urban Health*, 93(6), 940-952.
- Beardslee, W. R., Gladstone, T. R., Wright, E. J., & Cooper, A. B. (2003). A family-based

- approach to the prevention of depressive symptoms in children at risk: evidence of parental and child change. *Pediatrics*, 112(2), e119-e131.
- Beardselee, W. R., Versage, E. M., & Gladstone, T. R. (1998). Children of affectively ill parents: a review of the past 10 years. *Journal of the American Academy of Child & Adolescent Psychiatry*, 37(11), 1134-1141.
- Beaver, K. M. (2008). Nonshared environmental influences on adolescent delinquent involvement and adult criminal behavior. *Criminology*, 46(2), 341-369.
- Beaver, K. M., Schwartz, J. A., Connolly, E. J., Al-Ghamdi, M. S., & Kobeisy, A. N. (2015). The role of parenting in the prediction of criminal involvement: Findings from a nationally representative sample of youth and a sample of adopted youth. *Developmental Psychology*, 51(3), 301.
- Becker, K. D., Lee, B. R., Daleiden, E. L., Lindsey, M., Brandt, N. E., & Chorpita, B. F. (2015). The common elements of engagement in children's mental health services: Which elements for which outcomes?. *Journal of Clinical Child & Adolescent Psychology*, 44(1), 30-43.
- Bennett, D. A. (2001). How can I deal with missing data in my study? *Australian and New Zealand Journal of Public Health*, 25, 464-469.
- Berk, R. A., Kriegler, B., & Baek, J. H. (2006). Forecasting dangerous inmate misconduct: An application of ensemble statistical procedures. *Journal of Quantitative Criminology*, 22(2), 131-145.
- Berk, R., Sherman, L., Barnes, G., Kurtz, E., & Ahlman, L. (2009). Forecasting murder within a population of probationers and parolees: a high stakes application of statistical learning. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 172(1), 191-211.
- Berk, R., & Bleich, J. (2014). Forecasts of violence to inform sentencing

- decisions. *Journal of Quantitative Criminology*, 30(1), 79-96.
- Berk, R., & Hyatt, J. (2014). Machine learning forecasts of risk to inform sentencing decisions. *Fed. Sent'g Rep.*, 27, 222.
- Berk, R. A., Sorenson, S. B., & Barnes, G. (2016). Forecasting domestic violence: A machine learning approach to help inform arraignment decisions. *Journal of Empirical Legal Studies*, 13(1), 94-115.
- Berk, R. A. (2016). *Statistical learning from a regression perspective* (Second Edition). New York: Springer.
- Berk, R., Brown, L., Buja, A., George, E., & Zhao, L. (2018). Working with misspecified regression models. *Journal of Quantitative Criminology*, 34(3), 633-655.
- Bernstein, D. P., Stein, J. A., Newcomb, M. D., Walker, E., Pogge, D., Ahluvalia, T., ... & Zule, W. (2003). Development and validation of a brief screening version of the Childhood Trauma Questionnaire. *Child Abuse & Neglect*, 27(2), 169-190.
- Beronio, K., Glied, S., & Frank, R. (2014). How the Affordable Care Act and Mental Health Parity and Addiction Equity Act greatly expand coverage of behavioral health care. *The Journal of Behavioral Health Services & Research*, 41(4), 410-428.
- Berwick, D. M., Nolan, T. W., & Whittington, J. (2008). The triple aim: care, health, and cost. *Health Affairs*, 27(3), 759-769.
- Billings, J., Dixon, J., Mijanovich, T., & Wennberg, D. (2006). Case finding for patients at risk of readmission to hospital: development of algorithm to identify high risk patients. *BMJ*, 333(7563), 327.
- Billings, J., Blunt, I., Steventon, A., Georghiou, T., Lewis, G., & Bardsley, M. (2012). Development of a predictive model to identify inpatients at risk of re-admission within 30 days of discharge (PARR-30). *BMJ Ppen*, 2(4), e001667.

- Bilsker, D., Goldner, E. M., & Jones, W. (2007). Health service patterns indicate potential benefit of supported self-management for depression in primary care. *The Canadian Journal of Psychiatry*, 52(2), 86-95.
- Blanco, C., Patel, S. R., Liu, L., Jiang, H., Lewis-Fernández, R., Schmidt, A. B., ... & Olfson, M. (2007). National trends in ethnic disparities in mental health care. *Medical Care*, 45(11), 1012-1019.
- Boardman, J. D., & Alexander, K. B. (2011). Stress trajectories, health behaviors, and the mental health of black and white young adults. *Social Science & Medicine*, 72(10), 1659-1666.
- Bone, D., Goodwin, M. S., Black, M. P., Lee, C. C., Audhkhasi, K., & Narayanan, S. (2015). Applying machine learning to facilitate autism diagnostics: pitfalls and promises. *Journal of Autism and Developmental Disorders*, 45(5), 1121-1136.
- Bonde, J. P. E. (2008). Psychosocial factors at work and risk of depression: a systematic review of the epidemiological evidence. *Occupational and Environmental Medicine*, 65(7), 438-445.
- Boulware, L. E., Cooper, L. A., Ratner, L. E., LaVeist, T. A., & Powe, N. R. (2016). Race and trust in the health care system. *Public Health Reports*.
- Breiman, L. (2001). Random forests. *Machine Learning*, 45(1), 5-32.
- Breslau, N., & Anthony, J. C. (2007). Gender differences in the sensitivity to posttraumatic stress disorder: An epidemiological study of urban young adults. *Journal of Abnormal Psychology*, 116(3), 607.
- Bresnick, J. (2018, September 04). 10 High-Value Use Cases for Predictive Analytics in Healthcare. HealthITAnalytics. Retrieved from <https://healthitanalytics.com/news/10-high-value-use-cases-for-predictive-analytics-in-healthcare>

- Bridge, J. A., Asti, L., Horowitz, L. M., Greenhouse, J. B., Fontanella, C. A., Sheftall, A. H., ... & Campo, J. V. (2015). Suicide trends among elementary school-aged children in the United States from 1993 to 2012. *JAMA Pediatrics*, 169(7), 673-677.
- Brown, L. K., Tolou-Shams, M., Lescano, C., Houck, C., Zeidman, J., Pugatch, D., ... & Project SHIELD Study Group. (2006). Depressive symptoms as a predictor of sexual risk among African American adolescents and young adults. *Journal of Adolescent Health*, 39(3), 444-e1.
- Brown, C., Conner, K. O., Copeland, V. C., Grote, N., Beach, S., Battista, D., & Reynolds III, C. F. (2010). Depression stigma, race, and treatment seeking behavior and attitudes. *Journal of Community Psychology*, 38(3), 350-368.
- Buckner, J. C., & Bassuk, E. L. (1997). Mental disorders and service utilization among youths from homeless and low-income housed families. *Journal of the American Academy of Child & Adolescent Psychiatry*, 36(7), 890-900.
- Burns, B. J., Costello, E. J., Angold, A., Tweed, D., Stangl, D., Farmer, E. M., & Erkanli, A. (1995). Children's mental health service use across service sectors. *Health Affairs*, 14(3), 147-159.
- Bussing, R., Zima, B. T., Gary, F. A., Mason, D. M., Leon, C. E., Sinha, K., & Garvan, C. W. (2003). Social networks, caregiver strain, and utilization of mental health services among elementary school students at high risk for ADHD. *Journal of the American Academy of Child & Adolescent Psychiatry*, 42(7), 842-850.
- Carpenter, K. L., Sprechmann, P., Calderbank, R., Sapiro, G., & Egger, H. L. (2016). Quantifying risk for anxiety disorders in preschool children: A machine learning approach. *PloS one*, 11(11), e0165524.
- Carlson, E. B., & Dalenberg, C. J. (2000). A conceptual framework for the impact of

