Predicting Adverse Adolescent Outcomes:

Risk Tiering in a Safety Net Healthcare System

A DISSERTATION

Submitted to the Faculty of

Marywood University in partial fulfillment

of the requirements

for the Degree of Doctor of Philosophy

By

Remle Newton-Dame

Marywood University

Scranton, PA

August 2025

Dissertation Chair: Dr. Alan Levine

Dissertation Committee Members: Dr. Katherine Piwnica-Worms, Dr. Dave Chokshi

Dissertation Readers: Dr. Nichola Davis, Dr. Mary McCord

MARYWOOD UNIVERSITY

STRATEGIC LEADERSHIP AND ADMINISTRATIVE STUDIES

DISSERTATION APPROVAL

We hereby approve the dissertation

Predicting Adverse Adolescent Outcomes:

Risk Tiering in a Safety Net Healthcare System

by Remle Newton-Dame

Candidate for the Degree: Doctor of Philosophy

Date of Approval:	Approved by:	Signature:
8/7/2025	Dr. Alan Levine Committee Chair	Signed by: Alan M. Levine F8AE2BC776A14D3...
8/6/2025	Dr. Katherine Piwnica-Worms Committee Member	Signed by: Katherine Piwnica-Worms 702D9DC66CB2482...
8/6/2025	Dr. Dave Chokshi Committee Member	Signed by: 37549DE96B204F1...
8/6/2025	Dr. Nichola Davis Reader	Signed by: Nichola Davis 96881A2414F041E...
8/7/2025	Dr. Mary McCord Reader	DocuSigned by: AE6CCBD578C846E...

MARYWOOD UNIVERSITY

STRATEGIC LEADERSHIP AND ADMINISTRATIVE STUDIES

DISSERTATION APPROVAL

We hereby approve the dissertation

Predicting Adverse Adolescent Outcomes:

Risk Tiering in a Safety Net Healthcare System

by Remle Newton-Dame

Candidate for the Degree: Doctor of Philosophy

Date of Approval:	Approved by:	Signature:
8/7/2025	Dr. Alan Levine Committee Chair	Signed by: Alan M. Levine F8AE2BC776A14D3...
8/6/2025	Dr. Katherine Piwnica-Worms Committee Member	Signed by: Katherine Piwnica-Worms 702D9DC66CB2482...
8/6/2025	Dr. Dave Chokshi Committee Member	Signed by: 37549DE96B204F1...
8/6/2025	Dr. Nichola Davis Reader	Signed by: Nichola Davis 96881A2414F041E...
8/7/2025	Dr. Mary McCord Reader	DocuSigned by: AE6CCBD578C846E...

MARYWOOD UNIVERSITY

STRATEGIC LEADERSHIP AND ADMINISTRATIVE STUDIES

DISSERTATION APPROVAL

We hereby approve the dissertation

Predicting Adverse Adolescent Outcomes:

Risk Tiering in a Safety Net Healthcare System

by Remle Newton-Dame

Candidate for the Degree: Doctor of Philosophy

Date of Approval:	Approved by:	Signature:
8/7/2025	Dr. Alan Levine Committee Chair	Signed by: Alan M. Levine F8AE2BC776A14D3...
8/6/2025	Dr. Katherine Piwnica-Worms Committee Member	Signed by: Katherine Piwnica-Worms 702D9DC66CB2482...
8/6/2025	Dr. Dave Chokshi Committee Member	Signed by: 37549DE96B204F1...
8/6/2025	Dr. Nichola Davis Reader	Signed by: Nichola Davis 96881A2414F041E...
8/7/2025	Dr. Mary McCord Reader	DocuSigned by: AE6CCBD578C846E...

MARYWOOD UNIVERSITY

STRATEGIC LEADERSHIP AND ADMINISTRATIVE STUDIES

DISSERTATION APPROVAL

We hereby approve the dissertation

Predicting Adverse Adolescent Outcomes:

Risk Tiering in a Safety Net Healthcare System

by Remle Newton-Dame

Candidate for the Degree: Doctor of Philosophy

Date of Approval:	Approved by:	Signature:
8/7/2025	Dr. Alan Levine Committee Chair	Signed by: Alan M. Levine F8AE2BC776A14D3...
8/6/2025	Dr. Katherine Piwnica-Worms Committee Member	Signed by: Katherine Piwnica-Worms 702D9DC66CB2482...
8/6/2025	Dr. Dave Chokshi Committee Member	Signed by: 37549DE96B204F1...
8/6/2025	Dr. Nichola Davis Reader	Signed by: Nichola Davis 96881A2414F041E...
8/7/2025	Dr. Mary McCord Reader	DocuSigned by: AE6CCBD578C846E...

MARYWOOD UNIVERSITY

STRATEGIC LEADERSHIP AND ADMINISTRATIVE STUDIES

DISSERTATION APPROVAL

We hereby approve the dissertation

Predicting Adverse Adolescent Outcomes:

Risk Tiering in a Safety Net Healthcare System

by Remle Newton-Dame

Candidate for the Degree: Doctor of Philosophy

Date of Approval:	Approved by:	Signature:
8/7/2025	Dr. Alan Levine Committee Chair	Signed by: Alan M. Levine F8AE2BC776A14D3...
8/6/2025	Dr. Katherine Piwnica-Worms Committee Member	Signed by: Katherine Piwnica-Worms 702D9DC66CB2482...
8/6/2025	Dr. Dave Chokshi Committee Member	Signed by: 37549DE96B204F1...
8/6/2025	Dr. Nichola Davis Reader	Signed by: Nichola Davis 96881A2414F041E...
8/7/2025	Dr. Mary McCord Reader	DocuSigned by: AE6CCBD578C846E...

MARYWOOD UNIVERSITY

STRATEGIC LEADERSHIP AND ADMINISTRATIVE STUDIES

DISSERTATION APPROVAL

We hereby approve the dissertation

Predicting Adverse Adolescent Outcomes:

Risk Tiering in a Safety Net Healthcare System

by Remle Newton-Dame

Candidate for the Degree: Doctor of Philosophy

Date of Approval:	Approved by:	Signature:
8/7/2025	Dr. Alan Levine Committee Chair	Signed by: Alan M. Levine F8AE2BC776A14D3...
8/6/2025	Dr. Katherine Piwnica-Worms Committee Member	Signed by: Katherine Piwnica-Worms 702D9DC66CB2482...
8/6/2025	Dr. Dave Chokshi Committee Member	Signed by: 37549DE96B204F1...
8/6/2025	Dr. Nichola Davis Reader	Signed by: Nichola Davis 96881A2414F041E...
8/7/2025	Dr. Mary McCord Reader	DocuSigned by: AE6CCBD578C846E...

MARYWOOD UNIVERSITY

STRATEGIC LEADERSHIP AND ADMINISTRATIVE STUDIES

DISSERTATION APPROVAL

We hereby approve the dissertation

Predicting Adverse Adolescent Outcomes:

Risk Tiering in a Safety Net Healthcare System

by Remle Newton-Dame

Candidate for the Degree: Doctor of Philosophy

Date of Approval:	Approved by:	Signature:
8/7/2025	Dr. Alan Levine Committee Chair	Signed by: Alan M. Levine F8AE2BC776A14D3...
8/6/2025	Dr. Katherine Piwnica-Worms Committee Member	Signed by: Katherine Piwnica-Worms 702D9DC66CB2482...
8/6/2025	Dr. Dave Chokshi Committee Member	Signed by: 37549DE96B204F1...
8/6/2025	Dr. Nichola Davis Reader	Signed by: Nichola Davis 96881A2414F041E...
8/7/2025	Dr. Mary McCord Reader	DocuSigned by: AE6CCBD578C846E...

Chapter 1: The Problem and Its Setting

Introduction

New York City is a place of great wealth and great inequality for adolescents and their families. Of the 1.7 million children living in NYC in 2019, almost 747,000 were aged 10-17 (*Child Population*, n.d.). City-wide statistics often aggregate across younger and adolescent children. Nevertheless, looking at the conditions of families across the city is revealing. In New York City, an estimated 330,000 children were living in poverty as of 2020 (Aribot et al., 2022). Thirty eight percent of children were living in families experiencing "material hardship," including challenges in paying for housing, food, utilities or medical care.

These economic challenges translate to housing challenges. In 2017, 32% of families with children had fallen behind on their rent, a statistic that likely increased in the current inflation-related housing crisis in the city (Mironova & Bach, 2018). Homelessness is a significant challenge facing children in NYC. During the 2020-2021 school year, 101,000 children experiencing homelessness attended NYC public schools, including 28,000 children living in a shelter (Student Homelessness in New York City, 2021). Food insecurity is also common. The pandemic disrupted food security systems in the city, with 34% of households with children reporting difficulty in accessing emergency resources like food pantries in October of 2020 (Aldo Crossa et al., 2021). When it is difficult to afford food or housing, it can be challenging for families to access the healthcare that they need.

Enter NYC Health+Hospitals (H+H), the nation's largest safety net healthcare system. H+H offers primary, specialty, emergency and inpatient care to all New Yorkers, regardless of ability to pay. Across eleven hospitals and 70+ community health centers, H+H cares for majority-uninsured and Medicaid patients, serving more than a million New Yorkers each year. H+H offers social support as an embedded feature of overall health promotion. This includes on-site legal help for eviction prevention, immigration support or personal need. Referrals are available to community resources like food pantries during the care encounter via a platform

embedded in the medical record. Staff also support tax preparation and transportation coordination. Enrollers are on-site to assist with WIC and SNAP benefits or to register eligible patients for MetroPlus, a municipal Medicaid health plan. Health+Hospitals also created its own municipal health benefit program for uninsured patients, NYC Care. NYC Care provides >135,000 uninsured patients with access to care for a simple co-pay at any NYC Health+Hospitals location (James et al., 2021). These services, combined with the largest primary care Community Health Worker (CHW) program in the country, make Health+Hospitals a leader in socially responsible, community-oriented health care.

NYC Health+Hospitals has a robust pediatric population. In 2021, 209,600 children visited H+H, including 46,600 adolescent primary care patients (aged 12-17). Those adolescent patients were a diverse sample of New York, with 46.4% documented as Hispanic/Latinx, 34.4% as Black/African American, 5.0% as Asian/Pacific Islander and 2.4% as White. Among the H+H adolescent primary care population, 33.2% reported a primary language other than English.

Adolescents in care at H+H may be living with multiple medical and social stressors. The rate of uninsured adult patients across the system was 27.7% in 2021, suggesting many of the adolescent patients are living in families with uninsured adults. In New York State, children have universal health care access and undocumented immigrant children can be publicly insured through CHIP Medicaid plans, although some eligible patients are not enrolled. H+H is one of the primary care providers of choice for children and youths in the NYC foster care system. A true count of the number of H+H adolescents experiencing homelessness is unavailable, given the difficulty of measuring the kinds of strategies (couch surfing, doubling up) often deployed by families in response to housing stress. However, in 2021, H+H saw at least 881 adolescents documented as experiencing homelessness in its primary care clinics. Additionally, H+H adolescent patients are likely to live in neighborhoods that have faced decades of systematic underinvestment and excessive environmental hazards (Environmental Justice NYC Initiative, 2024). Exposure to pollution and poor housing quality impacts asthma. The rate of diagnosed

asthma in H+H adolescent primary care patients (16.7%) is two times higher than available city-wide prevalence estimates (Walters, Sarah et al., 2021).

For children and adolescents growing up amidst systemic inequity, connection to trusted, high-quality primary care is critical (Stirling et al., 2024). To develop into healthy adults, adolescents need healthy conditions and support at the individual, relational and community levels, and broader societal alignment to promote equity, respect and inclusion ("Youth 360°," n.d.). High-quality primary care which centers the experience and dignity of the adolescent patient while addressing physical and social needs is a critical piece of this healthy ecosystem (Alderman et al., 2019; Riley et al., 2018).

This kind of nuanced primary care takes time. The average H+H pediatric provider has 1,444 patients to manage in his or her panel (July 2024), and the average pediatric encounter nationally is just 16 minutes (Overhage & Johnson, 2020). Being able to quickly identify the adolescents and families who would most benefit from enhanced intervention and connection to community support would improve H+H's ability to target services to the patients who need them most. This would not substitute for clinician judgment, but rather provide a tool to flag adolescents for whom enhanced evaluation and monitoring is warranted.

However, defining adolescent risk in the general primary care population is challenging. In the last decade, research into pediatric risk scores has proliferated, spanning readmission to the hospital (Delaplain et al., 2020), cancer (Pinto et al., 2015), sepsis (Atreya & Wong, 2019), traumatic brain injury (Kuppermann et al., 2009), mortality after trauma (Muisyo et al., 2019), surgery risk (Ji et al., 2019) and more. However, primary care providers are tasked with taking care of the whole person, both broad physical health and "the behavioral, environmental, social, and other factors that influence health outcomes" (Thornton et al., 2024). Primary care providers must consider not just one narrow outcome but the spectrum of risks which adolescent patients face, from sexually transmitted infections to substance use, depression to poorly controlled

asthma. No holistic risk score is currently commercially or academically available that covers the primary sources of adolescent risk from a primary care perspective.

To create such a score, this inquiry sought to identify and predict adverse outcomes in adolescent health and development. Specifically, this study sought to harness data from the electronic medical record (EMR) and geographically matched sources to predict the risk of adverse health outcomes that an adolescent patient will experience in the coming year, using the concept of Adverse Childhood Experiences (ACEs) to bound a working definition of adolescent risk, among adolescents aged 12-17 in primary care in 2022 at NYC H+H.

Theoretical Framework: Adolescent Development

Adolescents are a distinct group of patients with distinct needs. As outlined in Rudolph's *Pediatrics*, during adolescence patients are experiencing hormonal puberty, "expand[ing] social relationships…transition[ing] from concrete to abstract thinking," developing sexual feelings and drives, turning to peers to identify behavioral standards, testing independence, defining identity, deepening compassion and increasing complex thinking (Ozer & Irwin Jr., 2011). This development is not happening in a vacuum. As Urie Bronfenbrenner theorized in 1977 in his ecological systems theory, pediatric development is influenced by structural factors within multiple nested environments that impact that child or adolescent. Although Bronfenbrenner coined these systems the microsystem, mesosystem, exosystem, macrosystem and chronosystem, subsequent theorists and public health practitioners have further adapted and refined the social-ecological model. The Healthy Teen Network has produced a version of the social-ecological model tailored to teen development which illustrates this dynamic for adolescent patients (Figure 1.1).

According to their social-ecological model, teenage development is influenced first by a series of individual factors. These include intrinsic characteristics, like genetic predispositions, sexual orientation or gender identity; social determinants of health like food and housing

security; and behavioral factors like self-efficacy and smart use of media which are shaped both by personal temperament and longitudinal environmental influences.

At the next level of the model are relationships, which are extremely important in a teenager's life. This level of the model extends beyond familial relationships to other trusted adults, friends and peers. The quality of these relationships is crucial. Relationships with caregivers and trusted adults that support healthy development are warm, supportive and foster open communication. These kinds of relationships can mitigate the impacts of trauma and adversity on a developing brain. Peer relationships and social networks which provide deep connection and reinforce healthy behaviors and decision-making are critical to a developing teen's mental health, sense of self and social connection or isolation.

At the community level, a developing teenager is influenced by the communities in which they live and learn. A healthy community provides safety and opportunities to socialize and exercise. A healthy community offers access to healthy food, transportation and non-stigmatizing healthcare. Healthy communities also support the transition to adulthood through access to educational resources, job opportunities, and culturally competent mentors. Barriers to these community-level factors, due to factors like systemic racism, lack of municipal investment and the long-term effects of redlining, adversely impact adolescent development.