- traumatic experiences. *Trauma, Violence, & Abuse*, 1(1), 4-28.
- Carr, C. P., Martins, C. M. S., Stingel, A. M., Lemgruber, V. B., & Juruena, M. F. (2013). The role of early life stress in adult psychiatric disorders: a systematic review according to childhood trauma subtypes. *The Journal of Nervous and Mental Disease*, 201(12), 1007-1020.
- Chandra, A., & Minkovitz, C. S. (2007). Factors that influence mental health stigma among 8th grade adolescents. *Journal of Youth and Adolescence*, 36(6), 763-774.
- Chaplin, T. M., Hong, K., Bergquist, K., & Sinha, R. (2008). Gender differences in response to emotional stress: an assessment across subjective, behavioral, and physiological domains and relations to alcohol craving. *Alcoholism: Clinical and Experimental Research*, 32(7), 1242-1250.
- Char, D. S., Shah, N. H., & Magnus, D. (2018). Implementing machine learning in health care—addressing ethical challenges. *The New England Journal of Medicine*, 378(11), 981.
- Charlson, F. J., Moran, A. E., Freedman, G., Norman, R. E., Stapelberg, N. J., Baxter, A. J., ... & Whiteford, H. A. (2013). The contribution of major depression to the global burden of ischemic heart disease: a comparative risk assessment. *BMC Medicine*, 11(1), 250.
- Chen, P., & Chantala, K. (2014). Guidelines for analyzing Add Health data. *Carol. Popul. Cent. Univ. NC Chap. Hill*, 1-53.
- Chen, W. Y., Corvo, K., Lee, Y., & Hahm, H. C. (2017). Longitudinal trajectory of adolescent exposure to community violence and depressive symptoms among adolescents and young adults: understanding the effect of mental health service usage. *Community Mental Health Journal*, 53(1), 39-52.

- Cheng, T. C., & Lo, C. C. (2017). Social Risk and Protective Factors in Adolescents' Reduction and Cessation of Alcohol Use. *Substance Use & Misuse*, 52(7), 916-928.
- Chorpita, B. F., Bernstein, A. D., Becker, K. D., Ebesutani, C., Park, A. L., Brown, T., Tsai, K., Moskowitz, A. L., & Daleiden, E. L. (2016). Measurement feedback and beyond: The place for evidence in clinical decision support. In J. Persons (Chair), Monitoring progress in psychotherapy: Why and how. Symposium presented at the 50th Annual Convention of the Association for Behavioral and Cognitive Therapies. New York, NY.
- Clement, S., Schauman, O., Graham, T., Maggioni, F., Evans-Lacko, S., Bezborodovs, N., ... & Thornicroft, G. (2015). What is the impact of mental health-related stigma on help-seeking? A systematic review of quantitative and qualitative studies. *Psychological Medicine*, 45(01), 11-27.
- Clements, M. S., Armstrong, B. K., & Moolgavkar, S. H. (2005). Lung cancer rate predictions using generalized additive models. *Biostatistics*, 6(4), 576-589.
- Collins, C., Hewson, D., Munger, R., & Wade, T. (2010). Evolving models of behavioral health integration in primary care. New York: Milbank Memorial Fund, Retrieved from <http://www.milbank.org/uploads/documents/10430EvolvingCare/EvolvingCare.pdf>
- Collishaw, S., Hammerton, G., Mahedy, L., Sellers, R., Owen, M. J., Craddock, N., ... & Thapar, A. (2016). Mental health resilience in the adolescent offspring of parents with depression: a prospective longitudinal study. *The Lancet Psychiatry*, 3(1), 49-57.
- Conley, C. S., Durlak, J. A., & Dickson, D. A. (2013). An evaluative review of outcome

- research on universal mental health promotion and prevention programs for higher education students. *Journal of American College Health*, 61(5), 286-301.
- Constantine, M. G. (2007). Racial microaggressions against African American clients in cross-racial counseling relationships. *Journal of Counseling Psychology*, 54(1), 1.
- Cook, T.B., Reeves, G.M., Teufel, J., & Postolache, T.T. (2015). Persistence of racial disparities in prescription of first-generation antipsychotics in the USA. *Pharmacoepidemiology and Drug Safety*, 24(11), 1197–1206.
- Cooper, J. L., & Aratani, Y. (2009). The status of states' policies to support evidence-based practices in children's mental health. *Psychiatric Services*, 60(12), 1672-1675.
- Cooper, L. A., Roter, D. L., Carson, K. A., Beach, M. C., Sabin, J. A., Greenwald, A. G., & Inui, T. S. (2012). The associations of clinicians' implicit attitudes about race with medical visit communication and patient ratings of interpersonal care. *American Journal of Public Health*, 102(5), 979-987.
- Copeland, W. E., Keeler, G., Angold, A., & Costello, E. J. (2007). Traumatic events and posttraumatic stress in childhood. *Archives of General Psychiatry*, 64(5), 577-584.
- Costello, E. J., Egger, H., & Angold, A. (2005). 10-year research update review: the epidemiology of child and adolescent psychiatric disorders: I. Methods and public health burden. *Journal of the American Academy of Child & Adolescent Psychiatry*, 44(10), 972-986.
- Costello, E. J., Copeland, W., & Angold, A. (2011). Trends in psychopathology across the adolescent years: what changes when children become adolescents, and when adolescents become adults?. *Journal of Child Psychology and Psychiatry*, 52(10), 1015-1025.

- Crosnoe, R., & Elder Jr, G. H. (2004). Family dynamics, supportive relationships, and educational resilience during adolescence. *Journal of Family Issues*, 25(5), 571-602.
- Culhane, D. P. (2016). The potential of linked administrative data for advancing homelessness research and policy. *European Journal of Homelessness*, 10(3), 109-126.
- Curran, E., Adamson, G., Stringer, M., Rosato, M., & Leavey, G. (2016). Severity of mental illness as a result of multiple childhood adversities: US National Epidemiologic Survey. *Social Psychiatry and Psychiatric Epidemiology*, 1-11.
- Davenport, T., & Harris, J. (2017). *Competing on Analytics: Updated, with a New Introduction: The New Science of Winning*. Harvard Business Press.
- DeWit, D. J., DuBois, D., Erdem, G., Larose, S., & Lipman, E. L. (2016). The Role of program-supported mentoring relationships in promoting youth mental health, behavioral and developmental outcomes. *Prevention Science*, 17(5), 646-657.
- Diab, M., Peltonen, K., Qouta, S. R., Palosaari, E., & Punamäki, R. L. (2015). Effectiveness of psychosocial intervention enhancing resilience among war-affected children and the moderating role of family factors. *Child Abuse & Neglect*, 40, 24-35.
- Diener, E., & Seligman, M. E. (2002). Very happy people. *Psychological Science*, 13(1), 81-84.
- Dong, Y., & Peng, C. Y. J. (2013). Principled missing data methods for researchers. *SpringerPlus*, 2(1), 222.
- Dore, M. M., Wilkinson, A. N., & Sonis, W. A. (1992). Exploring the relationship between a continuum of care and intrusiveness of children's mental health services. *Psychiatric Services*, 43(1), 44-48.

- Driscoll, A. K., Russell, S. T., & Crockett, L. J. (2008). Parenting styles and youth well-being across immigrant generations. *Journal of Family Issues*, 29(2), 185-209.
- DuBois, D. L., Portillo, N., Rhodes, J. E., Silverthorn, N., & Valentine, J. C. (2011). How effective are mentoring programs for youth? A systematic assessment of the evidence. *Psychological Science in the Public Interest*, 12(2), 57-91.
- Duhoux, A., Fournier, L., Gauvin, L., & Roberge, P. (2012). Quality of care for major depression and its determinants: a multilevel analysis. *BMC Psychiatry*, 12(1), 142.
- Durlak, J. A., Weissberg, R. P., Dymnicki, A. B., Taylor, R. D., & Schellinger, K. B. (2011). The impact of enhancing students' social and emotional learning: A meta-analysis of school-based universal interventions. *Child Development*, 82(1), 405-432.
- Easterbrooks, M. A., Ginsburg, K., & Lerner, R. M. (2013). Resilience among military youth. *The Future of Children*, 23(2), 99-120.
- Eckerson, W. W. (2007). Predictive analytics. *Extending the Value of Your Data Warehousing Investment. TDWI Best Practices Report*, 1, 1-36.
- Eisenberg, D., & Chung, H. (2012). Adequacy of depression treatment among college students in the United States. *General Hospital Psychiatry*, 34(3), 213-220.
- Eliacin, J., Rollins, A.L., Burgess, D.J., Salyers, M.P., & Matthias, M.S. (2016). Patient activation and visit preparation in African American veterans receiving mental health care. *Cultural Diversity and Ethnic Minority Psychology*, 22(4), 580–587.
- Elkington, K. S., Bauermeister, J. A., & Zimmerman, M. A. (2010). Psychological distress, substance use, and HIV/STI risk behaviors among youth. *Journal of Youth and Adolescence*, 39(5), 514-527.