Finally, teens are growing within a societal ecosystem. This includes formal and informal policy and funding (or lack of funding) for education and development. It also includes less tangible but critical factors like social norms and social equity or inequity. Growing up in a deeply inequitable societal ecosystem negatively impacts development across racial and socio-economic groups (Benner, 2017).

The social ecological model provides a high-level view of the rich web of interactions and factors guiding teen development at the individual, relationship, community and societal levels. Although each teenager will experience their own challenges, sustained stress at any level of this model can impact adolescent development, health and long-term wellbeing (Garner et al.,

2021). Although limited stress can be adaptive or tolerable, when stress becomes intense and sustained, it can become toxic (C. A. Nelson et al., 2020). According to Bucci et al, “prolonged activation of [the] stress response…disrupts brain architecture and increases risk of health disorders [as] prolonged allostasis establishes a chronic stress response” (Bucci et al., 2016). This chronic stress response has neurologic, immunologic, metabolic and epigenetic impacts (Danese & McEwen, 2012; Rogosch et al., 2011; Slopen et al., 2012). Adolescents who have experienced or are experiencing pervasive toxic stress at any level of this model may experience negative physical and behavioral health effects if they do not have mitigating sources of support from family, friends, trusted adults and communities (Boullier & Blair, 2018; Garner et al., 2021; Shonkoff et al., 2021). One way in which the healthcare community has begun to measure exposure to toxic stress is through the study of “Adverse Childhood Experiences,” or ACEs.

Conceptual Framework: Effects of Adverse Childhood Experiences

Adolescents are subject to a unique set of behavioral risks, compared to younger children or adults. As identified by Riley et al, “[m]ost adolescents are generally healthy but participate in risky behaviors, such as unsafe driving, violence, drug and alcohol use, and unprotected sex, at disproportionately high rates, which can lead to significant morbidity and mortality” (Riley et al., 2018). Although adolescent brains are innately hard wired to crave novelty and risk as compared to adult brains (Siegel, 2013), some adolescents are more disposed towards risk-taking than others. Adverse Childhood Experiences are an important etiologic factor.

Per the CDC, Adverse Childhood Experiences are “ potentially traumatic events that occur in childhood…[which] can change brain development and affect how the body responds to stress” (CDC, 2021). Examples of ACES include physical, sexual or emotional abuse; and exposure to violence, substance use or mental illness during childhood (Dube et al., 2003; Turner et al., 2020). ACEs are significantly associated with a wide range of poor outcomes

covering both physical and behavioral health, in early childhood, adolescence and adulthood. As illustrated by the Institute for Health Equity in the below conceptual framework, past exposure to ACEs impacts current and future health and wellbeing through both direct and mediated pathways. This conceptual framework is provided in Figure 1.2, in the Tables and Figures section of this work.

As this conceptual framework illustrates, exposure to ACEs impacts the production of health both directly and indirectly (Allen, Matilda & Donkin, Angela, 2015). The direct impact of ACEs on depression and anxiety is well-documented. However, other more surprising conditions such as allergies, asthma, eczema, diabetes and repeat infections have also been linked to ACEs and the many physical impacts of toxic stress (C. A. Nelson et al., 2020). ACEs also indirectly affect adolescent health by changing the brain. A growing body of research suggests that exposure to ACEs during crucial developmental windows changes the structure and functioning of the brain systems involved in the regulation of impulse and reward, establishing a pathway between ACEs exposure and increased behavioral risk-taking broadly (Shonkoff et al., 2021). For example, smoking is an independent health-harming behavior, which can impact asthma control, physical stamina and respiratory illness during adolescence itself (*Tobacco Use Among Children and Teens | American Lung Association*, n.d.). However, exposure to ACEs is associated with earlier smoking initiation and heavier smoking among adults (Indiana Department of Health, 2021).

The American Academy of Pediatrics and the CDC have given significant attention to the necessity of preventing and treating ACEs (CDC, 2021). One of many intervention levers is connection to high quality primary care, through which adolescents can receive integrated behavioral healthcare, appropriate specialty referrals, connection to social workers or Community Health Workers, engagement of community-level supports, fostering of trusted relationships with adults and peers, and other forms of support focused on building individual and family resilience (Burke Harris, 2018; Lloyd et al., 2021; Traub & Boynton-Jarrett, 2017).

However, many healthcare systems do not screen for ACEs, or do so heterogeneously (Barnes et al., 2020; C. Bethell et al., 2021; C. D. Bethell et al., 2017; Finkelhor, 2018); and while there is widespread recognition that ACEs have critical impacts on child and adolescent health, universal screening is not recommended by major national bodies at this time. Given the lack of consensus on screening, NYC Health+Hospitals does not routinely screen for ACEs. Instead, the system focuses on fostering a supportive environment where trauma can be addressed effectively through trauma-informed care, supportive resource connections and trusting relationships.

Although many adolescents at NYC Health+Hospitals are exposed to at least one systemic stressor, resource constraints mean that primary care practices must identify the highest risk teenagers to connect to enhanced supports like care referrals, CHWs or community resources. Risk scores can help providers understand where to invest time and augment clinician judgment. However, no satisfactory definition of measurable risk has been identified for adolescent primary care.

Tracking measurable outcomes associated with ACEs may be a fruitful approach. An expansive definition of ACEs which incorporates both relational and social stressors as outlined in the Institute for Health Equity conceptual framework above provides a compelling base to define adverse physical and behavioral outcomes of concern for adolescent development. By predicting adverse outcomes associated with past or current ACEs exposure, and flagging adolescents at highest risk of multiple adverse outcomes, healthcare systems like NYC Health+Hospitals can connect adolescents and families to services and additional support.

Purpose Statement

This quantitative inquiry built on the ample body of literature on the impact of ACEs and social determinants on adolescent development, health and wellbeing, filling a crucial gap in the risk prediction literature. The study aimed to create and validate a single tool to systematically flag adolescents in the top 5% of risk of negative health outcomes associated with Adverse

Childhood Events (ACEs). Specifically, the algorithm used data from pediatric patients aged 12-17 who received primary care at NYC Health+Hospitals (H+H) between January 1 and December 31, 2022 to predict the number of ACEs-associated health outcomes documented in the EMR in 2023. The algorithm used clinical data from NYC Health + Hospitals, an urban public hospital system, augmented with geographically linked datasets from other City agencies and the CDC.

The outcome of interest was adverse outcomes associated with ACEs exposure as measured in the EMR in 2023, as outlined by key informants at H+H (see Appendix A, B). The domains of adverse events included high acute utilization; pregnancy and sexually transmitted infections; challenges in school; depression and suicidality; disordered eating; BMI trajectory; alcohol and drug use; abuse, trauma and related diagnoses; and acute utilization for chronic and behavioral disorders. These domains were selected because of their association with exposure to ACEs in the child and adolescent literature, their ability to be measured in the electronic medical record, and their selection as important domains of risk by key informant interviews with clinicians and patients at NYC Health+Hospitals collected before this study (see Appendix A, B). The algorithm selected predictors associated with this outcome from candidate variables in demographics, diagnoses, routine screenings, vital statistics, utilization, laboratory results and geographically matched socioeconomic status data.

Research Question

How accurately can the risk of adverse health outcomes for adolescents in primary care at a large urban safety net system be predicted by the prior year's electronic medical record and public data across demographic, SDOH, utilization and clinical domains?

Sub-Problems

1. Which demographic data in the EMR in 2022 predict adverse outcomes in 2023 among adolescents in primary care at NYC Health+Hospitals?

2. Which SDOH data recorded in the EMR in 2022 or pulled from geographically-matched neighborhood datasets predict adverse outcomes in 2023 among adolescents in primary care at NYC Health+Hospitals?
3. Which utilization data in the EMR in 2022 predict adverse outcomes in 2023 among adolescents in primary care at NYC Health+Hospitals?
4. Which clinical data in the EMR in 2022 predict adverse outcomes in 2023 among adolescents in primary care at NYC Health+Hospitals? Clinical data here refer to diagnoses, screenings, lab results and vitals.
5. Which adverse outcomes among adolescents in primary care can be measured in the EMR and aggregated into a single score for 2023? Adverse outcomes here refer to domains associated with ACE exposure, as identified from the literature and previous key informants within Health+Hospitals. These fall within the broad categories of increased acute utilization; pregnancy and sexually transmitted infections; challenges in school; depression and suicidality; disordered eating; BMI trajectory (top or bottom 5%); alcohol and drug use; abuse, trauma and related diagnoses; and utilization for key chronic and behavioral disorders.
6. How accurately can the risk of adverse health outcomes for adolescents in primary care at a large urban safety net system be predicted by the prior year's electronic medical record and public data across demographic, socioeconomic, utilization and clinical domains?

Hypotheses

H_o: The risk of adverse health outcomes for adolescents in primary care at a large urban safety net system can be predicted by the prior year's electronic medical record and public data across demographic, socioeconomic, utilization and clinical domains, at $\geq$70% AUROC overall and with $\geq$90% accuracy, $\geq$20% sensitivity and $\geq$30% positive predictive value at the top 5% high risk threshold.

H_a: The risk of adverse health outcomes for adolescents in primary care at a large urban safety net system cannot be predicted by the prior year's electronic medical record and public data across demographic, socioeconomic, utilization and clinical domains, at ≥70% AUROC overall and with ≥90% accuracy, ≥20% sensitivity and ≥30% positive predictive value at the top 5% high risk threshold.

Definitions

Adverse Childhood Events: As mentioned above, the CDC defines Adverse Childhood Events (ACEs) as events that inflict trauma which "can change brain development and affect how the body responds to stress" (CDC, 2021). Although Felitti et al identified a more limited set of relational risks in their landmark 1998 study, the field of ACEs research has since expanded to include traumatic experiences arising from other sources of adversity, including the lived experience of racism, witnessing violence within one's community and experiencing material deprivation within the home that is not mitigated by a trusted adult or social support (Felitti et al., 1998; Koita et al., 2018). ACEs are not measured directly in this study. Rather, this study seeks to identify and predict outcomes that manifest during adolescence which are associated with exposure to ACEs during childhood and adolescence, as a way to systematically bound adverse outcomes and risk within primary care.

Adverse Health Outcomes: An adverse health outcome is defined by Sherwin as "the causation, promotion, facilitation and/or exacerbation of a structural and/or functional abnormality, with the implication that the abnormality produced has the potential of lowering the quality of life, contributing to a disabling illness, or leading to a premature death" (Sherwin, 1983). In this study, adverse health outcomes are defined as outcomes which can be measured in the electronic medical record that are significantly associated with exposure to 1+ ACEs during childhood and adolescence, as identified by Nelson et al or as added and modified by a key informant (C. A. Nelson et al., 2020).

Electronic medical records: Electronic medical records (EMRs) are “digital versions of the paper charts in clinician offices, clinics, and hospitals [which]...contain notes and information collected by and for the clinicians in that office, clinic, or hospital…for diagnosis and treatment” *(What Are the Differences between Electronic Medical Records, Electronic Health Records, and Personal Health Records? | HealthIT.Gov, n.d.)*. In this study, EMR refers to the Epic electronic medical record instance implemented across NYC Health+Hospitals’ ambulatory and acute care facilities in 2019.

Primary care: The Institute of Medicine defines primary care as “the provision of integrated, accessible health care services by clinicians who are accountable for addressing a large majority of personal health care needs, developing a sustained partnership with patients, and practicing in the context of family and community” (Care et al., 1996). In this study, primary care refers principally to pediatric and adolescent care delivered by providers (MDs, NPs and PAs) across NYC Health+Hospitals’ 11 hospitals, 6 ambulatory care centers and 70+ Community Health Centers.

Acute Care: According to the World Health organization, acute care includes “emergency medicine, trauma care, pre-hospital emergency care, acute care surgery, critical care, urgent care and short-term inpatient stabilization” (Hirshon et al., 2013). In this study, acute care refers principally to care delivered during an emergency room visit or inpatient stay at NYC Health+Hospitals.

Safety net system: According to the Institute of Medicine, safety net systems are those which “deliver a significant level of health care and other needed services to uninsured, Medicaid and other vulnerable patients” (Wang et al., 2023). NYC Health+Hospitals is the nation’s largest safety net system. This study uses the IOM definition when referring to safety net systems.

Social Determinants of Health: According to the CDC’s Healthy People 2030, social determinants of health (SDOH) are “the conditions in the environments where people are born,

live, learn, work, play, worship, and age that affect a wide range of health, functioning, and quality-of-life outcomes and risks" (*Social Determinants of Health - Healthy People 2030 | Health.Gov*, n.d.). In this study, social determinants are captured directly in the electronic medical record via housing status variables, socially oriented billing z-codes for including food insecurity and lived experience of foster care. Unmeasured social determinants are brought in via geographic address match (homelessness and public NYCHA housing residency) and via neighborhood-level variables from the Social Vulnerability Index (2016-2020) (*CDC SVI Documentation 2020 | Place and Health | ATSDR*, 2022).

Demographics: According to Miriam Webster, demographics are "the statistical characteristics of human populations (such as age or income)" (*Definition of DEMOGRAPHIC*, 2023). In this study, demographics refer to patient variables from the EMR that include age, race/ethnicity, primary language, insurance status (Medicaid, uninsured, and commercial/other) and address (used for matching).

Diagnosis: According to the Cambridge Dictionary, a diagnosis is "a judgment about what a particular illness or problem is, made after examining it" (*Diagnosis*, 2023). In this study, diagnosis refers to an International Statistical Classification of Diseases (ICD-10) alpha-numeric classification of a given disease. This study uses two grouper methodologies to classify disease. The first is the Healthcare Cost & Utilization Project (HCUP)'s Clinical Classifications Software Refined (CCSR) system, which groups diagnoses into intelligible categories. The second is the Pediatric Comorbidity Index (PCI), which Sun et al developed in 2020 to capture the most important categories of morbidity among children and adolescents, by ICD-10 code (Sun et al., 2021). This study incorporates into its data all PCI categories, as well as CCSR categories for abuse, neglect, trauma, PTSD, social or material adversity, overdose, suicidality, experience of violence, enuresis, encopresis and educational difficulties. These are the diagnosis groups associated with ACEs exposure in the literature which did not have a dedicated PCI definition.

Healthcare screenings: According to the federal Office of Disease Prevention and Health Promotion, screenings are “medical tests that doctors use to check for diseases and health conditions before there are any signs or symptoms” (*Get Screened - MyHealthfinder | Health.Gov*, n.d.). Among adolescents visiting primary care, screenings are primarily questionnaires which are provided via paper, electronic tablet or verbal questioning during a visit. In this study, screenings refer primarily to standard and customized questionnaires done during the SSHADESS (Strengths, School, Home, Activities, Drugs, Emotions, Sexuality, Safety) assessment within the primary care visit (*Chapter 32: The SSHADESS Screening: A Strength-Based Psychosocial Assessment*, n.d.). Screening data in this study cover safety, food insecurity, foster care status, gender and sexual identity, alcohol and substance use, smoking status and school performance. They also include standardized clinical screenings for anxiety (GAD-7), depression (PHQ-9) and suicidality (CCRS). The one screening not included in SSHADESS is Body Mass Index (BMI), which was trended from 2022 to 2023 to identify the top and bottom 5% of patients’ whose BMI changed from Year 1 to Year 2 (aka BMI trajectory).

Utilization: According to the Encyclopedia of Behavioral Medicine, health care utilization is “the quantification or description of the use of services by persons for the purpose of preventing and curing health problems, promoting maintenance of health and well-being, or obtaining information about one’s health status and prognosis...[often measured by] the number of services used over a period of time” (Carrasquillo, 2013, p910). In this study, utilization refers to the number of visits undertaken by a patient in a given year to departments within the Emergency Department, Inpatient, Primary Care, Behavioral Health or other Outpatient settings. In some instances, this study separated acute utilization by regular ED/Inpatient and Psych ED/Inpatient visits, which refer to visits to specialized psychiatric wards within the ED and Inpatient settings.

Laboratory Results (“Labs”): According to Medline, “a laboratory (lab) test checks a sample of your blood, urine (pee), or other body fluid or tissue to learn about your health” (*How*

to Understand Your Lab Results, n.d.). This study included lab results in several composite definitions, such as sexually transmitted infections (STIs).

Prediction and Outcome Periods: According to Rijnbeek and Reps, any prediction problem is defined by a an attempt to use information collected during an observation window (prediction period) to anticipate whether an outcome will occur during a defined time-at-risk (outcome period) (Rijnbeek & Reps, 2021). In this study, the outcome period refers to events observed during 2023, as documented in the electronic medical record. In this study, the prediction period refers to one or more visits which occurred during 2022 within an H+H primary care department. Only data recorded in a patients' record as of their final visit during the prediction period was used in EMR-based predictor variables.

Delimitations

This inquiry was limited to risk prediction among adolescents in primary care in a large urban safety net system. This inquiry did not explore younger children (ages 0-11) or adults. The inquiry was limited to data collected during the course of care, or public agency data that were matched to enhance that clinical data. No primary patient data collection was done in this study. While resilience is a well-documented mediator in the impact of ACEs on health in adolescence and beyond, it is poorly documented in primary care and was thus outside the purview of this inquiry (Stirling et al., 2024). This inquiry did not cover the implementation of its risk score, beyond an exploration of possible directions and impacts in Discussion. It did not examine longitudinal risk measurement or directly investigate intergenerational trauma. While ultimately multi-disciplinary in nature due to the complexity of the topic, this inquiry was grounded in an epidemiologic approach, rather than a psychological, sociological, demographic or ethnographic approach.

Large machine learning models have in recent years increasingly been leveraged in healthcare, particularly in academic medical centers with a long history of unified electronic medical record adoption like Mass General. Complex AI and large language models are on the

precipice of transforming healthcare.The data infrastructure of Health+Hospitals during the study period could not support the development or deployment of an AI or deep learning model, due to computational intensity; and Health+Hospitals has not yet implemented a large language model within its firewall. As such, AI and deep learning methods fell outside the scope of this inquiry.

Assumptions

This inquiry included key assumptions. The first and perhaps most important assumption is that data found in the EMR or matched from public sources can truly approximate real underlying risk among primary care adolescents. The inquiry assumed that identifying adverse health outcomes associated during adolescence with exposure to ACEs is a valid method to bound adolescent risk. It assumed that outcomes associated with ACEs in the literature are truly associated with ACE exposure. It also assumed that using key informants to tailor a risk outcome would improve that risk outcome's ability to reflect local practice and needs.

Within the statistical modeling, there were additional key assumptions. This inquiry assumed that data from 2022 used to predict outcomes in 2023 would still be relevant in and beyond 2024. It assumed that regularized machine learning, in the form of a LASSO model, was an appropriate statistical method for the underlying data. Finally, it assumed that in instances in which it is impossible to measure data missingness, underlying missingness is non-differentially distributed. For example, diagnosis variables cannot differentiate between an adolescent who does not have a disease and an adolescent who has not been evaluated for the disease. This assumption is common across studies in this discipline, although it is likely to be violated. The treatment of missing data is further explored in Chapter 3 (Methods) and Chapter 5 (Discussion).

Significance of the Study

This study is responding to a gap within the risk prediction literature. While disease-specific risk scores have proliferated, very few generalized risk scores exist for the

adolescent primary care population. There is some precedent for socio-medical risk scoring in practice. The HealthySteps model, in which providers manually risk tier children aged 0-4 years, includes both medical and social risk factors (*Tiers and Core Components*, 2022). However, the unique needs of young children cannot be extrapolated to adolescent age ranges. The work of Bethell et al on the Integrated Child Risk Index does incorporate both social and medical risk using an ACEs perspective, and this work will guide the inquiry (C. Bethell et al., 2021). However, the score produced by this work, while of very high value to research, is too time-intensive to be able to be implemented in primary care practice, given the multiple questionnaires required to generate an Integrated Child Risk Index score. Additionally, any manual risk screening or tiering protocol requires staff time to implement. A risk score optimized for general adolescent primary care would provide targeting guidance without the need for extensive new screening protocols.

This study also addresses a gap in the risk prediction literature. Scores which predict readmission to the hospital for very sick pediatric populations exist, such as the model developed by Leary et al for children with complex conditions, but they are not generalizable across a full population without inpatient utilization (Leary et al., 2019). Pediatric comorbidity scores are limited in scope, incorporate only diagnosis data and fail to adequately capture social determinants of health (Sun et al., 2021). Models which predict an individual ACEs-associated outcome exist, such as Su et al's work predicting suicide using electronic medical record data, but they fail to capture a broad spectrum of adverse outcomes (Su et al., 2020). Models predicting medical cost are available, as outlined in Bellis et al's excellent systematic review of literature from North America and Europe, but they are seldom focused on pediatrics, and utilization among adults is distinct (Bellis et al., 2019). Models measuring past ACEs experience exist, but they are not tuned to predict future adverse outcomes and are often engineered for research rather than clinical practice (Bhattarai et al., 2021). This study aimed to create a comprehensive tool which addresses this gap in the literature.

A strong adolescent risk score has the potential to have real-world impact within primary care at H+H. In 2021, the Office of Population Health at H+H launched a program embedding Community Health Workers into primary care practices across the system. As part of the American Rescue Plan, New York City received $35 million to create the Public Health Corps, a group of Community Health Workers (CHWs) focused on low income New Yorkers. Senator Gillibrand is now seeking $55 billion in funding to replicate this model in communities across the country (NYC Health+Hospitals, 2022). By flagging adolescents and families who would most benefit from additional support, NYC H+H could ensure that it uses its CHW workforce to best effect, particularly as that workforce is reshaped by the evolving Medicaid policy landscape. This presents an unprecedented opportunity to connect families to resources that address social needs and reduce ACEs risk, a key driver of healthcare outcomes and overall well-being. Implementing an adult high risk score has allowed H+H to garner additional resources from the city and state, and to develop more targeted internal programming. A functional pediatric risk score would extend that advocacy power to our adolescent population.

Chapter 2: Review of the Literature

Introduction

The United States spends heavily to address pediatric disease burden after it occurs. For example, the U.S. spends an estimated $1.59 billion in asthma-related hospitalizations alone annually (Perry et al., 2019). In contrast, spending on primary care has hovered at approximately 5% of all healthcare dollars since 2003. Health in children and families is produced holistically, and requires holistic solutions rooted in a robust primary care system. How have we evolved in addressing social and physical health status holistically over time? And how could we better identify children at risk, to be able to better support them through adolescence? This literature review begins to answer these questions by describing adolescent risk and wellbeing, tracing the history of the field of social determinants of health and summarizing the evolving study of Adverse Childhood Experiences (ACEs).

This study covers the trajectory of two major disciplines: social determinants of health (SDOH), and adverse childhood experiences (ACEs). The Centers for Disease Control (CDC) defines social determinants of health as “conditions in the places where people live, learn, work, and play that affect a wide range of health and quality-of life-risks and outcomes” (*Social Determinants of Health | CDC*, 2021). For many families, social determinants of health such as poverty, lack of education or neighborhood environmental exposures can be a distal cause of traumatic stress, decreasing resources for resilience and self-efficacy. Social determinants of health in turn impact the prevalence of Adverse Childhood Experiences (ACEs), a proximal cause of childhood traumatic stress with impacts on development and health throughout the life course.

Because these literatures are wide, this literature review prioritized meta-analyses, systematic reviews and canonical sources cited many times in the literature. The SDOH and ACEs overviews leveraged a peer-reviewed literature search in Pubmed, the best of breed

engine for public health, social science and medical literature curated by the National Library of Medicine. Pubmed's MeSH key terms mapped poorly on the domains under study. As a result, this review used a combination of keyword searches (terms: "social determinants of health" [MeSH] *and* health outcome; "Adverse Childhood Experiences") and snowball searches (identifying key sources from bibliographies of seminal works in the field and those that cite the seminal works). Google Scholar was the primary engine used to identify key reports and white papers cited often in the literature, such as Alma Ata and the Rio Declaration. Searches were conducted iteratively.

In the below review of the literature, adolescent development is described, with particular focus on risk identification and management in primary care. Next, the broad historical trajectories of social determinants and ACEs research are traced. Finally, major themes in the literature are highlighted to establish the strategic direction of this inquiry. By tracing the history of these disciplines and identifying relevant gaps, this review aims to place this dissertation within an ongoing, broader conversation on how best to identify and assist adolescents in need.

Defining and Assessing Risk in Adolescence

Adolescent Development

Adolescence is a critical window of development in which social cognition accelerates, puberty blossoms and new areas of the brain develop (Blakemore & Mills, 2014). Adolescence is marked by transformation. It is a period in which children must navigate a complex set of new competencies, which include "becom[ing] independent…understand[ing] their changing social-sexual roles and sexual identity…chang[ing] their relationships with their parents…transform[ing] acquaintances into deeper friendships…focus[ing] their ambitions on their futures, and…transform[ing] their images of themselves to accommodate their physical and psychosocial changes" (Health et al., 2000). Although adolescence is defined as 12-17 in this inquiry, it actually refers to a window of development which may vary by individual and extend through early adulthood (Jaworska & MacQueen, 2015). During adolescence, most children go

through the five Tanner stages of puberty, including breast development and menarche among girls, and voice and genital changes for boys (Breehl & Caban, 2024). These hormonal transitions accompany a range of social development in this critical window.

Beyond the visible physical changes, there is a profound transformation happening in the brain, with implications for behavior, emotional wellbeing and decision-making (Lahey, 2021; Siegel, 2013). During adolescence, individuals gradually transition from concrete to abstract thinking (Jaworska & MacQueen, 2015). The brain trims unnecessary neural connections forged during childhood while strengthening areas devoted to reasoning. Although adolescents are capable of increasingly sophisticated thought, the prefrontal cortex is still developing through early adulthood (BALOCCHINI et al., 2013; National Research Council, 2011). Adolescence is a period of heightened risk seeking and emotional reactivity, with high impulsivity and sensitivity to rewards (Jaworska & MacQueen, 2015). This translates to more frequent risk-taking behaviors in the adolescent period, in human and animal species alike (Adolescence, 2011).

Adolescent Risk

Most parents would agree that adolescence is a time of heightened risky behaviors. As Balocchini et al noted, “adolescents and young adults are more likely than adults to binge drink, smoke cigarettes, have casual sex partners, engage in violent and other criminal behavior, and have fatal or serious accidents” (BALOCCHINI et al., 2013, p191). However, defining adolescent risk itself can be tricky. Frequently, the term “at risk” appears with “no consistent definition” and can encompass any manner of negative possibilities (Moore, 2006). Internationally, the biggest drivers of mortality among adolescents and young adults include injuries, interpersonal violence and self-harm (World Health organization, 2023). Injury, violence and self-harm were also the top three causes of death among adolescents aged 15-19 in the US from 2018-2022 (Centers for Disease Control and Prevention (CDC), 2024). However, anchoring on mortality, which is a low incidence outcome among adolescents, fails to grasp the full spectrum of risks which may impact adolescent wellbeing. The IOM’s Committee on the Science of Adolescence identified

“sexual risk-taking, substance use, illegal behavior, and risky driving” as the top areas of adolescent risk, with additional focus on “mental health risk” from disorders such as depression or disordered eating (National Research Council, 2011). This list, while important, is hardly comprehensive. A more granular definition of risk is required to measure it in primary care.

When searching for an enumeration of components for a broad definition of adolescent risk, this review identified a new gold standard for global adolescent risk measurement. In June of 2024, the WHO published the first global set of indicators to measure adolescent health and wellbeing. The domain of health behaviors and risks includes indicators for healthy weight; diet; exercise; alcohol, cannabis and nicotine use; early sexual initiation; use of contraceptives; exposure to bullying and exposure to violence (*The Adolescent Health Indicators Recommended by the Global Action for Measurement of Adolescent Health*, n.d.). Additional indicators cover health services use; HPV vaccination; HIV and STI prevalence; injuries; depression and anxiety; suicidality; educational opportunities; poverty; and food insecurity. This list was compiled for measurement at a national level, across large survey datasets, and cannot be translated to a clinical context. However, it enumerates key domains of adolescent risk which can be measured and are broadly applicable across populations and contexts. It also aligns closely to domains of concern from exposure to Adverse Childhood Experiences, as will be delineated subsequently in this literature review.

A more localized definition of adolescent risk might draw from the top challenges facing adolescents in New York City. The 2019 NYC Youth Risk Behavior Survey (YRBS) by the NYC Department of Health provides a population-representative overview of top stressors facing middle and high school students in New York City. Violence and bullying were big concerns. Among NYC middle and high school students in 2019, 22% reported getting into a physical fight, 10% had missed school because they felt unsafe, 17% had been bullied on school property, 11% had experienced physical dating violence and 8% reported carrying a weapon in the last 12 months (NYC Department of Health and Mental Hygiene, n.d.). Mental health was

also a top issue. Even before the pandemic, 36% reported feeling sad, and 9% reported a suicide attempt. Substance use was also important. While only 3% reported cigarette smoking, 15% reported vaping in the last month, 21% reported alcohol use in the last month, 9% reported binge drinking, 12% reported using a prescription pain medicine without a prescription, 5-6% reported trying heroin, cocaine, methamphetamines or ecstasy, and 18% reported using marijuana in the last 30 days. Sexual health was highlighted, with 45% of sexually active teenagers in New York City reporting not using a condom at last sexual contact, 21% not using contraception, and 14% having initiated sex before 13 years. Social health was another important theme. Twenty two percent reported some level of food insecurity, 12% reported that they had run away or been kicked out or abandoned, 21% reported parental incarceration and 13% reported a history of arrest. Finally, physical health issues also surfaced, with 25% reporting an asthma diagnosis and 14% reporting an asthma attack in the past 12 months. Only 20% of high school students surveyed reported getting enough sleep, 13% reported eating sufficient fruits and vegetables (4+ servings), and 14% reported getting 60+ minutes of physical activity a day. By prevalence, top areas of adolescent risk in New York City may therefore include sexual health; vaping, alcohol and substance use; violence and bullying; neglect and incarceration; mental health; food insecurity; asthma; insomnia; and diet and physical activity.

Although excessive TV viewing was captured in the YRBS, it did not ask about excessive phone or social media use, which has been linked to poor sleep quality, decreased mental health and increased risk of self-harm (Brautsch et al., 2023; Dibben et al., 2023; Khalaf et al., n.d.). Additionally, significant literature has covered the role of the pandemic in worsening social isolation and mental health challenges among adolescents, which is not captured in this 2019 data (Almeida et al., 2021; Loades et al., 2020; Wolf & Schmitz, 2024). It is reasonable to hypothesize that these effects hold true within New York City, even though they were not measured in the Youth Behavior Risk Survey.