- English, D. J., Marshall, D. B., & Stewart, A. J. (2003). Effects of family violence on child behavior and health during early childhood. *Journal of Family Violence*, 18(1), 43-57.
- Engstrom, M., El-Bassel, N., & Gilbert, L. (2012). Childhood sexual abuse characteristics, intimate partner violence exposure, and psychological distress among women in methadone treatment. *Journal of Substance Abuse Treatment*, 43(3), 366-376.
- Eschenbeck, H., Lehner, L., Hofmann, H., Bauer, S., Becker, K., Diestelkamp, S., ... & Salize, H. J. (2019). School-based mental health promotion in children and adolescents with StresSOS using online or face-to-face interventions: study protocol for a randomized controlled trial within the ProHEAD Consortium. *Trials*, 20(1), 64.
- Esposito, M. H., Lee, H., Hicken, M. T., Porter, L. C., & Herting, J. R. (2017). The consequences of contact with the criminal justice system for health in the transition to adulthood. *Longitudinal and Life Course Studies*, 8(1), 57.
- Executive Office of the President, Munoz, C., Director, D. P. C., Megan (US Chief Technology Officer Smith (Office of Science and Technology Policy)), & DJ (Deputy Chief Technology Officer for Data Policy and Chief Data Scientist Patil (Office of Science and Technology Policy)). (2016). Big data: A report on algorithmic systems, opportunity, and civil rights. Executive Office of the President.
- Fenwick-Smith, A., Dahlberg, E. E., & Thompson, S. C. (2018). Systematic review of resilience-enhancing, universal, primary school-based mental health promotion programs. *BMC Psychology*, 6(1), 30.
- Fiorilli, C., Capitello, T. G., Barni, D., Buonomo, I., & Gentile, S. (2019). Predicting

- adolescent depression: The interrelated roles of self-esteem and interpersonal stressors. *Frontiers in Psychology*, 10.
- Fishbein, D. H., Ridenour, T. A., Stahl, M., & Sussman, S. (2016). The full translational spectrum of prevention science: facilitating the transfer of knowledge to practices and policies that prevent behavioral health problems. *Translational Behavioral Medicine*, 6(1), 5-16.
- Fletcher, J. M. (2009). Childhood mistreatment and adolescent and young adult depression. *Social Science & Medicine*, 68(5), 799-806.
- Fothergill, A., & Peek, L. (2015). *Children of Katrina*. University of Texas Press.
- Fortuna, L.R., Alegria, M., & Gao, S. (2010). Retention in depression treatment among ethnic and racial minority groups in the United States. *Depression and Anxiety*, 27(5), 485–494.
- Fossum, S., Mørch, W., Handegård, B. H., & Drugli, M. B. (2007). Childhood disruptive behaviors and family functioning in clinically referred children: Are girls different from boys? *Scandinavian Journal of Psychology*, 48(5), 375-382.
- Franco, L., J Pottick, K., & Huang, C. C. (2010). Early parenthood in a community context: Neighborhood conditions, race–ethnicity, and parenting stress. *Journal of Community Psychology*, 38(5), 574-590.
- Frank, R. G., Beronio, K., & Glied, S. A. (2014). Behavioral health parity and the Affordable Care Act. *Journal of Social Work in Disability & Rehabilitation*, 13(1-2), 31-43.
- Frieden, T. R. (2010). A framework for public health action: the health impact pyramid. *American Journal of Public Health*, 100(4), 590-595.
- Galatzer-Levy, I. R., Karstoft, K. I., Statnikov, A., & Shalev, A. Y. (2014). Quantitative

- forecasting of PTSD from early trauma responses: A machine learning application. *Journal of Psychiatric Research*, 59, 68-76.
- Garg, P. & Glick, S. (2018, October 24). AI's Potential to Diagnose and Treat Mental Illness. Retrieved from <https://hbr.org/2018/10/ais-potential-to-diagnose-and-treat-mental-illness>
- George, C., Herman, K. C., & Ostrander, R. (2006). The family environment and developmental psychopathology: The unique and interactive effects of depression, attention, and conduct problems. *Child Psychiatry and Human Development*, 37(2), 163-177.
- Gillingham, P. (2015). Predictive risk modelling to prevent child maltreatment and other adverse outcomes for service users: Inside the 'black box' of machine learning. *British Journal of Social Work*, bcv031.
- Ginsburg, G. S., Becker, E. M., Keeton, C. P., Sakolsky, D., Piacentini, J., Albano, A. M., ... & Peris, T. (2014). Naturalistic follow-up of youths treated for pediatric anxiety disorders. *JAMA psychiatry*, 71(3), 310-318.
- Gopalan, G., Goldstein, L., Klingenstein, K., Sicher, C., Blake, C., & McKay, M. M. (2010). Engaging families into child mental health treatment: updates and special considerations. *Journal of the Canadian Academy of Child & Adolescent Psychiatry*, 19(3).
- Gould, M., Lake, A., Kleinman, M., Galfalvy, H., Chowdhury, S., & Madnick, A. (2018). Exposure to suicide in high schools: Impact on serious suicidal ideation/behavior, depression, maladaptive coping strategies, and attitudes toward help-seeking. *International Journal of Environmental Research and Public Health*, 15(3), 455.
- Graham, M. M., James, M. T., & Spertus, J. A. (2018). Decision support tools: realizing the potential to improve quality of care. *Canadian Journal of Cardiology*.

- Greeff, A. P., Vansteenwegen, A., & Ide, M. (2006). Resiliency in families with a member with a psychological disorder. *The American Journal of Family Therapy*, 34(4), 285-300.
- Greene, D., Hoffmann, A. L., & Stark, L. (2019, January). Better, Nicer, Clearer, Fairer: A Critical Assessment of the Movement for Ethical Artificial Intelligence and Machine Learning. In Proceedings of the 52nd Hawaii International Conference on System Sciences.
- Green, J. G., McLaughlin, K. A., Berglund, P. A., Gruber, M. J., Sampson, N. A., Zaslavsky, A. M., & Kessler, R. C. (2010). Childhood adversities and adult psychiatric disorders in the national comorbidity survey replication I: associations with first onset of DSM-IV disorders. *Archives of General Psychiatry*, 67(2), 113-123.
- Gulliver, A., Griffiths, K. M., & Christensen, H. (2010). Perceived barriers and facilitators to mental health help-seeking in young people: a systematic review. *BMC Psychiatry*, 10(1), 113.
- Guo, Y., Wei, Z., Keating, B. J., & Hakonarson, H. (2015). Machine learning derived risk prediction of anorexia nervosa. *BMC medical genomics*, 9(1), 4.
- Gyani, A., Shafran, R., Myles, P., & Rose, S. (2014). The gap between science and practice: How therapists make their clinical decisions. *Behavior Therapy*, 45(2), 199-211.
- Haine-Schlagel, R., & Walsh, N. E. (2015). A review of parent participation engagement in child and family mental health treatment. *Clinical Child and Family Psychology review*, 18(2), 133-150.
- Hall, M. A., Dugan, E., Zheng, B., & Mishra, A. K. (2001). Trust in physicians and