New York City also has some very specific risks to health and wellbeing that vary by geography. Because of systematic underinvestment, redlining and environmental racism, there is a large disparity in neighborhood-level environmental exposures, poor quality housing, street safety, green space and heat-reducing tree cover (Environmental Justice NYC Initiative, 2024). Exposure to air pollution impacts lung health, development and chronic disease over the life course (*Air Pollution and Your Health*, n.d.). Racialized zoning has historically concentrated sources of air pollution in low income neighborhoods of color across the US and in New York City (Jbaily et al., 2022). Ozone is a gas produced by volatile organic compounds and nitrogen oxides which irritates the upper respiratory tract when inhaled (US EPA, 2016). Although East Harlem and the Upper East Side are adjacent neighborhoods, the risk of an ozone-related asthma ED visit is starkly different by neighborhood, with an estimated 21 annual visits per 100,000 children in the Upper East Side, and 214 visits per 100,000 children in East Harlem (*Health Impacts of Air Pollution*, n.d.). These neighborhoods also differ wildly on percent of homes with cracks or holes (6.1% vs. 20.2%), leaks (8.3% vs. 17.5%), mice or rats (37.6% vs. 54.8%) or cockroaches (17.5% vs. 39.2%) (*Cockroaches Data for NYC*, n.d.; *Housing Maintenance Data for NYC | Environment and Health Data Portal*, n.d.; *Mice and Rats Data for NYC*, n.d.). Life is hotter in East Harlem, with higher surface temperatures and less green cover (Interactive Heat Vulnerability Index., n.d.). Life is less secure, with 31.9% of residents in East Harlem living in poverty, as compared to 6.8% of the Upper East Side (*Economic Conditions Data for NYC*, n.d.). Life is also less safe, with 904.9 violence-related annual ED visits per 100,000 people in East Harlem, as compared to 126.3 visits in the Upper East Side (*Violence Data for NYC*, n.d.). As these neighbors illustrate, it is critical to consider the role of geography and place-based health when looking to describe and understand adolescent risk, particularly in New York City.

Finally, any enumeration of adolescent risk must consider the lived experience of discrimination on the growth and development of the individual. High school adolescents report

distress after experiencing racial or ethnic discrimination or prejudice from peers, teachers or their communities (Fisher et al., 2000). Black and Hispanic youth, particularly boys, are most likely to report experiencing discrimination from adults such as police, educators or store workers, and Asian youth are most likely to report discrimination from peers (Benner, 2017). Gibbons et al found that experience of discrimination was associated with higher risk sexual behavior among African American adolescents and lower resilience among young adults, with higher impacts associated with higher exposures (Gibbons et al., 2012). Benner's 2017 review found that racial and ethnic discrimination profoundly impacted adolescent development and adjustment. Adolescents experiencing racial or ethnic discrimination suffered increased anxiety, depression, stress and loneliness; reported more academic challenges; increased behavioral risk-taking such as substance use or shoplifting; and reported more somatic complaints and insomnia (Benner, 2017). Adolescents whose parents experienced discrimination were also adversely affected. Resilience in the form of support from friends, parents, trusted adults and community institutions blunted these effects.

Among LGBTQ+ youth, experience of discrimination also impacts development and long-term health. Using the minority stress model, Russel and Fish illustrate how youth coming out in adolescence may experience discrimination during a crucial window of social development, leading to higher rates of depression, anxiety and substance use among LGBTQ+ youth (Russell & Fish, 2016). In 2022, 73% of LGBTQ youth experienced anxiety, 58% experienced depression and 37% reported receiving threats or harm because of their identity, according to a nationally representative survey by the Trevor Project (*2022 National Survey on LGBTQ Youth Mental Health*, n.d.). Risk of a suicide attempt is three times higher among LGBTQ youth than cisgendered or heterosexual youth, and rates are highest among Black, Middle Eastern and Indigenous youth (*2022 National Survey on LGBTQ Youth Mental Health*, n.d.; Allred, 2024). However, growing up in a supportive family or community halves the risk of suicide attempt (Allred, 2024). According to 2019 estimates from the NYC Youth Risk Behavior

Survey, 2% of middle and high school students in New York identified as transgender, and 13% of students reported identifying as gay or bisexual (NYC Department of Health and Mental Hygiene, n.d.). These are important risks to consider.

Minority and LGBTQ+ adolescents are both at increased risk of experiencing discrimination, and many youth of color who identify as LGBTQ+ experience intersectional minority stress (Salerno, 2023). Data truly capturing experience of discrimination and sources of familial or social resilience and support are lacking in the electronic medical record. Place, race, language, sexual and gender identity variables are proxy measures which may allow us to identify adolescents at risk of discrimination. Any assessment of adolescent risk must take these socio-geographic exposures into account, however imperfectly we may be able to measure them.

Role of Primary Care

Primary care has an important role in supporting adolescents as they develop. Although adolescents engage in outpatient care less often than adults or young children, they should receive annual preventive care visits (L & K, 2023; Riley et al., 2018). Within primary care, adolescents should receive “family planning and reproductive health services, diagnosis and treatment of STIs and HIV, mental health counseling and treatment, and substance abuse counseling and treatment” (Committee on Adolescence, American Academy of Pediatrics, 2008). In Bright Futures: Guidelines for Health Supervision of Infants, Children and Adolescents (4th edition), the American Academy of Pediatrics recommends that in addition to routine physical exams and immunizations, adolescents should receive screenings for blood pressure, BMI, vision and hearing; anemia, tuberculosis, HIV and sexually transmitted infections; developmental progression; behavioral, social and emotional wellbeing; alcohol, tobacco and drug use; and depression and suicidality (Bright Futures/American Academy of Pediatrics, 2023). The mnemonic SSHADESS is commonly referenced to identify key psychosocial screening areas for adolescents in primary care, covering strengths, school, home, activities,

drugs, emotions/eating, sexuality and safety (Klein et al., 2020). Primary care providers should also offer directional guidance on healthy decision-making, employing a strength-based approach to cultivate resilience and confidence in adolescent patients as they navigate new skills (Riley et al., 2018). Comprehensive primary care should be adolescent-centered, cultivating trust through cultural humility, participatory decision-making, non-judgmental listening, respect for autonomy and protection of confidentiality (Alderman et al., 2019; Riley et al., 2018).

For patients who are struggling, primary care providers can be a key gateway to enhanced services (Burke Harris, 2018; Stirling et al., 2024). Many primary care providers work with families to document learning disabilities or mental health needs and to navigate the landscape of school-based services (Fitch, 2024). Trained providers can manage mild to moderate depression, anxiety and ADHD in the primary care setting, particularly with the support of collaborative care or embedded behavioral healthcare providers (Burkhart et al., 2020; Child and Adolescent Mental Health Coalition, 2023). Some primary care practices offer counseling within the medical home itself. Others help patients find in-network or community-based providers to address mental health or substance use issues that exceed primary care capacity. Social referral platforms such as UniteUs or FindHelp facilitate this process within the electronic medical record itself (Calvo-Friedman et al., 2023). Primary care providers can provide long-acting contraceptives in the primary care setting, treat sexually transmitted infections, screen patients for high risk sexual behaviors and connect patients to additional sexual health services (Klein et al., 2020). Primary care providers can identify patients who need specialty care to manage a chronic condition like juvenile diabetes or lupus. In cases of maltreatment, neglect, abuse or foster care placement, primary care providers can increase monitoring, connect patients to enhanced mental health services or protective services, and support resilience to provide tertiary prevention of harm (Stirling et al., 2024).

These activities have a cost in staff time. Providing a full suite of screenings and interventions in 12-18 minutes is unrealistic, and national estimates suggest adolescents aren't screened for risky behaviors in 50-85% of primary care interactions (Riley et al., 2018). While models like the patient-centered medical home are intended to provide additional reimbursement to support enhanced care (Bailit et al., 2010), resources remained stretched. High quality telehealth options can improve access to care for low-income adolescents, but while they may improve efficiency or reduce no-show, they are not a silver bullet for staffing challenges (Child and Adolescent Mental Health Coalition, 2023). Staffing models including community health workers, social workers or care managers can relieve pressure on primary care providers and support holistic care (Sandhu et al., 2021; Scott et al., 2018). However, reimbursement for community health workers and ancillary staff vary by state (National Academy for State Health Policy, 2024). Particularly in practices serving lower income populations, need often outstrips supply. Predictive models have been used elsewhere in healthcare to target scarce resources to those who might benefit most and may offer promise in prioritizing demand in pediatrics to better address the social determinants of health.

Social Determinants of Health

Evolution of the Field

The evolution of public health is rooted in the social determinants of health. In 1854, John Snow advised the municipal government to remove the handle of a tainted water pump in a low-income area where residents were all required to use the same tainted public water source, slowing the infamous south London cholera outbreak (Paneth, 2004). This famous anecdote in the annals of epidemiology is a key example of the early importance of place-based interventions and population-level context. Throughout the nineteenth and early twentieth centuries, health commissions recognized the impact of poor housing quality, overcrowding, environmental exposure, contamination and poverty itself on the production of health (Fairchild

et al., 2010). Public health practitioners were often social advocates and organizers like Alice Hamilton and Florence Kelley, working to address the spread of communicable disease at its upstream source. However, in the twentieth century, as communicable disease outbreaks responded to sanitation and housing improvements, the focus shifted to laboratory research and chronic disease control (Yong, 2021). In the second half of the twentieth century, the expanding power of the medical system re-oriented public health conversations from the community to the individual (Frieden, 2010). Medical interventions overtook policy interventions in the public health playbook, de-prioritizing social determinants over biological determinants of health.

New Policy Frames

The Declaration of Alma Ata catalyzed a re-thinking of the medical model, re-inserting the importance of community and context on the production of health and foregrounding the importance of social determinants. The Declaration of Alma Ata was signed in 1978 during the International Conference on Primary Health Care, which was convened by UNICEF and the World Health organization (WHO) (“Declaration of Alma-Ata International Conference on Primary Health Care, Alma-Ata, USSR, 6–12 September 1978,” 2004). The conference included 134 countries and 67 international organizations, nearly all of whom signed onto the declaration and its goal of “Health for All” by the year 2000. It served as a clarion call for enhanced, global access to primary care that treated both downstream health conditions and upstream social factors. Although the exact phrase “social determinants” doesn't appear, the intent was analogous. Alma Ata aimed to establish world-wide recognition of the interlinked nature of social and economic development with health inequalities and health outcomes. It also argued that primary care could play an important role in addressing both social factors and health outcomes. Although the ensuing four decades saw continued overreliance on disease-specific models of treatment and prevention, Alma Ata offered a vision of what holistic healthcare could be. Alma Ata is still cited in both human rights and SDOH frames of argument in public health. It

undergirded the millennium development goals (Exworthy, 2008) and remained a key framework by which community health progress (or lack thereof) was measured (Rifkin, 2018).

Following Alma Ata, the phrase “social determinants of health” gained prominence. A search of the National Library of Medicine for articles with the phrase “social determinants of health” in the title yields 83 articles published in the 1970’s, 471 published in the 1980s, 1,522 published in the 1990s, 3,497 published from 2000-2009, and 29,990 published since 2010. Another milestone came two decades after Alma Ata with the Marmot report, “Social Determinants of Health Inequalities” (Marmot, 2005). Sir Michael Marmot chaired the WHO Commission on the Social Determinants of Health in 2000. Marmot summarized the then-emerging body of evidence showing the impact of social status on communicable disease, non-communicable disease and mortality. The report outlined the impact of SDOH on rates of violence, exposure to environmental contaminants, opportunity for education, access to food and nutrition and other key cornerstones of health. Marmot identified how social determinants are socially constructed and as such are amenable to policy interventions. His report served as the cornerstone to the emerging field of SDOH and has been cited more than 5,000 times in the peer reviewed literature. The “Healthy People 2020” goals (written in 2010) were anchored in social determinants under the five domains of economic stability, education, health and health care, neighborhood and built environment, and social and community context, building directly on the work of Marmot and the Commission (*Social Determinants of Health | Healthy People 2020*, n.d.).

By 2011, when the WHO convened the World Conference on Social Determinants of Health, the importance of social determinants was widely acknowledged. However, this knowledge did not translate into more equitable access to care and outcomes. The Social Determinants of Health conference took aim at this issue with the Rio Declaration, which expanded Alma Ata’s goals and tied them to the concept of equity (*Rio Political Declaration on Social Determinants of Health | National Collaborating Centre for Determinants of Health*, n.d.).

The Declaration moved beyond health as a human right and declared that accessible, available, acceptable, affordable and high quality healthcare was a human right. The Declaration was rooted in a health equity approach. It argued that social determinants are essential to the production or amelioration of health disparities. It called for a Health in All approach to multi-sectoral policy making. Crucially for this inquiry and the Adverse Childhood Experiences field developing simultaneously, the social determinants targeted by the Rio declaration included "early years' experiences, education, economic status, employment and decent work, housing and environment, and effective systems of preventing and treating ill health" (*Rio Political Declaration on Social Determinants of Health | National Collaborating Centre for Determinants of Health*, n.d., p. 1). The Rio Declaration called upon governments to embed health equity into policy design and to “further reorient the health sector towards reducing health inequities,” a call to which this inquiry aims to respond.

SDOH in Medical Practice

In the decade since the Rio Declaration, the importance of social determinants has become enshrined in public health practice and pedagogy. However, the medical system is still struggling to effectively bring social determinants into the care setting. In 2016, the Institute of Medicine released “A Framework for Educating Health Professionals to Address the Social Determinants of Health,” recognizing the increased need for integration of SDOH into medical practice (Committee on Educating Health Professionals to Address the Social Determinants of Health et al., 2016). The Affordable Care Act introduced required community health needs assessments (CHNAs), to be done by non-profit healthcare systems every three years, in an effort to bring community health and social determinant needs closer to the healthcare system (*Community Health Needs Assessment for Charitable Hospital organizations - Section 501(r)(3) | Internal Revenue Service*, n.d.).

A key aspect of integrating social determinants of health into primary care has revolved around screening, characterized by heterogeneous and uneven adoption. The Institute of

Medicine recommended collection of housing status, income and education in EMR data in 2015, but data collection has not been routinely implemented (Medicine et al., 2015). The Institute for Medicaid Innovation identified nine validated tools for SDOH screening and recommended standard ICD-10 z-codes to document identified social needs but stopped short of recommending a single tool (Lekisha Daniel-Robinson & Jennifer E. Moore, 2019). The American Academy of Pediatrics' has not endorsed any single tool but offers multiple SDOH screeners through the Screening Technical Assistance and Resource Center. Some systems have created their own tools, tailored to their unique populations and available community resources (Calvo-Friedman et al., 2023; Hunt, 2021). The American Academy of Pediatrics Pilots such as the upcoming New York State Medicaid 1115 Waiver are exploring a mandated single screener (Office of Health Insurance Programs, 2021). Screening workflows and roles also vary. A 2017 review across 67 SDOH intervention studies found that social needs screening was happening heterogeneously, with screening done by a mix of providers, social workers, community health workers and volunteers (Gottlieb et al., 2017). The Center for Medicare and Medicaid Innovation launched a five-year SDOH screening and intervention challenge grant in 2017 which is yielding concrete guidance for systems looking to expand screening now.

Care models to address social determinants of health are also heterogeneous and expanding rapidly. SDOH interventions in the ambulatory care setting have been shown to improve chronic disease control, mental health, healthy behaviors and overall quality of life, including among pediatric populations (Gottlieb et al., 2017). States like California and Oregon have used Medicaid pilots to fund housing navigation, food support, care coordination and legal aid (Alderwick et al., 2019), and New York is piloting a new 1115 waiver targeting social services for Medicaid patients outside the care setting. In the pediatrics space, models focused on young children and families offer approaches which could be extended to older populations. For example, the University of Pennsylvania created resource and implementation guides for

addressing social determinants in pediatrics for their HealthySteps program (Piotrowski et al., 2009). The CLEAR collaboration has outlined strategies like targeted social determinant screening, social service referrals, hiring social navigators and benefit enrollers, reducing barriers to access by providing childcare and transportation, adding interpreter services, extending clinic hours, locating clinics within patient communities and adjusting panel size to accommodate more complex case mixes to target the upstream barriers to improved health and healthcare, building on well-established models for complex adult patients such as the Camden Coalition (Andermann, 2016; Truchil et al., 2018). The COVID-19 pandemic highlighted the impact of social determinants on health and many health systems more purposefully integrated social services like legal help, food assistance, benefits enrollment and direct cash transfers into clinical practice (Clapp et al., 2020).

Identifying and effectively addressing social determinants of health is a powerful opportunity to improve the production of health within the healthcare system. However, the impact of adverse social determinants on health varies by family and individual. One way to understand why some families thrive and some face terrible difficulties in adverse conditions is the study of traumatic stress produced by Adverse Childhood Experiences.

Adverse Childhood Experiences and Their Impact Through the Life Course

Evolution of the Field

Adverse Childhood Experiences (ACEs) research examines the impact of traumatic stress during childhood on the individual throughout the life course. ACEs research documents the unique and crucial role of trauma on childhood development, including endocrine, neurological and immunological systems and future health-related behaviors (Burke Harris, 2018). Over time, the list of commonly cited ACEs expanded from household or familial trauma such as parental neglect and death to include traumatic stress stemming from social determinants, including exposure to racism, violence and poverty.

Although the 20th century saw diverse research on child maltreatment and neglect, the genesis of the ACE-specific field can be traced to the landmark 1998 study “Relationship of Childhood Abuse and Household Dysfunction to Many of the Leading Causes of Death in Adults: The Adverse Childhood Experiences (ACE) Study,” in which Felitti et al first studied the impact of cumulative ACEs on adult health outcomes (Felitti et al., 1998). It has subsequently been cited over 16,000 times. In this baseline paper of the Adverse Childhood Experiences (ACE) Study, Felitti et al surveyed 9,508 Kaiser Permanente San Diego members about their ACE exposures across seven domains and evaluated the strength of association and the dose-response relationship between single/cumulative ACEs, adult risk behaviors and disease states. The authors found that as ACE exposures increased, risk of smoking, severe obesity, physical inactivity, depressed mood, suicide attempts, alcoholism, substance use, >=50 sexual partners, and history of one or more sexually transmitted infections also increased significantly. They also found significant relationships with diabetes, ischemic heart disease, emphysema/bronchitis, hepatitis/jaundice, skeletal fracture and self-reported poor health. The authors evaluated the generalizability of their sample to the US population using national surveys and conducted sensitivity analyses to evaluate the impact of non-response bias.

In the quarter century since Felitti, the ACEs field has produced multiple systematic reviews quantifying the impact of childhood trauma on adult health outcomes. For example, Hughes et al’s 2017 Lancet systematic review combined populations from 37 studies to evaluate the impact of having four or more documented ACEs on adult health outcomes, as compared to having no documented ACEs (Hughes et al., 2017). This methodologically rigorous study followed the gold standard PRISMA reporting guidelines (Shamseer et al., 2015). The meta-analysis included a large sample size (N=253,719) and 23 outcomes of interest. The outcomes with the strongest relationship to four or more ACES included problematic drug use, interpersonal and self-directed violence, sexual risk taking and mental health. Moderate associations were observed for use of alcohol or tobacco, poor reported health, heart disease,

cancer and respiratory disease. However, all outcomes studied were significantly associated with multiple ACE exposures. This article represented the first systematic review of the impact of multiple ACEs on health and consolidated the then-emerging consensus that ACEs have wide-ranging impacts on adult health not just for risky behaviors but also for chronic conditions. These findings have been replicated across multiple reviews and systematic analyses for adults (Boullier & Blair, 2018; Herzog & Schmahl, 2018; Kalmakis & Chandler, 2015; Petruccelli et al., 2019).

In 2019, Bellis and Hughes established that this relationship also held for cost. Published in the Lancet, “Life course health consequences and associated annual costs of adverse childhood experiences across Europe and North America: A systematic review and meta-analysis” estimated the cost of 1 or 2+ ACE exposures annually in North America and Europe (Bellis et al., 2019). Pooling data from 23 studies, they estimated the population-attributable fraction (PAF) due to ACE exposure for high risk behavioral conditions, including obesity, harmful alcohol use, illicit drug use and smoking, and for non-communicable diseases including cardiovascular disease, cancer, diabetes, respiratory disease, depression and anxiety. The authors found that a quarter of harmful alcohol use could be attributable to ACEs, as well as 30% of anxiety and 40% of depression. While the relationship between 2+ ACES and the conditions studied was attenuated as compared to Hughes et al’s findings for 4+ ACES, the economic findings were staggering. The authors estimated that ACES are responsible for $748 billion annually in costs in North America, which is approximately 3.55% of GDP. Reducing ACES exposure by 10% would save $56 billion across North America annually, the authors found. The authors’ strategy to capture the financial impact of multiple ACES was a novel one and was crafted explicitly with an aim towards influencing economic and social policy.

Although economic and social policies to target ACEs are still developing, funding for US research and data collection at the federal level has increased in recent years. The Centers for Disease Control (CDC) underwrote the original Kaiser Permanente/Felitti study, built a major

portfolio around violence prevention anchored in the science of ACEs (CDC, 2021), published an MMWR outlining the impact of ACEs on adult health across 25 states in 2019 (Merrick et al., 2019), and released a funding opportunity (RFA-CE-20-003) to support ACEs prevention in 2019/2020. Since 2009, 48 states included an ACE survey module in at least one Behavior Risk Factor Surveillance System (BRFSS) survey implementation, recognizing the importance of tracking these upstream factors to the production of health across the United States (*Behavioral Risk Factor Surveillance System ACE Data |Violence Prevention|Injury Center|CDC*, 2022). In 2016, the National Survey of Children's Health also incorporated an ACEs module (C. D. Bethell et al., 2017) and it is now possible to quantify ACEs exposure nationally (Women and Children's Health, 2020).

This investment in research has yielded strong causal evidence. Across studies, the relationship between ACES, health risk factors and adult disease have been shown to satisfy five categories of Hill's causal criteria used to suggest causality in epidemiology. These criteria are: strength of association; consistency across populations; specificity of effect to exposure (less relevant for population-wide and multi-factorial effects); temporality, or cause preceding effect; dose-response relationship, or more exposure leading to greater prevalence or intensity of effect; plausibility, or proposed mechanism by which cause might lead to effect; coherence, or replicating effects in the laboratory setting; experiment, or observing effects in trials where possible; and analogy, or similar causes producing similar effects (Rothman & Greenland, 2005). Across the studies included in this review, research on the impact of ACEs on health outcomes and behaviors throughout the life course currently suggests strength, consistency, temporality, dose-response and coherence.

Biological Mechanisms for the Impact of ACEs on Health

Researchers have shown that ACEs exposure is associated with traumatic physical stress. In the last twenty years, scientists have uncovered many of the physical mechanisms whereby trauma impacts the body. In their authoritative 2012 review, Danese and McEwen

collated the literature on the biological impact of adverse childhood experiences, exploring the role of allostatic load (Danese & McEwen, 2012). Allostatic load is in essence a quantification of chronic stress in the body. The authors detail how chronic exposure to adversity changes the brain structurally and functionally, including attention and emotional regulation, fear conditioning and memory. These changes make activation in response to stress more likely, creating a feedback loop. Exposure to chronic stress increases cortisol production via the endocrine system and induces chronic inflammation via the immune system. Chronic stress exposure has been shown to increase risk of chronic disease; insulin resistance, metabolic syndrome and diabetes; and cellular aging, cognitive decline and dementia. Exposure to chronic stress during childhood impacts development of the nervous, endocrine and immune systems, creating life-long consequences in the body and the brain.

Exposure to traumatic stress has impacts in the genome itself. A systematic review by Parade et al found that childhood adversity was significantly associated with increased DNA methylation across the genome and in key genes impacting glucocorticoid receptors, oxytocin and genes involved in systemic inflammation (Parade et al., 2021). These methylations impact social behavior, attachment, stress response and internalizing behavioral problems. Childhood adversity (but not maltreatment) was also associated with advanced molecular aging, both in children aged 8-16 and among adults. Finally, the reviewers highlighted that the epigenetic impacts of prenatal exposures and stress create intergenerational transmission of toxic stress at a genome level (Babenko et al., 2015). In their review of cardiovascular disease and mental health in children and adolescents, Xu et al also found evidence for transmission of cardiometabolic risk across generations due to exposure to adversity and stress (L. Xu et al., 2022).

ACEs and Health Outcomes During Childhood and Adolescence

The literature documenting the association of ACEs exposure and adult physical and behavioral health outcomes is robust and mature. However, adolescents are not small adults.

Adolescents have unique developmental needs and risk factors, but the scholarship around the impact of ACEs on that development during adolescence is newer. For example, a 2019 review by Petruccelli et al found only 12 studies examining the health impact of ACEs before age 18 through 2016, as compared to 84 studies examining health effects among adults (Petruccelli et al., 2019).

There is an expanding body of research suggesting that the metabolic, immunological and neurological impacts of ACEs exposure create observable health and behavior impacts during childhood and adolescence itself (Burke Harris, 2018). ACEs impact adolescent mental health, resilience, decision-making and development of higher-level thinking among adolescents. Significant associations have been observed between exposure to ACEs and adolescent ADHD (Bomysoad & Francis, 2020), anxiety (Balistreri & Alvira-Hammond, 2016; Bomysoad & Francis, 2020; Elmore & Crouch, 2020), asthma (Bellis et al., 2018), depression (Anderson et al., 2022; Blum et al., 2019; Bomysoad & Francis, 2020; Brockie et al., 2015; Elmore et al., 2020; Elmore & Crouch, 2020), overall poor mental health (Baldwin et al., 2021; C. Bethell et al., 2019; Boch et al., 2019; Folk, Ramos, et al., 2021), overall poor physical health (Baldwin et al., 2021; Balistreri & Alvira-Hammond, 2016; Bellis et al., 2018) and chronic comorbidity (Kerker et al., 2015), risky sexual behavior (Brown et al., 2015; Hillis et al., 2001; Lin et al., 2011; Richter et al., 2014), PTSD (Brockie et al., 2015), substance use (Bomysoad & Francis, 2020; Brockie et al., 2015; Folk, Kemp, et al., 2021), suicidality (Anderson et al., 2022; Brockie et al., 2015; Folk, Kemp, et al., 2021), learning or behavioral problems (Burke et al., 2011), justice system contact (Graf et al., 2021), and inadequate healthcare access and utilization (Baron-Lee et al., 2015; Berg et al., 2018; C. Bethell et al., 2022). These myriad outcomes reflect that ACEs impact the physical, behavioral and social health of children and adolescents (C. A. Nelson et al., 2020).

Systematic reviews identifying the impact of ACEs on adolescent health, writ large, are starting to emerge. Many of these reviews are focused on a particular sub-population or

endpoint, with more development happening around justice-involvement, which is in and of itself a social determinant of health. A 2021 systematic review looking at the association of ACEs experience with adolescent and young adult justice system contact found a dose-response relationship (Graf et al., 2021). A scoping review from Folk et al in 2021 found that among justice-involved youth, ACE exposure was significantly associated with mental health issues, substance use, pregnancy and academic challenges, across eight studies (Folk, Kemp, et al., 2021). Draxler's 2022 narrative synthesis framework review found a relationship between ACEs and treatment non-adherence, across six studies (Draxler & Ruppar, 2022). A systematic review and meta-analysis by Morgan et al in 2021 found a 63% reduction in the likelihood of having high psychological resilience in youths exposed to one or more ACEs, using data pooled across three studies, and noted that six additional studies identified a negative relationship between ACE exposure and resilience (Morgan et al., 2021). In the sexual health space, Moussaoui's 2022 systematic review found a significant association across 19 studies between ACEs exposure and dysmenorrhea, but not necessarily pelvic pain or dyspareunia (Moussaoui & Grover, 2022). Oh et al's 2018 systematic review found that ACEs were associated with asthma, infection, somatic issues such as stomach ache, headache or insomnia and cognitive delay, with a subset of ACEs also associated with weight gain and increased cortisol levels (Oh et al., 2018). Additional reviews included pediatric and adult populations, but did not adequately disaggregate these very different populations to identify health outcomes which emerge during childhood and adolescence itself (Petruccelli et al., 2019).

Evidence is also emerging about the national burden of ACEs during childhood and adolescence. Per the National Survey of Children's Health (NSCH) has identified that 33% of children nationally have experienced at least one ACE, and 14% of children experienced 2+ ACEs (Women and Children's Health, 2020), although prevalence was found to be lower (9.8%) among adolescents aged 15-19 in New York (*Explore Adverse Childhood Experiences in New York | AHR*, n.d.). Rates were highest among lower-income families and Black or American

Indian families. The NSCH also found a cross-sectional dose-response relationship between number of ACEs and depression, anxiety, special healthcare needs, poorly rated physical health, behavioral issues and difficulty with friendship. These findings may be an underestimate, as parents may be uncomfortable disclosing household adversity, particularly if they have lived experience of racism (Stirling et al., 2024). These findings suggest ACEs are critical to consider when trying to identify families and children requiring additional support from the medical system.

ACES Screening

Given the focus on integrating social determinants of health into clinical practice in the last twenty years, it is not surprising that there are nascent efforts to evaluate and treat the downstream effects of Adverse Childhood Experiences in primary care. One strategy is screening for ACEs exposure. There has been significant interest in ACEs screening since the Felitti study. The American Academy of Pediatrics recommended that providers consider screening adults and children for traumatic stress in 2012 but stopped short of recommending universal screening (Barnes et al., 2020). The American Academy of Family Physicians currently recommends that providers be conversant on the impact of ACEs on patient health and the need for trauma-informed care but called for more research on ACE screening in lieu of a recommendation. The United States Preventive Services Task Force determined in 2024 that there was insufficient evidence to recommend primary care intervention to prevent child maltreatment, given the uneven literature on the benefits of ACE screening and intervention pathways. This literature review covers ACEs screening tools and practices to establish that screening universally for ACEs in primary care is not a viable alternative to risk identification in current practice.

ACE screening efforts have focused on both parents and children. In “Methods to Assess Adverse Childhood Experiences of Children and Families: Toward Approaches to Promote Child Well-being in Policy and Practice,” Bethell et al systematically reviewed the

literature for ACE assessment methods (C. D. Bethell et al., 2017). They compared the 14 methods identified by domains included, setting of administration and methodology. They found comparable domains across most instruments, with 12 including the core domains of sexual, emotional and physical abuse and the majority including divorce (10), physical neglect (11) and emotional neglect (9). Tools that went beyond the original CDC/Kaiser study questionnaire also included neighborhood violence (6), school bullying (4), experiencing racism or discrimination (4) and death of a parent or caregiver (4). This review also evaluated in more depth the validity of the ACEs module of the 2016 National Survey of Children's Health (NSCH), finding that questions included in the module represented a single construct (childhood adversity), did not present acceptability issues, and had no significant redundancy.