- medical institutions: what is it, can it be measured, and does it matter?. *The Milbank Quarterly*, 79(4), 613-639.
- Hammen, C. (2009). Adolescent depression: Stressful interpersonal contexts and risk for recurrence. *Current Directions in Psychological Science*, 18(4), 200-204.
- Harris, K.M., C.T. Halpern, E. Whitsel, J. Hussey, J. Tabor, P. Entzel, and J.R. Udry. 2009. The National Longitudinal Study of Adolescent to Adult Health: Research Design [WWW document]. URL: <http://www.cpc.unc.edu/projects/addhealth/design>.
- Harris, K. M. (2013). The add health study: Design and accomplishments. *Chapel Hill: Carolina Population Center, University of North Carolina at Chapel Hill*.
- Harrison, M. E., McKay, M. M., & Bannon, W. M. (2004). Inner-city child mental health service use: The real question is why youth and families do not use services. *Community Mental Health Journal*, 40(2), 119-131.
- Hasnain-Wynia, R., Baker, D. W., Nerenz, D., Feinglass, J., Beal, A. C., Landrum, M. B., ... & Weissman, J. S. (2007). Disparities in health care are driven by where minority patients seek care: examination of the hospital quality alliance measures. *Archives of Internal Medicine*, 167(12), 1233-1239.
- Hayes, J. F., Marston, L., Walters, K., King, M. B., & Osborn, D. P. (2017). Mortality gap for people with bipolar disorder and schizophrenia: UK-based cohort study 2000–2014. *The British Journal of Psychiatry*, 211(3), 175-181.
- Herrman, H. (2012). The status of mental health promotion. *Public Health Reviews*, 34(2), 1-21.
- Hilbert, S., Goerigk, S., Padberg, F., Nadjiri, A., Übleis, A., Jobst, A., ... & Sarubin, N. (2019). The role of self-esteem in depression: A longitudinal Study. *Behavioural and Cognitive Psychotherapy*, 47(2), 244-250.

- Hoagwood, K. E. (2005). Family-based services in children's mental health: a research review and synthesis. *Journal of Child Psychology and Psychiatry*, 46(7), 690-713.
- Hoagwood, K., & Olin, S. S. (2002). The NIMH blueprint for change report: Research priorities in child and adolescent mental health. *Journal of the American Academy of Child & Adolescent Psychiatry*, 41(7), 760-767.
- Hoberman, H. M. (1992). Ethnic minority status and adolescent mental health services utilization. *The Journal of Mental Health Administration*, 19(3), 246-267.
- Hogan, M. F. (2003). New Freedom Commission report: The President's New Freedom Commission: Recommendations to transform mental health care in America. *Psychiatric Services*, 54(11), 1467-1474.
- Holway, G. V., Umberson, D., & Thomeer, M. B. (2017). Binge drinking and depression: the influence of romantic partners in young adulthood. *Society and Mental Health*, 7(1), 36-49.
- Horowitz, K., McKay, M., & Marshall, R. (2005). Community violence and urban families: experiences, effects, and directions for intervention. *American Journal of Orthopsychiatry*, 75(3), 356.
- Hunt, D. L., Haynes, R. B., Hanna, S. E., & Smith, K. (1998). Effects of computer-based clinical decision support systems on physician performance and patient outcomes: a systematic review. *Jama*, 280(15), 1339-1346.
- Indu, P. S., Anilkumar, T. V., Pisharody, R., Russell, P. S. S., Raju, D., Sarma, P. S., ... & Andrade, C. (2017). Prevalence of depression and past suicide attempt in primary care. *Asian Journal of Psychiatry*, 27, 48-52.
- Jackson, J. M. (2015). Multi-family groups for multi-stressed families initial outcomes

- and future implications. *Research on Social Work Practice*, 25(5), 537-548.
- Jenkins, R., Baingana, F., Ahmad, R., McDaid, D., & Atun, R. (2011). Social, economic, human rights and political challenges to global mental health. *Mental Health in Family Medicine*, 8(2), 87.
- Johannesen, J. K., Bi, J., Jiang, R., Kenney, J. G., & Chen, C. M. A. (2016). Machine learning identification of EEG features predicting working memory performance in schizophrenia and healthy adults. *Neuropsychiatric Electrophysiology*, 2(1), 3.
- Johnson, E., Mellor, D., & Brann, P. (2008). Differences in dropout between diagnoses in child and adolescent mental health services. *Clinical Child Psychology and Psychiatry*, 13(4), 515-530.
- Jorm, A. F. (2012). Mental health literacy: Empowering the community to take action for better mental health. *American Psychologist*, 67(3), 231.
- Kallus, N., & Zhou, A. (2018). Residual unfairness in fair machine learning from prejudiced data. arXiv preprint arXiv:1806.02887.
- Katz, S. J., Kessler, R. C., Lin, E., & Wells, K. B. (1998). Medication management of depression in the United States and Ontario. *Journal of General Internal Medicine*, 13(2), 77-85.
- Kawamoto, K., Houlihan, C. A., Balas, E. A., & Lobach, D. F. (2005). Improving clinical practice using clinical decision support systems: a systematic review of trials to identify features critical to success. *Bmj*, 330(7494), 765.
- Kazdin, A. E., & Mazurick, J. L. (1994). Dropping out of child psychotherapy: Distinguishing early and late dropouts over the course of treatment. *Journal of Consulting and Clinical Psychology*, 62(5), 1069.
- Kazdin, A. E. (1995). *Conduct disorders in childhood and adolescence* (Vol. 9). Sage Publications.

- Kazdin, A. E., Holland, L., & Crowley, M. (1997). Family experience of barriers to treatment and premature termination from child therapy. *Journal of Consulting and Clinical Psychology, 65*(3), 453.
- Kazdin, A. E., & Wassell, G. (1999). Barriers to treatment participation and therapeutic change among children referred for conduct disorder. *Journal of Clinical Child Psychology, 28*(2), 160-172.
- Kazdin, A. E. (2000). Perceived barriers to treatment participation and treatment acceptability among antisocial children and their families. *Journal of Child and Family Studies, 9*(2), 157-174.
- Kazdin, A. E., & Wassell, G. (2000). Predictors of barriers to treatment and therapeutic change in outpatient therapy for antisocial children and their families. *Mental Health Services Research, 2*(1), 27-40.
- Kelly, M. M., Tyrka, A. R., Anderson, G. M., Price, L. H., & Carpenter, L. L. (2008). Sex differences in emotional and physiological responses to the Trier Social Stress Test. *Journal of Behavior Therapy and Experimental Psychiatry, 39*(1), 87-98.
- Kerkorian, D., Bannon, W. M., & McKay, M. (2006). Seeking help a second time: Parents'/caregivers' characterizations of previous experiences with mental health services for their children and perceptions of barriers to future use. *American Journal of Orthopsychiatry, 76*(2), 161-166.
- Kessler, R. C., Avenevoli, S., & Merikangas, K. R. (2001). Mood disorders in children and adolescents: an epidemiologic perspective. *Biological Psychiatry, 49*(12), 1002-1014.
- 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.

- Kessler, R. C., Demler, O., Frank, R. G., Olfson, M., Pincus, H. A., Walters, E. E., ... & Zaslavsky, A. M. (2005). Prevalence and treatment of mental disorders, 1990 to 2003. *New England Journal of Medicine*, 352(24), 2515-2523.
- Kessler, R. C., Amminger, G. P., Aguilar-Gaxiola, S., Alonso, J., Lee, S., & Ustun, T. B. (2007). Age of onset of mental disorders: a review of recent literature. *Current Opinion in Psychiatry*, 20(4), 359.
- Kessler, R. C., Angermeyer, M., Anthony, J. C., De Graaf, R., Demyttenaere, K., Gasquet, I., ... & Kawakami, N. (2007). Lifetime prevalence and age-of-onset distributions of mental disorders in the World Health Organization's World Mental Health Survey Initiative. *World Psychiatry*, 6(3), 168.
- Kessler, R. C., Heeringa, S., Lakoma, M. D., Petukhova, M., Rupp, A. E., Schoenbaum, M., ... & Zaslavsky, A. M. (2008). Individual and societal effects of mental disorders on earnings in the United States: results from the national comorbidity survey replication. *American Journal of Psychiatry*.
- Kessler, R. C., van Loo, H. M., Wardenaar, K. J., Bossarte, R. M., Brenner, L. A., Cai, T., ... & Nierenberg, A. A. (2016). Testing a machine-learning algorithm to predict the persistence and severity of major depressive disorder from baseline self-reports. *Molecular Psychiatry*, 21(10), 1366.
- Keyes, K. M., Eaton, N. R., Krueger, R. F., McLaughlin, K. A., Wall, M. M., Grant, B. F., & Hasin, D. S. (2012). Childhood maltreatment and the structure of common psychiatric disorders. *The British Journal of Psychiatry*, 200(2), 107-115.
- Kilbourne, A. M., Beck, K., Spaeth-Rublee, B., Ramanuj, P., O'Brien, R. W., Tomoyasu, N., & Pincus, H. A. (2018). Measuring and improving the quality of mental health care: a global perspective. *World Psychiatry*, 17(1), 30-38.

- Kim-Cohen, J., Moffitt, T. E., Caspi, A., & Taylor, A. (2004). Genetic and environmental processes in young children's resilience and vulnerability to socioeconomic deprivation. *Child Development, 75*, 651-668.
- Kim, H., Tracy, E. M., Biegel, D. E., Min, M. O., & Munson, M. R. (2015). The effects of organizational culture on mental health service engagement of transition age youth. *The Journal of Behavioral Health Services & Research, 42*(4), 466-485.
- Kindig, D., & Stoddart, G. (2003). What is population health?. *American Journal of Public Health, 93*(3), 380-383.
- Klein, S., & Hostetter, M. (2014). In focus: Integrating behavioral health and primary care. *The Commonwealth Fund*.
- Koegl, C. J., Farrington, D. P., Augimeri, L. K., & Day, D. M. (2008). Evaluation of a targeted cognitive-behavioral program for children with conduct problems—The SNAP® under 12 outreach project: Service intensity, age and gender effects on short-and long-term outcomes. *Clinical Child Psychology and Psychiatry, 13*(3), 419-434.
- Koning, P., & Heinrich, C. J. (2013). Cream-skimming, parking and other intended and unintended effects of high-powered, performance-based contracts. *Journal of Policy Analysis and Management, 32*(3), 461-483.
- Koppelman, J. (2004, June). Children with mental disorders: making sense of their needs and the systems that help them. National Health Policy Forum.
- Kucharska, J. (2017). Sex differences in the appraisal of traumatic events and psychopathology. *Psychological Trauma: Theory, Research, Practice, and Policy, 9*(5), 575.
- Langer, C. S., Antonelli, R. C., Chamberlain, L., Pan, R. J., & Keller, D. (2018). Evolving

- federal and state health care policy: Toward a more integrated and comprehensive care-delivery system for children with medical complexity. *Pediatrics*, 141(Supplement 3), S259-S265.
- Lavoie, F., & Hodgins, S. (1994). Mental disorders among children with one parent with a lifetime diagnosis of major depression. *A critical review of the literature on children at risk for major affective disorders*, 37-82.
- Layous, K., Chancellor, J., & Lyubomirsky, S. (2014). Positive activities as protective factors against mental health conditions. *Journal of Abnormal Psychology*, 123(1), 3.
- Lieberman, A. F., & Knorr, K. (2007). The impact of trauma: A developmental framework for infancy and early childhood. *Pediatric Annals*, 36(4), 209-215.
- Leaf, P. J., Alegria, M., Cohen, P., Goodman, S. H., Horwitz, S. M., Hoven, C. W., ... & Regier, D. A. (1996). Mental health service use in the community and schools: Results from the four-community MECA study. *Journal of the American Academy of Child & Adolescent Psychiatry*, 35(7), 889-897.
- Leslie, K. R., & Chike-Harris, K. (2018). Patient-administered screening tool may improve detection and diagnosis of depression among adolescents. *Clinical Pediatrics*, 57(4), 457-460.
- Lester, P., Peterson, K., Reeves, J., Knauss, L., Glover, D., Mogil, C., & Beardslee, W. (2010). The long war and parental combat deployment: Effects on military children and at-home spouses. *Journal of the American Academy of Child & Adolescent Psychiatry*, 49(4), 310-320.
- Levitt, J. M., Saka, N., Romanelli, L. H., & Hoagwood, K. (2007). Early identification of mental health problems in schools: The status of instrumentation. *Journal of School Psychology*, 45(2), 163–191.

- Lewandowski, R. E., O'Connor, B., Bertagnolli, A., Beck, A., Tinoco, A., Gardner, W. P., ... & Brace, N. E. (2016). Screening for and diagnosis of depression among adolescents in a large health maintenance organization. *Psychiatric Services*, 67(6), 636-641.
- Lindsey, M. A., Browne, D. C., Thompson, R., Hawley, K. M., Graham, J. C., Weisbart, C., ... & Kotch, J. B. (2008). Caregiver mental health, neighborhood, and social network influences on mental health needs among African American children. *Social Work Research*, 32(2), 79-88.
- Li, Y. H., Cuccaro, P., Chen, H., Abughosh, S., Mehta, P. D., & Essien, E. J. (2018). HIV-related sexual decisions made by African-American adolescents living in different family structures: study from an ecodevelopmental perspective. *HIV/AIDS (Auckland, NZ)*, 10, 19.
- Lindsey, M. A., Brandt, N. E., Becker, K. D., Lee, B. R., Barth, R. P., Daleiden, E. L., & Chorpita, B. F. (2014). Identifying the common elements of treatment engagement interventions in children's mental health services. *Clinical Child and Family Psychology Review*, 17(3), 283-298.
- Logan, D. E., & King, C. A. (2002). Parental identification of depression and mental health service use among depressed adolescents. *Journal of the American Academy of Child and Adolescent Psychiatry*, 41, 296–304.
- Marcus, G., & Davis, E. (2014). Eight (no, nine!) problems with big data. *The New York Times*, 6(04), 2014.
- Marr, B. (2017). How Machine Learning is Transforming Healthcare. Retrieved from <http://data-informed.com/how-machine-learning-is-transforming-healthcare/>
- Marschner, I. C., & Bosch, R. J. (1998). Flexible assessment of trends in age-specific

- HIV incidence using two-dimensional penalized likelihood. *Statistics in Medicine*, 17(9), 1017-1031.
- Martínez-Hernández, A., Carceller-Maicas, N., DiGiacomo, S. M., & Ariste, S. (2016). Social support and gender differences in coping with depression among emerging adults: a mixed-methods study. *Child and Adolescent Psychiatry and Mental Health*, 10(1), 2.
- Masten, A. S. (2007). Resilience in developing systems: Progress and promise as the fourth wave rises. *Development and Psychopathology*, 19(3), 921–930.
- Masten, A. S., & Narayan, A. J. (2012). Child Development in the Context of Disaster, War, and Terrorism: Pathways of Risk and Resilience. *Annual Review of Psychology*, 63(63), 227–257.
- Matheson, F. I., Smith, K. L., Fazli, G. S., Moineddin, R., Dunn, J. R., & Glazier, R. H. (2014). Physical health and gender as risk factors for usage of services for mental illness. *J Epidemiol Community Health*, 68(10), 971-978.
- Mathers, C. D., & Loncar, D. (2006). Projections of global mortality and burden of disease from 2002 to 2030. *PLoS medicine*, 3(11), e442.
- Matthews, A. K., Corrigan, P. W., Smith, B. M., & Aranda, F. (2006). A qualitative exploration of African-Americans' attitudes toward mental illness and mental illness treatment seeking. *Rehabilitation Education*, 20(4).
- Matud, M. P. (2004). Gender differences in stress and coping styles. *Personality and Individual Differences*, 37(7), 1401-1415.
- Maulik, P. K., Eaton, W. W., & Bradshaw, C. P. (2011). The effect of social networks and