Measurement tools tailored to primary care are still developing. Bethell et al found that five of 14 ACE exposure assessment tools reviewed in 2017 had been developed for the clinical setting (C. D. Bethell et al., 2017). No one assessment tool was ubiquitous and accepted as routine practice across clinical applications. Only one of the five clinical ACE tools was developed to measure child or adolescent responses, rather than parental ACEs. The lack of uniform ACEs screening amongst children and adolescents makes uniform screening recommendations difficult to develop and implement. A concurrent study by Wade et al piloted a much simplified two item screener, from the 11-item Behavioral Risk Factor Surveillance System (BRFSS) ACE screener using BRFSS data from 2011-2012 (Wade et al., 2017). The goal of the study was to create a screener with acceptable validity that would be rapid to deliver in a clinical setting with adult patients. The study ultimately selected parental alcohol abuse and emotional abuse as the two items in the tested screener. The two item screener performed as well in logistic regression in predicting the outcomes of interest as the 11 item screener, after adjusting for patient-level social and demographic factors. Although this screener was not included in the Bethell review, it had high potential for clinical application due to its parsimony and strong performance. Subsequent to the Bethell review, the PEARLs tool was developed to capture

ACEs and “related life events” including four social determinants of health domains (food insecurity, housing instability, exposure to violence, experience of discrimination) (Koita et al., 2018). This instrument shows promise, particularly when leveraged in the context of a trusting longitudinal relationship (Long et al., 2022), but has yet to be widely implemented.

ACEs screening practices and workflows have been shown to be heterogeneous across pediatric primary care settings. A 2016 study by Kerker found that although more than 90% of providers agreed that chronic stress in childhood impacted future stress management and impacted brain development, three in ten providers didn’t screen for any ACEs, and only 4% usually screened across the 7 major ACE domains. Clinical implementation has lagged due to heterogeneous screening instruments, lack of clear guidelines for when, whom and how to screen in clinical practice, and unclear intervention pathways (Barnes et al., 2020). To date, the American Academy of Pediatrics has not endorsed universal ACE screening in pediatrics (Stirling et al., 2024).

One implementation challenge is the lack of a clear, validated intervention pathway. As Finkelhor et al warned in 2018, implementing universal screening in primary care without first establishing intervention workflows to address identified needs has the potential to increase distress and decrease trust (Finkelhor, 2018; Stirling et al., 2024). While multiple toolkits have been produced, primarily focusing on supporting family resilience with economic policies and community programming, coherent recommendations for treatment of positive ACE screens in primary care are lacking (Folk, Kemp, et al., 2021; ortiz & Sibinga, 2017), and development of evidence-based primary care interventions for childhood adversity has been identified as a priority area for research (Shonkoff et al., 2021). A systematic review by Loveday et al assessing whether ACE screening in children aged 0-12 improved adversity-related referrals or service uptake found limited evidence of impact (Loveday et al., 2022). Screening may progress absent AAP, AAFP or USPSTF recommendation if incentivized through insurance, such as California’s ACE-Aware Medi-Cal program (ACE Aware, 2024). However, uptake of ACEs

screening is unlikely to be universal in the near term. It is possible that statistical models which estimate risk of ACE exposure or ACE-related outcomes could fill an information gap without large-scale changes in clinician practice.

Risk Prediction in Pediatrics and Adolescent Medicine

As established above, universal screening for ACEs is not current practice in primary care, and it is not recommended by major bodies such as the United States Preventive Services Task Force. However, identifying patients who need the most support to develop into healthy adults remains critical. As outlined by Shonkoff et al in their 2020 overview of the psychosocial impact of ACEs on development, “Measuring demographic risk factors or adverse childhood experiences (e.g., by calculated adverse childhood experience scores) at a population level can provide valuable data for policy makers, but the identification of individual risk and allocation of resources would be enhanced substantially by direct measures of child health and development within the context of primary health care to determine priorities for intervention and match specific services to identified needs” (Shonkoff et al., 2021). Risk prediction algorithms may play a role, using the proxy data that are currently collected in the electronic medical record, to approximate true risk and flag patients for enhanced screenings or interventions.

Such a score does not currently exist in the literature. Although adolescents are subject to myriad risks across emotional, behavioral and physical loci, most risk scores developed in pediatric populations are extremely narrow. There are focused risk scores for specific events or conditions such as obesity (Gerald et al., 1994), diabetes (Castensøe-Seidenfaden et al., 2017), asthma (Eum et al., 2019), cancer (Tolkkinen et al., 2018), injury (Bradbury et al., 1999), surgery (Cromhout et al., 2022), hospital readmission (Nacht et al., 2022; Sills et al., 2017), and primary care engagement (Lefchak et al., 2022). Sun et al developed a pediatric comorbidity index based on ICD-10 code diagnoses, but it only includes diagnosed conditions and does not incorporate screening or symptom data (Sun et al., 2021). There are models which predict specific behavioral health outcomes in children, such as Spencer et al’s work investigating

social risk and mental health of children aged 6-11 in primary care (Spencer et al., 2020) or Scholes-Balong et al's work predicting incident cannabis use in adolescents (Scholes-Balog et al., 2020). There are high quality models which highlight mental health risk, such as Su et al's work predicting suicide using electronic medical record data (Su et al., 2020). These models do not help a primary care provider to flag children on whom they want to do more comprehensive screening and follow-up, across a variety of risk factors.

Examples of strong generalized risk scores which capture social, behavioral and medical risk do exist, such as the Integrated Child Risk Index developed by Bethell et al in 2022 (C. Bethell et al., 2022). The authors used a nationally representative survey dataset to create an integrated risk assessment tool that scores medical, social and relational risk. The score defines medical risk as 1+ complex healthcare need, 2+ chronic conditions, 1+ functional difficulties or fair/poor overall health status. It defines social risk as food insecurity, income insecurity, unsafe neighborhood/violence exposure, or experienced racism. It defines relational risk as 2+ household ACES, caregiver with fair/poor mental health, caregiver-reported aggravation with child, or poor coping of caregiver. The authors identified these domains from a comprehensive review of the pediatric and adolescent risk literature. Each domain was independently associated with the outcomes of interest, which included emergency room days, score on the child flourishing index, and educational preparedness and engagement. This score can reasonably be considered a gold standard of risk assessment within research. However, this score has yet to be simplified for clinical implementation or applied to prediction tasks within primary care.

Among adults, there is a robust body of literature on generalized risk prediction in the clinical setting. Robust models predicting all-cause mortality and all-cause hospital admission or readmission among adults which are often applied to the outpatient context, including the Charlson comorbidity index (Sundararajan et al., 2004) and the LACE index (Fry et al., 2020). Multiple electronic medical records have automated all-cause admission risk prediction

algorithms for adults. Models are available not only for the general population but also more vulnerable sub-populations, like the low income and uninsured patients of NYC Health+Hospitals (Ziring et al., 2018) or Denver Health, a large safety net system in Colorado (T. L. Johnson et al., 2015) (Rinehart et al., 2018). Generalized risk prediction models and accompanying complex care programs are far less robust in the pediatric space as compared to adult medicine. This is a key gap with implications for the health of families and children.

Discussion

Summary of Findings

In the 40+ years since the International Conference on Primary Health Care published the Declaration of Alma Ata, there has been an increasing trend to recognize the importance of social determinants to the production of health at the population level. SDOH research suggests complex, meaningful interactions between the individual, their family, their community and their physical and policy environments. Calls to integrate social determinants into primary care have become ubiquitous, but implementation of screening and intervention has been heterogeneous. It is critical to address social determinants of health to support healthy families and adolescents and avoid the adverse impacts of chronic stress caused by things like financial uncertainty, discrimination, food insecurity and housing instability.

The field of Adverse Childhood Experiences research has sought to quantify the effects of toxic stress on the production of health and healthy behavior in adults and children, with a particular focus on relational dysfunction. ACEs exposure is linked to developmental impacts in the neurological, immune and endocrine systems via the physiological impacts of cumulative stress or “allostatic load.” In adults, exposure to ACEs have been shown to be associated with poor physical, mental and behavioral health and are responsible for an estimated $748 billion dollars annually. While younger, the literature on pediatric ACEs has shown that similar health effects emerge during childhood or adolescence itself, impacting sexual health, chronic disease control, behavioral risk taking, alcohol and substance use and school performance. While some

jurisdictions like California have moved to make ACE screening routine, consensus on the benefit and impact of ACEs screening remains elusive, and evidence on efficacy of associated interventions is weak.

Interventions to improve adolescent health and wellbeing are important because adolescents are in a sensitive developmental period. Adolescents are going through hormonal puberty, developing social awareness and undergoing changes in the brain which impact mood, behaviors, reasoning and decision-making. Primary care providers can be important partners in supporting adolescent growth and self-actualization, in addition to supporting adolescent physical and sexual health. When adolescents are struggling, high quality primary care offers a gateway to enhanced services like embedded behavioral health support, referrals to specialty providers, connection to community supports, advocacy for school-based services, enrollment in social benefits programs or engagement with community health workers or social workers. To properly care for adolescents in need, however, primary care practices must identify which children might benefit from enhanced connection.

Major Themes

In reviewing the literature of social determinants of health, adverse childhood experiences and adolescent development, several key themes emerged. This discussion elucidates two main themes: the increasing convergence of concerns across SDOH and ACEs disciplines; and screening and prediction of risk in adolescent primary care. By tracing these themes and identifying key gaps, this review aims to situate the inquiry study within the literature.

Theme 1: Overlap of Domains across SDOH and ACEs Fields of Study - Relational and Social Stressors Impact Allostatic Load

As Allen et al so eloquently illustrated, the health of families is not produced in a vacuum (Allen, Matilda & Donkin, Angela, 2015). Social determinants impact ACEs, and ACEs impact social determinants, across generations. Social stressors may cause ACEs directly, such as a

lived experience of racism or discrimination. Social stressors may also cause ACEs indirectly. As Conger and Conger outlined in their Family Stress Model of Economic Hardship, financial insecurity increases parental stress and decreases parental resilience, leading to increased parent-adolescent conflict and harsher discipline (Conger et al., 2010), which may ultimately increase the risk of abuse or maltreatment. Additionally, ACEs impact lifetime earning potential through increased risk of incarceration, emotional volatility and poorer overall health (Hughes et al., 2017), contributing to an intergenerational cycle of poverty which further increases risk of ACEs (Shonkoff et al., 2021). Resilience in the form of family or relational assets, strong communities, social safety net policies or trusted relationships with key adults can mitigate the impact of SDOH and ACEs on health and wellbeing (Boullier & Blair, 2018; Garner et al., 2021; Shonkoff et al., 2021).

An adolescent coming into primary care brings these many sources of risk and resilience with them, and providers must address sources of social and relational risk to treat the whole patient. It is therefore gratifying to see that the ACEs literature has expanded to include more social sources of toxic stress over time. The canonical 1998 Felitti study of adverse childhood experiences classified seven major domains of ACEs: abuse of a psychological, physical or sexual nature; violence against a parent; substance use, mental illness or suicidality in the family; and incarceration of a parent (Felitti et al., 1998). Over time, socially mediated sources of toxic stress started appearing in ACE studies. Finkelhore et al included community violence exposure and low SES in their revised inventory of adverse childhood experiences (Finkelhor et al., 2015). From 2016, the National Survey of Children's health included "was treated or judged unfairly because of his or her race or ethnic group" and "witnessed violence in his or her neighborhood" in their ACEs survey module (Women and Children's Health, 2020). The PEARLS screener includes food insecurity, housing instability, exposure to violence, experience of discrimination in the ACES and related experiences assessment (Koita et al., 2018). Byomsad et al included lived experience of racism or discrimination, witnessing neighborhood

violence or material deprivation in the last year in their 2020 study of the impact of ACEs on adolescent health (Bomysoad & Francis, 2020), and Bethell et al included the same three factors in their 2017 systematic review (C. D. Bethell et al., 2017). These researchers were recognizing overlap between social and relational toxic stress that had long been recognized in other disciplines with a more holistic childhood adversity focus, such as research focused on incarceration. This trend towards synchronicity supports the inclusion of both social and relational domains of risk in any attempt to capture holistic adolescent adversity and wellbeing in primary care.

Theme 2: Defining, Screening for and Predicting Risk in Primary Care

The literature on adolescent development lacked a single, validated, canonical definition of adolescent risk. Decision-making around sexual health, alcohol and substance use was a concern across sources. Similarly, exposure to violence, abuse and neglect, incarceration or bullying was highlighted by most sources as a key area of adolescent risk. All sources highlighted mental health risk, including depression, anxiety and suicidality. Asthma was the only chronic disease referenced across sources, although developmental disorders and delays which impact school performance did appear frequently. Areas of healthy living and habits, including healthy eating, exercise, sleep and phone use were frequently mentioned. Select sources also mentioned lived experience of discrimination as a key risk. These risk factors encompass a wide range of both adverse outcomes and predicting factors of adverse outcomes, with many elements acting as both an outcome and a predictor of further harm.

The WHO, the Youth Risk Behavior Survey and the Integrated Child Risk Index offer three models of concrete risk measurement across large adolescent populations. Taken together, they offer a valuable insight as to what risks can reasonably be measured in adolescent patients. However, these frameworks are too complex and time-intensive to realistically deploy in a primary care setting. Adapting any of these models would require

significant de novo patient data collection, in the form of waiting room surveys, MyChart patient questionnaires or provider, nurse or medical assistant time during a visit.

Another barrier to adopting a risk classification system from the WHO, YRBS or Integrated Child Risk Index is the heterogeneous nature of screening in primary care. To create a risk score, it would be ideal to screen directly for both ACEs and social determinants of health. However, this literature review surfaced that implementation of these screenings in primary care has been heterogeneous. Although social determinants of health screening is advancing at the state level, consensus on screening domains and effective intervention pathways within the healthcare system has lagged. Implementation of ACE screening in pediatrics has been uneven, with limited evidence available as to the efficacy of potential health system interventions for positive screens.

Beyond SDOH and ACEs, adolescents in primary care may not receive the full spectrum of existing screenings for depression, suicidality, behavioral risk taking or substance use (Riley et al., 2018). Even if screening were universal, providers looking to understand their full patient panel are met with an overwhelming amount of information which is poorly organized within the electronic medical record, contributing to overwhelm and burnout among providers (Li et al., 2022). Manually balancing an ACE score against results of screenings for social determinants, suicidality, depression, anxiety, school performance, sexual health, alcohol use and substance use is not scalable. A generalized risk score would reduce provider load inside the EMR.

The Pediatric Comorbidity Index (PCI), developed by Boston Children's Hospital, does capture many chronic conditions using existing EMR data (Sun et al., 2021). However, it has several significant limitations. It does not include measures of adversity such as homelessness, experience of violence or abuse and neglect, which we know are extremely important to pediatric trajectories. It is diagnosis-based, meaning it does not capture screening data, social determinants, laboratory results or vitals in the electronic medical record. It also lacks any measure of neighborhood disadvantage or experience of discrimination, which are important

predictors of long-term health. An ideal score would take advantage of the full spectrum of data available in the record, including both proxy and direct measures, when assessing adolescent risk.

A predictive risk score, leveraging available EMR data and supplementing that data with appropriate geographic data, might help to address this gap. In adult medicine, predictive risk scores help primary care practices outreach overdue patients, decide who to prioritize for 7 day follow up after hospitalization or fill Community Health Worker program spots. Although there are many condition-specific predictive models available in pediatrics, this literature review was unable to identify any prediction models that quantified generalized risk across medical, social and behavioral domains.

This may be due to the difficulty of defining an outcome of interest for adolescent patients in primary care. While cost or acute utilization are reasonable outcomes in an adult population, ED and inpatient visits among adolescents underestimate the true burden of morbidity and risk in this age range. As Shonkoff stated, “reductions in health care use (e.g., emergency department visits and hospitalizations) are clearly important, but they are not a sufficient proxy for well-documented direct effects on child health or development.” (Shonkoff et al., 2021). Optimal adolescent outcome data in the electronic medical record are often incompletely recorded and or missing entirely. For example, thriving in school is correlated with adolescent mental health and wellbeing across studies (Lee et al., 2024). However, educational outcomes are seldom robust inside the electronic medical record. Data related to school performance such as attendance have been historically unavailable due to FERPA data exchange limitations which require individual written consent and are not feasible to meet across a large patient panel (*HIPAA and FERPA Basics*, n.d.). Absent educational measures such as attendance, few single outcome measures in the medical record are applicable across ages 12-17 and offer appropriate discrimination between high, medium and low risk individuals.