- social support on mental health services use, following a life event, among the Baltimore Epidemiologic Catchment Area cohort. *The Journal of Behavioral Health Services & Research*, 38(1), 29-50.
- Mayo Clinic. (2019, January 29). Women's Increased Risk of Depression. Retrieved from <https://www.mayoclinic.org/diseases-conditions/depression/in-depth/depression/art-20047725>
- McDonald, S., Kehler, H., Bayrampour, H., Fraser-Lee, N., & Tough, S. (2016). Risk and protective factors in early child development: Results from the All Our Babies (AOB) pregnancy cohort. *Research in Developmental Disabilities*, 58, 20-30.
- McKay, M. M., Stoewe, J., McCadam, K., & Gonzales, J. (1998). Increasing access to child mental health services for urban children and their caregivers. *Health & Social Work*, 23(1), 9-15.
- McKay, M. M., & Bannon Jr, W. M. (2004). Engaging families in child mental health services. *Child and Adolescent Psychiatric Clinics*, 13(4), 905-921.
- McKay, M. M., Hibbert, R., Hoagwood, K., Rodriguez, J., Murray, L., Legerski, J., & Fernandez, D. (2004). Integrating evidence-based engagement interventions into "real world" child mental health settings. *Brief Treatment and Crisis Intervention*, 4(2), 177.
- McKay, M. M., Harrison, M. E., Gonzales, J., Kim, L., & Quintana, E. (2002). Multiple-family groups for urban children with conduct difficulties and their families. *Psychiatric Services*, 53(11), 1467-1468.
- McKay, M. M., Lynn, C. J., & Bannon, W. M. (2005). Understanding inner city child mental health need and trauma exposure: implications for preparing urban service providers. *American Journal of Orthopsychiatry*, 75(2), 201.
- Meichenbaum, D., & Turk, D. C. (1987). Facilitating treatment adherence: A

practitioner's guidebook. New York: Plenum.

Mental Health in America (2019) The State of Mental Health in America. Retrieved April 3 2019, from <http://www.mentalhealthamerica.net/issues/state-mental-health-america>

Merikangas, K. R., Ames, M., Cui, L., Stang, P. E., Ustun, T. B., Von Korff, M., & Kessler, R. C. (2007). The impact of comorbidity of mental and physical conditions on role disability in the US adult household population. *Archives of General Psychiatry*, 64(10), 1180-1188.

Merikangas, K. R., He, J. P., Burstein, M., Swanson, S. A., Avenevoli, S., Cui, L., ... & Swendsen, J. (2010). Lifetime prevalence of mental disorders in US adolescents: results from the National Comorbidity Survey Replication–Adolescent Supplement (NCS-A). *Journal of the American Academy of Child & Adolescent Psychiatry*, 49(10), 980-989.

Milevsky, A. (2005). Compensatory patterns of sibling support in emerging adulthood: Variations in loneliness, self-esteem, depression and life satisfaction. *Journal of Social and Personal Relationships*, 22(6), 743-755.

Miller, L. M., Southam-Gerow, M. A., & Allin Jr, R. B. (2008). Who stays in treatment? Child and family predictors of youth client retention in a public mental health agency. In *Child & Youth Care Forum* (Vol. 37, No. 4, pp. 153-170). Springer US.

Murphy, J. M., Ichinose, C., Hicks, R. C., Kingdon, D., Crist-Whitzel, J., Jordan, P., ... & Jellinek, M. S. (1996). Utility of the Pediatric Symptom Checklist as a psychosocial screen to meet the federal Early and Periodic Screening, Diagnosis, and Treatment (EPSDT) standards: a pilot study. *The Journal of Pediatrics*, 129(6), 864-869.

Murphy, K. P. (2012). *Machine learning: a probabilistic perspective*. MIT press.

- Moore, K., Redd, Z., Burkhauser, M., Mbwana, K., & Collins, A. (2009). Children in poverty. *Child Trends Research Brief*, 11.
- Nadeem, E., Lange, J. M., & Miranda, J. (2009). Perceived need for care among low-income immigrant and US-born black and Latina women with depression. *Journal of Women's Health*, 18(3), 369-375.
- NAMI. Mental Health Conditions. (n.d.). Retrieved April 18, 2016, from <https://www.nami.org/Learn-More/Mental-Health-Conditions>
- Narendorf, S. C., Munson, M. R., Ben-David, S., Cole, A. R., & Scott Jr, L. D. (2018). Race and gender differences in attitudes toward help seeking among marginalized young adults with mood disorders: A mixed-methods investigation. *Psychiatric Rehabilitation Journal*, 41(4), 277.
- National Center for Health Statistics (US. (2016). Health, United States, 2015: with special feature on racial and ethnic health disparities.
- National Research Council. (2009). *Preventing mental, emotional, and behavioral disorders among young people: Progress and possibilities*. National Academies Press.
- National Survey on Drug Use and Health (NSDUH). (2017). Retrieved from <https://www.samhsa.gov/data/report/2017-nsduh-annual-national-report>
- NIMH (2017). Technology and the Future of Mental Health Treatment. Retrieved March 30, 2019, from <https://www.nimh.nih.gov/health/topics/technology-and-the-future-of-mental-health-treatment/index.shtml>
- Nock, M. K., & Ferriter, C. (2005). Parent management of attendance and adherence in child and adolescent therapy: A conceptual and empirical review. *Clinical Child and Family Psychology Review*, 8(2), 149-166.
- Norman, R. E., Byambaa, M., De, R., Butchart, A., Scott, J., & Vos, T. (2012). The long-

- term health consequences of child physical abuse, emotional abuse, and neglect: A systematic review and meta-analysis. *PLoS Medicine*, 9(11), 1–31.
- O'Connell, M. E., Boat, T., & Warner, K. E. (2009). *Preventing mental, emotional, and behavioral disorders among young people: Progress and possibilities* (Vol. 7). Washington, DC: National Academies Press.
- O'Reilly, M., Svirydzenka, N., Adams, S., & Dogra, N. (2018). Review of mental health promotion interventions in schools. *Social Psychiatry and Psychiatric Epidemiology*, 53(7), 647-662.
- Orru, G., Pettersson-Yeo, W., Marquand, A. F., Sartori, G., & Mechelli, A. (2012). Using support vector machine to identify imaging biomarkers of neurological and psychiatric disease: a critical review. *Neuroscience & Biobehavioral Reviews*, 36(4), 1140-1152.
- Owens, P. L., Hoagwood, K., Horwitz, S. M., Leaf, P. J., Poduska, J. M., Kellam, S. G., & Ialongo, N. S. (2002). Barriers to children's mental health services. *Journal of the American Academy of Child & Adolescent Psychiatry*, 41(6), 731-738.
- Ozer, E. M., Zahnd, E. G., Adams, S. H., Husting, S. R., Wibbelsman, C. J., Norman, K. P., & Smiga, S. M. (2009). Are adolescents being screened for emotional distress in primary care?. *Journal of Adolescent Health*, 44(6), 520-527.
- Paradis, A. D., Giaconia, R. M., Reinherz, H. Z., Beardslee, W. R., Ward, K. E., & Fitzmaurice, G. M. (2011). Adolescent family factors promoting healthy adult functioning: A longitudinal community study. *Child and Adolescent Mental Health*, 16(1), 30-37.
- Parks, J., Svendsen, D., Singer, P., Foti, M. E., & Mauer, B. (2006). Morbidity and

mortality in people with serious mental illness. *Alexandria, VA: National Association of State Mental Health Program Directors (NASMHPD) Medical Directors Council*, 25(4).

Passos, I. C., Mwangi, B., Cao, B., Hamilton, J. E., Wu, M. J., Zhang, X. Y., ... & Soares, J. C. (2016). Identifying a clinical signature of suicidality among patients with mood disorders: a pilot study using a machine learning approach. *Journal of Affective Disorders*, 193, 109-116.

Patel, V., Flisher, A. J., Hetrick, S., & McGorry, P. (2007). Mental health of young people: a global public-health challenge. *The Lancet*, 369(9569), 1302-1313.

Patel, S.R., & Bakken, S. (2010). Preferences for participation in decision making among ethnically diverse patients with anxiety and depression. *Community Mental Health Journal*, 46(5), 466–473.

Pellerin, K. A., Costa, N. M., Weems, C. F., & Dalton, R. F. (2010). An examination of treatment completers and non-completers at a child and adolescent community mental health clinic. *Community Mental Health Journal*, 46(3), 273-281.

Perou, R., Bitsko, R. H., Blumberg, S. J., Pastor, P., Ghandour, R. M., Gfroerer, J. C., ... & Parks, S. E. (2013). Mental health surveillance among children—United States, 2005–2011. *MMWR Surveill Summ*, 62(Suppl 2), 1-35

Petersen, L. A. (2002). Racial differences in trust: reaping what we have sown?. *Medical Care*, 40(2), 81-84.

Petitclerc, A., Boivin, M., Dionne, G., Zoccolillo, M., & Tremblay, R. E. (2009). Disregard for rules: The early development and predictors of a specific dimension of disruptive behavior disorders. *Journal of Child Psychology and Psychiatry*, 50(12), 1477-1484.

Pidano, A. E., Kimmelblatt, C. A., & Neace, W. P. (2011). Behavioral health in the

- pediatric primary care setting: Needs, barriers, and implications for psychologists. *Journal of Developmental and Behavioral Pediatrics*, 11, 1541-1559.
- Pincus, H. A., Scholle, S. H., Spaeth-Rublee, B., Hepner, K. A., & Brown, J. (2016). Quality measures for mental health and substance use: gaps, opportunities, and challenges. *Health Affairs*, 35(6), 1000-1008.
- Pincus, H. A., Spaeth-Rublee, B., & Watkins, K. E. (2011). The case for measuring quality in mental health and substance abuse care. *Health Affairs*, 30(4), 730-736.
- Popa, T. A., & Ladea, M. (2012). Nutrition and depression at the forefront of progress. *Journal of Medicine and Life*, 5(4), 414.
- Porche, M. V., Costello, D. M., & Rosen-Reynoso, M. (2016). Adverse family experiences, child mental health, and educational outcomes for a national sample of students. *School Mental Health*, 1-17.
- Puyat, J. H., Kazanjian, A., Goldner, E. M., & Wong, H. (2016). How often do individuals with major depression receive minimally adequate treatment? A population-based, data linkage study. *The Canadian Journal of Psychiatry*, 61(7), 394-404.
- Radloff, L. S. (1977). The CES-D scale: A self report depression scale for research in the general population. *Applied Psychological Measurements*, 1, 385-401.
- Raviv, A., Sharvit, K., Raviv, A., & Rosenblat-Stein, S. (2009). Mothers' and fathers' reluctance to seek psychological help for their children. *Journal of Child and Family Studies*, 18(2), 151-162.
- Rawal, P., & McCabe, M. A. (2016). Health care reform and programs that provide opportunities to promote children's behavioral health. *NAM Perspectives*.
- Russell, J. (2015). Predictive analytics and child protection: Constraints and opportunities. *Child Abuse & Neglect*, 46, 182-189.

- Rosenberg, M. (1989). *Society and the adolescent self-image* (Rev. ed.). Middletown, CT: Wesleyan University Press.
- Safran, M. A., Mays, R. A., Huang, L. N., McCuan, R., Pham, P. K., Fisher, S. K., McDuffie, K. Y., ... Trachtenberg, A. (2009). Mental health disparities. *American Journal of Public Health, 99*(11), 1962-6.
- Salk, R. H., Hyde, J. S., & Abramson, L. Y. (2017). Gender differences in depression in representative national samples: Meta-analyses of diagnoses and symptoms. *Psychological Bulletin, 143*(8), 783.
- Samet, J. H., Friedmann, P., & Saitz, R. (2001). Benefits of linking primary medical care and substance abuse services: patient, provider, and societal perspectives. *Archives of Internal Medicine, 161*(1), 85-91.
- Santiago, C. D., Kaltman, S., & Miranda, J. (2013). Poverty and mental health: how do low-income adults and children fare in psychotherapy?. *Journal of Clinical Psychology, 69*(2), 115-126.
- Sayal, K. (2006). Annotation: Pathways to care for children with mental health problems. *Journal of Child Psychology and Psychiatry, 47*(7), 649-659.
- Schoenwald, S. K., Chapman, J. E., Kelleher, K., Hoagwood, K. E., Landsverk, J., Stevens, J., ... & Research Network on Youth Mental Health. (2008). A survey of the infrastructure for children's mental health services: Implications for the implementation of empirically supported treatments (ESTs). *Administration and Policy in Mental Health and Mental Health Services Research, 35*(1-2), 84-97.
- Schreck, C. J., Fisher, B. S., & Miller, J. M. (2004). The social context of violent victimization: A study of the delinquent peer effect. *Justice Quarterly, 21*(1), 23-47.

- Seidler, Z. E., Dawes, A. J., Rice, S. M., Oliffe, J. L., & Dhillon, H. M. (2016). The role of masculinity in men's help-seeking for depression: a systematic review. *Clinical Psychology Review, 49*, 106-118.
- Shams, I., Ajorlou, S., & Yang, K. (2014). A predictive analytics approach to reducing avoidable hospital readmission. *arXiv preprint arXiv:1402.5991*.
- Shahar, G., & Henrich, C. C. (2010). Do depressive symptoms erode self-esteem in early adolescence?. *Self and Identity, 9*(4), 403-415.
- Shaw, D. S., Criss, M. M., Schonberg, M. A., & Beck, J. E. (2004). The development of family hierarchies and their relation to children's conduct problems. *Development and Psychopathology, 16*(03), 483-500.
- Shortreed, S. M. (2018). Predicting suicide attempts and suicide deaths following outpatient visits using electronic health records. *American Journal of Psychiatry, 175*(10), 951-960.
- Simon, G. E., Johnson, E., Lawrence, J. M., Rossom, R. C., Ahmedani, B., Lynch, F. L., ... & Simonian, S. J. (2006). Screening and identification in pediatric primary care. *Behavior Modification, 30*(1), 114-131.
- Singal, A. G., Mukherjee, A., Elmunzer, B. J., Higgins, P. D., Lok, A. S., Zhu, J., ... & Waljee, A. K. (2013). Machine learning algorithms outperform conventional regression models in predicting development of hepatocellular carcinoma. *The American Journal of Gastroenterology, 108*(11), 1723.
- Smokowski, P. R., Guo, S., Rose, R., Evans, C. B., Cotter, K. L., & Bacallao, M. (2014). Multilevel risk factors and developmental assets for internalizing symptoms and self-esteem in disadvantaged adolescents: Modeling longitudinal trajectories from the Rural Adaptation Project. *Development and psychopathology, 26*(4pt2), 1495-1513.

- Smith, K. (2014). Mental health: a world of depression. *Nature News*, 515(7526), 180.
- Southam-Gerow, M. A., Weisz, J. R., Chu, B. C., McLeod, B. D., Gordis, E. B., & Connor-Smith, J. K. (2010). Does cognitive behavioral therapy for youth anxiety outperform usual care in community clinics? An initial effectiveness test. *Journal of the American Academy of Child & Adolescent Psychiatry*, 49(10), 1043-1052
- Sowislo, J. F., & Orth, U. (2013). Does low self-esteem predict depression and anxiety? A meta-analysis of longitudinal studies. *Psychological Bulletin*, 139(1), 213.
- Stagman, S. M., & Cooper, J. L. (2010). Children's mental health: What every policymaker should know. National Center for Children in Poverty.
- Stein, B. D., Sorbero, M. J., Dalton, E., Ayers, A. M., Farmer, C., Kogan, J. N., & Goswami, U. (2013). Predictors of adequate depression treatment among Medicaid-enrolled youth. *Social Psychiatry and Psychiatric Epidemiology*, 48(5), 757-765.
- Stockdale, S. E., Lagomasino, I. T., Siddique, J., McGuire, T., & Miranda, J. (2008). Racial and ethnic disparities in detection and treatment of depression and anxiety among psychiatric and primary health care visits, 1995–2005. *Medical Care*, 46(7), 668.
- Substance Abuse and Mental Health Services Administration. (2011). Identifying mental health and substance use problems of children and adolescents: A guide for child-serving organizations (HHS Publication No. SMA 12-4670). Rockville, MD: Author.
- Sussman, S., & Ames, S. L. (2008). *Drug abuse: Concepts, prevention, and cessation*. Cambridge University Press.
- Tanenbaum, S. J. (2005). Evidence-based practice as mental health policy: Three controversies and a caveat. *Health Affairs*, 24(1), 163-173.