Applying the framework of adolescent ACEs exposure is one way to elevate which risks to quantify as adverse outcomes for inclusion in a generalized primary care score. ACEs are a mediator on the causal pathway of many of the aforementioned risks. A history of ACEs is also predictive of which adolescents and families will be most impacted by further adversity. By prioritizing outcomes associated with ACEs exposure to create a continuous adverse outcome score, it may be possible to better determine which adolescents require additional support to thrive.

Responding to the Literature: Towards a Pragmatic Risk Score for Primary Care Adolescents

This literature review has traced the increasingly interconnected study of social determinants of health and adverse childhood experiences. It has established that adolescence is a crucial developmental period typified by enhanced social porousness and increased novelty and risk-seeking. It has identified the role of primary care providers in supporting healthy adolescent development and supporting adolescents experiencing adversity. It has also identified the lack of adequate ways of identifying these adolescents through direct screening or existing risk prediction. To address these gaps, this inquiry developed a risk-prediction algorithm identifying adolescents at highest risk of experiencing ACE-associated adverse outcomes in the following year, using data from the electronic medical record enriched with key social determinants of health data as measured at the neighborhood level. If implemented, this algorithm would allow primary care providers at NYC Health+Hospitals to connect vulnerable patients to the services and interventions briefly described in Chapter 1 and fully described in Chapter 5 (Discussion), addressing a key gap in both literature and practice.

Chapter 3: Methods

Research Design

This inquiry used an epidemiologic approach to investigate the distribution of socio-medical risk in an urban safety net population of adolescents. The goal was to create a parsimonious risk score derived from the robust body of literature that could be implemented programmatically in a real-world setting (Goldstein et al., 2017). This inquiry used data from patients aged 12-17 who visited a primary care practice in 2022 to predict the risk of adverse health outcomes experienced in 2023. It generated a list of suggested adolescent health outcomes associated with Adverse Childhood Events from the literature, and associated risk predictor variables available inside the electronic medical record (EMR) or via geographic match. It synthesized previously conducted key informant interviews to craft a final set of adverse outcomes that were a) measurable in the EMR or via geographic match, and b) defined by providers and patients at NYC Health+Hospitals as important to adolescent risk in the primary care context.

To train an algorithm predicting 2023 adverse events using 2022 observed data, the study split the data geographically into a 70/30 train/validate partition. Training data variables were examined for collinearity, skewness, sparseness and missing data, and transformed, combined or dropped where necessary. Within the Training data, the study selected a modeling modality by tuning linear, negative binomial and logistic LASSO models. These preliminary models were selected by minimizing the Schwartz Bayesian Criteria (SBC), chosen for computational parsimony. Logistic LASSO provided the best balance of parsimony, discrimination and explainability while offering moderate to strong calibration. The final model was produced using logistic LASSO regression.

To train the final logistic model, five-fold cross-validation LASSO was performed. Predictors which were retained in at least two folds at the optimal LASSO step (minimum

average squared error + 1 standard error) were retained in the final model. Where two versions of the same variable were retained across folds, the version with the strongest beta was retained. Additionally, two variables (sickle cell diagnosis, eating disorder diagnosis) which were flagged as important for buy-in by the key informants but which were not retained by LASSO were tested and retained in the final logistic model.

To assess the impact of unmeasured missing data on model fit, a sensitivity analysis examined final model performance by loss to outpatient follow-up status. To assess bias, equal opportunity and predictive parity were assessed by race/ethnicity, language, sex at birth and insurance status. Finally, the model was deployed within the Validation set to derive the Validation calibration, discrimination and threshold-specific performance.

Sample

This study captured 37,126 adolescent patients with at least one visit to a primary care department in one of NYC Health+Hospitals' 11 hospitals, 6 ambulatory care centers or 33+ clinics during 2022, as measured in the Epic electronic medical record system.

This study's inclusion criteria were:

1. A visit to a pediatric, adolescent, family medicine or adult medicine primary care department at one of NYC Health + Hospitals' 11 hospitals, 6 ambulatory care centers and 70+ community health centers.
2. Being an adolescent, defined as age 12-17 years during 2022; and
3. Residing in New York City, defined as having an NYC zip code listed in the patient address field.

The study's exclusion criteria were:

1. Having an ancillary visit at H+H in 2022 but not having one or more visits with a full provider (MD, NP, PA). Ancillary visits include blood draws, x-rays, immunization-only visits and COVID-19 test visits.

2. Missing sex at birth demographic data or address data. Although other missing data were imputed or dropped, these two variables were required for upstream data matches to citywide datasets to generate homelessness, NYCHA status and neighborhood-level Social Vulnerability Index.

Inclusion and exclusion criteria were applied during the routine generation of these secondary use datasets. All eligible patients were used in the analysis and sample size was not calculated.

Because this study leveraged existing electronic medical record data for secondary data analysis, no participant recruitment or consenting was done. The NYC Health+Hospitals BRANY IRB found this study to be exempt. The Marywood IRB approved the study under expedited review with waiver of informed consent.

Validation Sample Etiologic modeling traditionally samples data to maximize generalizability. However, predictive modeling and machine learning usually leverage all available data to maximize signal, and instead use partitions to evaluate performance. This means that the model is developed using 70-80% of the available sample, and 20-30% of the sample is held back to validate algorithm performance (Bedoya et al., 2022). In this study, the prediction dataset was first partitioned in a 70/30% split, with 70% of the included population used to train the model and 30% used to assess internal validity and to produce final fit statistics.

This study used a non-random partition, which more closely mimics the real-world non-random population variation, or a type 2B validation per the TRIPOD statement on standard reporting of prediction models (Collins et al., 2015). This study does not include a robust external validation component because the goal was to produce a locally tailored model for the NYC Health+Hospitals primary care population rather than a nationally generalizable model. This study used geographic partition, which is a form of pseudo-external validation (Steyerberg & Vergouwe, 2014). All patients whose most recent primary care visit in 2022 happened in the Bronx were partitioned into the Validation dataset (n=10,764). The Bronx was chosen because it

represents an appropriate proportion of the total dataset (29.0%). The demographics and environmental characteristics of the Bronx are distinct as compared to other boroughs, which is ideal to mimic real-world variation. The prediction dataset retained patients receiving primary care at H+H sites based in Queens, Brooklyn, Manhattan and Staten Island (n=26,362).

Instrumentation

In *Introduction*, this inquiry identified five key sub-problems with definitional dimensions: Which demographic, SDOH, utilization and clinical data in the electronic medical record in 2022 predict adverse outcomes in 2023 among adolescents in primary care at NYC Health+Hospitals (Sub-Problems 1-4)? Which adverse outcomes among adolescents in primary care can be measured in the electronic medical record and aggregated into a single score for 2023 (Sub-Problem 5)? The *Instrumentation* section outlines the approach of defining candidate demographic, SDOH and clinical predictors and the process undertaken to define an outcome score. Details on selection of final predictors is available in *Analysis of Data*.

Instrumentation Overview: Sources of Data There was no pediatric primary data collection associated with this study. This study used secondary data from the Epic electronic medical record (EMR) to capture patient risk factors and outcomes. These variables were derived from the literature and from previously collected H+H key informant interviews (see appendix). In addition to EMR data, the study used publicly available geographic data as a proxy for unmeasured but important social factors such as family income. The study used geocoded patient addresses to link to a public use Centers for Disease Control Social Vulnerability Index dataset for neighborhood-level SDOH data (*CDC SVI Documentation 2020 | Place and Health | ATSDR*, 2022). Geographic linking was also used to identify patients living in NYC Housing Authority (NYCHA) public housing and patients experiencing homelessness, via NYCHA and homelessness flags routinely stored in the NYC Health+Hospitals data warehouse.

Definitional Process To find predictor and adverse outcome candidates to evaluate for the model, the adult and pediatric ACEs literature was reviewed to identify mental, physical and

behavioral health issues. Included outcomes were a) associated with exposure to one or more relational, environmental or social ACEs, as articulated in the literature; b) measurable in data from a clinical EMR or via a geographic match; and c) time-bound, aka able to be measured in 2023 specifically. Although adverse outcomes were collected across studies, the most useful resource was Nelson et al's systematic review of the association of exposure to ACEs with adverse outcomes in childhood and adolescence (C. A. Nelson et al., 2020). Adverse outcomes identified in the literature were translated into discrete measures that could be captured with data in the Health+Hospitals enterprise data warehouse or matched geographically. For example, Nelson et al identified "adolescent high risk sexual activity" as an outcome domain associated with ACEs exposure. That outcome domain can be measured in 2023 as a pregnancy (live birth, prenatal or postnatal visit, abortion or miscarriage) or a positive sexually transmitted infection or STI (Diagnosis or lab result). High association outcomes that cannot be measured in available data, such as high absenteeism in school, were not considered.

Previously collected key informant interviews were a guiding light in defining outcome domains and their candidate predictors. The collection, format and key findings from these key informant interviews are described in the Appendix. Key informants rated adverse outcomes associated with ACEs in the literature in terms of their importance to a provider, social worker or patient's experience of risk in the adolescent primary care setting. Key informants also identified domains of risk that did not surface in the literature review but which are important in primary care, such as avoiding emergency visits related to poor control of epilepsy or diabetes. Key informant interview transcript notes were abstracted and aligned by domain and measure. Level of consensus on importance of domain inclusion was evaluated across key informants. Finally, measure ratings were tabulated across key informants to create an average rating per measure. Key informant methods are described in Appendix A. Key informant themes, findings and average measure ratings are presented in Appendix B.

Using key informant feedback and literature findings, the dissertation author drafted candidate predictor and outcome of interest definitions for the NYC Health+Hospitals EMR. Draft variable definitions were reviewed with two NYC Health+Hospitals subject matter experts in EMR data to ensure uniformity, viability and construct validity. Variable definitions about which there was a disagreement were discussed until consensus was reached. For some variables, additional expert feedback from the H+H system was requested. Specifically, the study augmented a pre-existing definition crafted by the HIV and Sexual Health team for both pregnancy and positive STIs and leveraged an Ambulatory Care definition of abortion-related visits. Variable definitions were refined after data cleaning, as outlined in Analysis of Data and documented in the Data Dictionary. Variables were not stratified demographically.

Outcomes of Interest: 2023 Adverse Outcomes Score Because there was no single outcome score measuring social and medical risk available from the literature, this study sought to identify measurable events in the EMR which were identified by key informants as important risk outcomes for adolescents in primary care at NYC Health+Hospitals. It is not a validated score, but rather a reflection of local priorities. As such, it may have limited generalizability to other contexts. The limitations of this approach are further explored in Discussion.

Candidate outcomes were sourced from the ACEs literature and selected using the full key informant findings outlined in Appendix B. Key informant feedback was translated to 57 potential outcome sub-components, as outlined in the Data Dictionary. The final score summed outcomes across 45 composite outcomes. Each positive outcome recorded in 2023 received a single point, except the Pediatric Comorbidity Index (PCI). The PCI flag was coded categorically (0-2 comorbidities = 0 points, 3-4 comorbidities = 1 point, 5+ comorbidities = 2 points), based on cut points drawn from the 75th and 95th percentile PCI counts among H+H adolescents in primary care. The domains captured by the 2023 Adverse Outcome Score are detailed in Table 3.1 and described briefly below. For detailed definitions, please see the Data Dictionary.

The 2023 Adverse Outcome Score included four variables in the Abuse, Trauma, Neglect domain, which were a visit diagnosis or screening in 2023 for abuse or neglect; a visit diagnosis of enuresis or encopresis; and visit z-code diagnoses of experiencing psychosocial difficulties, including family conflict or interpersonal violence. The Mental Health domain included variables capturing 2023 visit diagnoses or screenings for depression, anxiety or suicidality. It also included a visit diagnosis flag for serious mental illness, defined as having a visit for obsessive compulsive disorder, psychosis, specified personality disorders or conduct disorder per the pediatric comorbidity index, or being the perpetrator of violence per CCSR). The Mental Health domain also included having 2+ somatic symptoms (headache, stomach problems and insomnia) documented via visit diagnosis. The Unmet Behavioral Health Needs domain included 1+ ED visits with a diagnosis of anxiety, depression, conduct disorder, OCD, psychosis, specified personality disorders or ADHD in 2023. The Alcohol & Substance Use domain included visit diagnoses or screenings for high alcohol use, any illicit substance use, overdose diagnosis, and tobacco use via cigarette, vape or other tobacco product. The Healthy Eating domain included being in the top or bottom 5% of BMI change within H+H from 2022 to 2023 (BMI Trajectory), having a diagnosis of disordered eating, and having 1+ ED visits for weight loss in 2023. The Social Determinants of Health domain included 1+ social needs (screening positive for 1+ social need on the social needs screener, receiving a social need z-code, screening positive for food insecurity on SSHADESS screening) or receiving a lifestyle-related z code diagnosis in 2023. Key informants frequently mentioned that social needs add family stressors and impact adolescent wellbeing, findings which were echoed in the literature review and the Conceptual and Theoretical frameworks. The School Challenges domain included screening positive or receiving a diagnosis for difficulty in school, a diagnosis of ADHD, or a diagnosis of a developmental disorder (PCI categories of chromosomal abnormalities, congenital malformation, pervasive developmental disorders or developmental delays) in 2023. Key informants felt that this domain was critical and lamented the limited

information available within the EMR. The Unmet Developmental Needs domain included 1+ ED visits with a diagnosis of chromosomal or congenital malformations, developmental delay or pervasive developmental disorders (including autism) in 2023. The sexual health domain included evidence of a pregnancy in 2023 or screening positive for a sexually transmitted infection in 2023. The general Chronic Disease domain included receiving a sickle cell diagnosis in 2023 and comorbidity count (0-1, 3-4 or 5+ comorbidities per the Pediatric Comorbidity Index). The Uncontrolled Chronic Disease domain included 1+ ED visits with a visit diagnosis of epilepsy, asthma, diabetes, cardiovascular disease and chronic GI illness. Although outside of asthma these conditions were not sourced from the literature as being associated with ACE exposure, the key informants felt that ED visits for these chronic diseases indicated sub-optimal management of the condition and therefore elevated risk. The Concerning Utilization domain included having 3+ ED visits, any inpatient psychiatric unit visit, 2+ generalized inpatient visits, 5+ specialty outpatient visits or no well child visit in 2023, with cutoff rationales specified in Table 3.1.

These component variables were added to create the final outcome score. The continuous 2023 Adverse Outcome Score was used for negative binomial and linear modeling, which require count and continuous data respectively. For logistic modeling and threshold-specific performance assessment, the score was binarized at the top 5% threshold. This translated to 6+ adverse outcomes, meaning that 5% of the Training sample had ≥ 6 adverse outcomes recorded in the EMR in 2023. These patients were flagged as “true high risk.”

There are several key limitations of this outcome as specified. Because several overlapping concepts were captured (condition-specific ED visits and a flag for 3+ ED visits, for example), some amount of double-counting was likely present in the score. This outcome score weighted depression, anxiety and ADHD slightly more heavily than other conditions, because it included both a flag for any visit diagnosis, and a separate flag for ED utilization for these

conditions. This was chosen because many adolescents with these conditions have no emergent utilization but are still at elevated risk. Adolescents who *do* have behavioral health-related emergent utilization are at extremely elevated risk. The overall impact of this overlap was low, impacting n=331 patients with anxiety, n=257 patients with ADHD, and n=360 patients with depression.