- Tennant, R., Goens, C., Barlow, J., Day, C., & Stewart-Brown, S. (2007). A systematic review of reviews of interventions to promote mental health and prevent mental health problems in children and young people. *Journal of Public Mental Health, 6*(1), 25-32.
- Thompson, S. J., Bender, K., Lantry, J., & Flynn, P. M. (2007). Treatment engagement: Building therapeutic alliance in home-based treatment with adolescents and their families. *Contemporary Family Therapy, 29*(1-2), 39-55.
- Thompson, L., & McCabe, R. (2012). The effect of clinician-patient alliance and communication on treatment adherence in mental health care: a systematic review. *BMC psychiatry, 12*(1), 87.
- Tolan, P. H., Henry, D. B., Schoeny, M. S., Lovegrove, P., & Nichols, E. (2014). Mentoring programs to affect delinquency and associated outcomes of youth at risk: A comprehensive meta-analytic review. *Journal of Experimental Criminology, 10*(2), 179-206.
- Trude, S., & Stoddard, J. J. (2003). Referral gridlock. *Journal of General Internal Medicine, 18*(6), 442-449.
- U.S. Department of Health & Human Services. (1999). Mental Health: A Report of the Surgeon General. Rockville, Md: U.S. Department of Health and Human Services, Substance Abuse and Mental Health Services Administration, Center for Mental Health Services, National Institutes of Health, National Institute of Mental Health.
- US Department of Health and Human Services. (2000). Report of the Surgeon General's Conference on Children's Mental Health: A national action agenda.
- Vander Stoep, A., Weiss, N. S., Kuo, E. S., Cheney, D., & Cohen, P. (2003). What

- proportion of failure to complete secondary school in the US population is attributable to adolescent psychiatric disorder?. *The Journal of Behavioral Health Services & Research*, 30(1), 119-124.
- Vogel, D. L., Heimerdinger-Edwards, S. R., Hammer, J. H., & Hubbard, A. (2011). "Boys don't cry": Examination of the links between endorsement of masculine norms, self-stigma, and help-seeking attitudes for men from diverse backgrounds. *Journal of Counseling Psychology*, 58(3), 368.
- Waller, R., Gardner, F., Dishion, T., Sitnick, S. L., Shaw, D. S., Winter, C. E., & Wilson, M. (2015). Early parental positive behavior support and childhood adjustment: Addressing enduring questions with new methods. *Social Development*, 24(2), 304-322.
- Wang, P. S., Berglund, P., & Kessler, R. C. (2000). Recent care of common mental disorders in the United States. *Journal of General Internal Medicine*, 15(5), 284-292.
- Wang, P. S., Demler, O., & Kessler, R. C. (2002). Adequacy of treatment for serious mental illness in the United States. *American Journal of Public Health*, 92(1), 92-98.
- Wang, P. S., Lane, M., Olfson, M., Pincus, H. A., Wells, K. B., & Kessler, R. C. (2005). Twelve-month use of mental health services in the United States: results from the National Comorbidity Survey Replication. *Archives of General Psychiatry*, 62(6), 629-640.
- Watson Health. (2016) Social and behavioral determinants of health. Retrieved from <https://www.ibm.com/blogs/watson-health/social-and-behavioral-determinants-of-health/>
- Weare, K., & Nind, M. (2011). Mental health promotion and problem prevention in

- schools: what does the evidence say?. *Health Promotion International*, 26(suppl 1), i29-i69.
- Weissman, M. M., & Klerman, G. L. (1977). Sex differences and the epidemiology of depression. *Archives of General Psychiatry*, 34(1), 98-111.
- Weisz, J. R., Chorpita, B. F., Palinkas, L. A., Schoenwald, S. K., Miranda, J., Bearman, S. K., ... & Gray, J. (2012). Testing standard and modular designs for psychotherapy treating depression, anxiety, and conduct problems in youth: A randomized effectiveness trial. *Archives of General Psychiatry*, 69(3), 274-282.
- Wellander, L., Wells, M. B., & Feldman, I. (2016). Does prevention pay? Costs and potential cost-savings of school interventions targeting children with mental health problems. *Journal of Mental Health Policy and Economics*, 19(2), 91-101.
- Werner, E., & Smith, R. (1982). *Vulnerable but invincible: A longitudinal study of resilient children and youth*. New York: McGraw-Hill.
- Whiteman, K., Ruggiano, N., & Thomlison, B. (2016). Transforming mental health services to address gender disparities in depression risk factors. *Journal of Women & Aging*, 28(6), 521-529.
- Wight, R. G., Botticello, A. L., & Aneshensel, C. S. (2006). Socioeconomic context, social support, and adolescent mental health: A multilevel investigation. *Journal of Youth and Adolescence*, 35(1), 109.
- Williams, J., Klinepeter, K., Palmes, G., Pulley, A., & Foy, J. M. (2004). Diagnosis and treatment of behavioral health disorders in pediatric practice. *Pediatrics*, 114(3), 601-606.
- Wilson, S. J., & Lipsey, M. W. (2007). School-based interventions for aggressive and

- disruptive behavior: Update of a meta-analysis. *American Journal of Preventive Medicine*, 33(2), S130-S143.
- Wissow, L. S., Brown, J., Fothergill, K. E., Gadamski, A., Hacker, K., Salmon, P., & Zelkowitz, R. (2013). Universal mental health screening in pediatric primary care: a systematic review. *Journal of the American Academy of Child & Adolescent Psychiatry*, 52(11), 1134-1147.
- Wolk, C. B., Kendall, P. C., & Beidas, R. S. (2015). Cognitive-behavioral therapy for child anxiety confers long-term protection from suicidality. *Journal of the American Academy of Child & Adolescent Psychiatry*, 54(3), 175-179.
- Wood, J. J., McLeod, B. D., Sigman, M., Hwang, W. C., & Chu, B. C. (2003). Parenting and childhood anxiety: Theory, empirical findings, and future directions. *Journal of Child Psychology and Psychiatry*, 44(1), 134-151.
- Wright, J., Beaver, K., Delisi, M., & Vaughn, M. (2008). Evidence of negligible parenting influences on self-control, delinquent peers, and delinquency in a sample of twins. *Justice Quarterly*, 25(3), 544-569.
- Yap, M. B. H., Pilkington, P. D., Ryan, S. M., & Jorm, A. F. (2014). Parental factors associated with depression and anxiety in young people: A systematic review and meta-analysis. *Journal of Affective Disorders*, 156, 8–23.
- Yoo, W., Ference, B. A., Cote, M. L., & Schwartz, A. (2012). A comparison of logistic regression, logic regression, classification tree, and random forests to identify effective gene-gene and gene-environmental interactions. *International Journal of Applied Science and Technology*, 2(7), 268.
- Young, A. S., Klap, R., Sherbourne, C. D., & Wells, K. B. (2001). The quality of care for

- depressive and anxiety disorders in the United States. *Archives of General Psychiatry*, 58(1), 55-61.
- Zarley, D. (2019, January 28). Meet the scientists who are training AI to diagnose mental illness. Retrieved from <https://www.theverge.com/2019/1/28/18197253/ai-mental-illness-artificial-intelligence-science-neuroimaging-mri>
- Zhu, L. (2018). A multi-level analysis on school connectedness, family support, and adolescent depression: Evidence from the National Longitudinal Study of Adolescent Health, 1995–1996. *Social Sciences*, 7(5), 72.
- Zima, B. T., Hurlburt, M. S., Knapp, P., Ladd, H., Tang, L., Duan, N., ... & Wells, K. B. (2005). Quality of publicly-funded outpatient specialty mental health care for common childhood psychiatric disorders in California. *Journal of the American Academy of Child & Adolescent Psychiatry*, 44(2), 130-144.
- Zuckerbrot, R. A., Cheung, A., Jensen, P. S., Stein, R. E., Laraque, D., & GLAD-PC STEERING GROUP. (2018). Guidelines for adolescent depression in primary care (GLAD-PC): part I. Practice preparation, identification, assessment, and initial management. *Pediatrics*, e20174081.