The outcome score also weighed 3+ ED visits and 3+ chronic comorbidities slightly more heavily than other risk factors. The practical impact of these overlaps was also modest. N=256 patients received a point for 3+ ED visits and also received points for 3+ condition-specific ED visits. N=145 patients received 1-2 points for chronic comorbidity which was also captured in condition-specific ED visits. Due to the many patients *not* captured in these overlaps, this outcome score retained these slightly overlapping concepts.

Candidate Predictors: To create the list of candidate predictors across demographics, social determinants, utilization and clinical data (Sub-Problems 1-4), a full census of available data routinely collected in clean enterprise data warehouse tables was undertaken. High value variables were collected for consideration, ultimately yielding 104 candidate variables. These variables were mapped to predictor concepts identified as important in the literature or by key informants. Variables which were not time-bound or which were considered by NYC Health+Hospitals clinical stakeholders to be unreliable were removed from consideration. As outlined in Analysis of Data, additional predictor variables were removed during data cleaning due to issues of missingness. The full set of tested predictor variables is available in the Data Dictionary in the Appendix.

Defining demographic and utilization predictors (Sub-Questions 1, 3) was relatively straightforward, drawing from standard fields in heavy use. Defining social determinants of health predictors (Sub-Problem 2) involved reviewing the level of completeness of current social determinants of health screenings at H+H, identifying commonly used social determinants ICD-10 “z codes,” and reviewing candidate geographic datasets to augment variables such as

income which were missing in the record. Ultimately, the CDC Social Vulnerability Index data, available at the census tract level (averaging 2016-2020 data) for NYC provided broadly applicable measures that are likely to be available longitudinally. Defining clinical predictors (Sub-Problem 4) involved heavy consultation of key informant interview transcripts, discussion with clinical experts within the Office of Population Health and Ambulatory Care, and thorough review of available diagnostic, lab, screening and vitals data. The Data Dictionary provides a full definition of each final candidate predictor variable.

Instrumentation Details: The scores of multiple instruments were referenced in the course of this study. The instruments and broad variable categories are presented here under the domains identified in the research sub-questions. Full variable definitions are available in the Data Dictionary, within the Appendix. A brief overview of variables requiring more explanation is offered below, organized by category of predictor:

Demographics and Social Determinants of Health:

- Age (continuous, 12-17 as of Dec 31, 2022)
- Sex at birth (male/female - nonbinary and trans categories not collected routinely in this field)
- Race/ethnicity (Hispanic/Latinx and Non-Hispanic: African-American, Asian/Pacific Islander, White, Something Else;)
- Preferred language (English, Spanish, Other)
- Insurance status, as of last visit (Commercial, Medicaid, Uninsured)
- Sexual orientation and Gender Identity (SOGI) Screening
- Social needs screening yield (0 vs. 1+ social needs, as identified across housing, food insecurity and other social domains on homegrown Health+Hospitals social screener that is semi-routinely administered in primary care)
- Homelessness (history of homelessness in last year, denoted by membership in H+H homelessness registry in the clinical data warehouse)

- Public housing (geographic match to NYCHA public housing, as of last visit in 2022)
- Social Vulnerability Index and associated sub-domains (geographically matched to patient home address census tract, using CDC's 2020 SVI data for New York City (*CDC SVI Documentation 2020* | *Place and Health* | *ATSDR*, 2022))

Utilization (H+H only - no external utilization available):

- Primary care visits (including visits to pediatrics, adolescent medicine, family medicine) in 2022
- Specialty visits (outpatient visits outside primary care)
- Behavioral Health visits (outpatient visits to a behavioral health clinic)
- Emergency department visits (Psych ED and General ED)
- Inpatient stays (Psych Inpatient and General Inpatient)

Clinical

- Diagnosis variables - 2022 (ICD-10 code diagnoses of persistent asthma, sickle cell, diabetes, hyperlipidemia, depression, anxiety, developmental delay, autism, ADHD, mood disorder, suspected or confirmed abuse or neglect, enuresis, encopresis, social and school-based difficulties, SDOH-related z codes, pregnancy, live birth, abortion, sexually transmitted infection)
- Pediatric Comorbidity Index count of comorbidities (Sun et al., 2021))
- Pregnancy-related visits (prenatal, labor and delivery, abortion)
- Depression via Patient Health Questionnaire-9 (PHQ-9) screening (*The PHQ-9 - PMC*, n.d.) or depression visit diagnosis
- Anxiety via Generalized Anxiety Disorder 7-Item (GAD-7) anxiety screening (*Generalized Anxiety Disorder 7-Item (GAD-7) - Mental Health Screening - National HIV Curriculum*, n.d.) or anxiety visit diagnosis
- Tobacco use screening (Current/Former, Never or Missing screening) or ICD-10 code diagnosis

- Alcohol or Substance Use captured via the CRAFFT substance use screening (*About the CRAFFT – CRAFFT*, n.d.; Knight et al., 1999) or visit diagnosis
- Suicidality captured via Columbia Suicide Severity Rating Scale (C-SSRS) suicidality screening (*Columbia Suicide Severity Rating Scale (C-SSRS)*, 2021), PHQ-9 question or visit diagnosis
- BMI trajectory (95th, 5th percentile of change in BMI from 2022 to 2023, as compared to all adolescent primary care patients (Jackson, 2018))
- Lab evidence of sexually transmitted infection (gonorrhea, chlamydia, syphilis)

Procedure (Data Collection)

No primary data were collected for this study. Secondary clinical data were pulled from the H+H Office of Population Health's Enterprise Data Warehouse primary schema by NYC Health+Hospitals analysts for the dissertation author. Base data were stored in an access-controlled drive by credentialled Office of Population Health staff. The study used industry standards for handling electronic medical record data throughout (Goldstein et al., 2017). Social Vulnerability Index data were downloaded from the CDC/ATSDR website July 2023. The study also referenced previously collected Key Informant Interviews with H+H clinicians and patients - for more details, see Appendix A.

Analysis of Data

The *Instrumentation* section of methods outlined the definition of candidate predictors and the selection process for the adverse outcome score, partially answering Sub-Problems 1-4 (Which demographic, SDOH, utilization and clinical data in the electronic medical record in 2022 predict adverse outcomes in 2023 among adolescents in primary care at NYC Health+Hospitals?) and fully addressing Sub-Problem 5 (Which adverse outcomes among adolescents in primary care can be measured in the electronic medical record and aggregated into a single score for 2023?). *Analysis of Data* addresses the processes used to select final predictors from among the candidates and outlines the analytic approach used to answer

Sub-Problem 6 (How accurately can the risk of adverse health outcomes for adolescents in primary care at a large urban safety net system be predicted by the prior year's electronic medical record and public data across demographic, SDOH, utilization and clinical domains?)

Analytic Overview: First, data was partitioned geographically into a 70/30 train/validate split. Within the Training partition, candidate predictors were examined for collinearity, skewness, sparseness and missing data, and transformed, combined or dropped where necessary. Preliminary logistic, negative binomial and linear regression models were fitted using LASSO variable selection to predict the 2023 adverse outcome score with the 2022 candidate predictors described in Instrumentation and defined in the appendix. Model fit and predictive power were assessed across all three modeling approaches within the Training data to identify the most appropriate modeling distribution.

Logistic regression was chosen as the optimal LASSO modeling modality, as described below. A five-fold cross-validation LASSO selection was then performed to identify final LASSO model variables, using the approach outlined by Hastie et al, and results were compared to those obtained via SBC selection. Variables retained across at least two of five folds were added to the final model, and qualitative review of collinear and essential variables was done. The final model was then retrained to obtain the final logistic model parameters. As a sensitivity analysis, performance was assessed among those not lost to outpatient follow-up status. To identify model bias, equal opportunity and predictive parity were assessed by race/ethnicity, language, insurance and sex. As the final step, the final logistic model was deployed on the Validation partition data to obtain Validation statistics.

Software Data were cleaned and analyzed using SAS Enterprise Guide 8.3. Code used for this project is available upon request. Data are available with legal DUA. The data dictionary was built and maintained in Excel. Figures were created in SAS, Excel and Powerpoint.

Data Cleaning The candidate dataset included 161 variables, including 104 candidate predictors from 2022 data and 57 candidate outcome score sub-components from 2023 data.

Per standard practice, the dataset was first partitioned into Training and Validation datasets before any cleaning and exploratory data analysis so as not to overfit or bias the data analysis process. The partition is described in Sample. After data cleaning, 88 predictor variables and 46 outcome score variables were retained. Data were cleaned without demographic stratification.

Within the Training data, all predictor variables were examined for skewness, heteroskedasticity, missingness and sparsity via frequency tables, histograms and descriptive statistics (mean, SD, median, interquartile range and spread), using proc freq and proc univariate in SAS. Outliers were not removed at this stage, because this population shows true heterogeneity that is important to retain. Instead, highly skewed variables were censored or transformed to meet skewness <|2| and kurtosis <|7| thresholds recommended by Hair and Bryne (Byrne, 2013; Hair et al., 2013). Further details are in *Data Transformations* below.

Variables which primarily took a single value (aka small n had non-zero values) were dropped or collapsed with similar variables belonging to the same domain (Luo et al., 2016). These included the candidate predictors for pregnancy (n=46, combined with STI flag), diagnosis of specified personality disorders (n=32, combined into Serious Mental Illness diagnosis category) and diagnosis of malignancy (n=38, no natural grouping, dropped). No variables suffered from the problem of perfect separation, whereby a variable is fully predictive of the outcome. Variables belonging to a common concept or which were determined to be highly collinear using chi square or Spearman's correlation coefficient were also combined into composite flags.

Data Missingness Missing data can be handled in several ways. For variables with large amounts of missing data, it is often most appropriate to either model missing as a category of the variable or combine variables with high missingness with other more complete variables as a composite flag (M. K. and K. Johnson, n.d.). Within the Training data, multiple candidate predictors had large amounts of missing data. They were handled as follows: the SSHADESS exercise screening (90.6% missing) and SSHADESS social support screening (83.1% missing)

were combined to create a Resilience flag, and patients missing both screeners were labeled as “Missing.” The food insecurity SSHADESS screening (82.8% missing) was combined with the Social Determinants Screener and visit-level social need diagnoses to create a composite SDOH Need flag. The foster care SSHADESS screening (82.0% missing) and the safety screener (63.3% missing) were added to the Abuse and Neglect flag. Foster care will be augmented with a new foster care social services flag in future, but these data were not available for 2022. The school performance SSHADESS screening (73.7% missing) was combined with an AHRQ CCSR diagnosis variable for school challenges to create a School Challenges flag. The sexual orientation (65.6% missing) and gender identity (57.8% missing) screenings were combined into a SOGI flag, and missing was modeled as a category.

Variables with less missingness were imputed. Per Steyerberg and Vergouwe, simple or multiple imputation is preferable to complete case analysis (i.e. removing patients with any missing data from the dataset) (Steyerberg & Vergouwe, 2014). Multiple imputation can be prohibitive to implement within LASSO modeling in the SAS software in terms of working server memory and was thus out of scope for this dissertation due to system constraints (Musoro et al., 2014). K nearest neighbor imputation is another method which preserves variability while imputing missing values (Gunn et al., 2023). K nearest neighbor was attempted on these data but failed to run due to insufficient server memory in the Health+Hospitals SAS environment. This dissertation therefore used hot deck imputation to address missing values in the data, which while a weaker imputation approach is still preferable to complete case analysis or substitution of mean/mode value (Bechtel et al., n.d.; Mukhopadhyay, n.d.). BMI Trajectory was the only outcome variable which required imputation. First, loss to follow-up patients, who had no 2023 visits, were set to a BMI trajectory of 0, to avoid providing information for patients who had not come in. The impact of these patients’ missing data on model performance was examined in the loss to follow-up sensitivity analysis. Next, missing 2022 and 2023 BMIs were imputed using simple random sampling without replacement (SRSWOR) using a single donor

for each observation (due to memory limitations) based on cells constructed with age, sex, race/ethnicity, eating disorder diagnosis and diabetes diagnosis. This resolved BMI trajectory missingness for n=2,878 patients, and failed to resolve missingness for n=7 patients. Remaining missingness was minimal across SVI variables (Themes n=29; Theme1, Theme2, Unemployment, Uninsured n=28; No High School Diploma n=27), Language (n=2), Homelessness (n=1), NYCHA Housing (n=1) and Insurance (n=1). Hot deck imputation was conducted for these variables using simple random sampling without replacement (SRSWOR) using a single donor for each observation based on cells using age, sex, race and SDOH screening. This successfully resolved missingness across these variables.

Data Transformations Utilization variables showed heavy skewness and kurtosis (>|2|, >|7|). Candidate utilization variables for behavioral health visits, ED psych visits, inpatient psych and non-psych visits were binarized as specified in the data dictionary, using key informant feedback to identify thresholds of concern. Primary care visits showed a long tail (max=145). Two versions of primary care visits were tested, one censoring at 5+ visits and one censoring at 20+ visits. Outpatient specialty visits were censored at 10 and also tested as a flag for 0-4 vs. 5+ visits (top 85% of adolescent utilization at H+H). ED visits were tested as a continuous variable and also as a censored 0 to 5+ visits variable.

Social vulnerability index census tract data showed evidence of both skewness and kurtosis. All SVI variables were log-transformed after adding a very small number (0.0001) to facilitate log transformation of 0 values. SVI Theme variables, which combine sub-variables into thematic domains, were also multiplied by 100 before log transformation to adjust their 0-1 scale. After transformation, skewness and kurtosis were brought to <|2| across all SVI variables.

Sample Description: After the model was finalized, Training and Validation sample characteristics were compared. Characteristics reported for 2022 baseline status include age, sex, race/ethnicity, preferred language, LGBTQ+ status, social needs, homelessness, public housing, census tract poverty level and count of comorbidities using PCI categories. Utilization

for 2022 was also compared across primary care, specialty care, the ED and inpatient settings. Outcome distribution was also compared across Training and Validation samples. The mean, median and range of 2023 outcome scores is reported, as is the distribution of scores from 0 to 6+ (top 5%, aka “true high risk). Significance of differences between the samples was assessed via chi square (categorical variables) and Kruskal-Wallis test (continuous variables). Within the Training sample only, the same comparisons were replicated for predicted vs. not predicted high risk (final model) and lost to outpatient follow-up vs. not lost to follow-up. The tablegen SAS macro developed by Jeffrey Meyers was used to generate all sample and population comparison tables (Meyers, 2022). These included Tables 4.1 (Demographic, Clinical and Utilization Characteristics of Training and Validation Samples (2022)), 4.2 (Distribution of Adverse Outcomes among Training and Validation Samples (2023)), 4.16 (Patient Characteristics by Predicted High Risk Status (Training)), 4.17 (Outcome Distribution by Predicted High Risk Status (Training)), and 4.18 (Table 4.18 Patient Characteristics by Lost to Outpatient Care Status (Training)).

Data Modeling This study aimed to train an algorithm to predict the number of 2023 adverse outcomes documented in the EMR, using 2022 EMR and geographic SDOH data. To identify the right prediction modeling approach, a preliminary LASSO model was trained using linear, negative binomial and logistic approaches. LASSO selection is a simple form of machine learning which uses “shrinkage” to choose model features. While there has been significant attention on heavier machine learning applications in healthcare in recent years, including neural networks and random forest approaches, the underlying infrastructure at Health+Hospitals is not yet mature enough to support these applications. LASSO selection offered a leaner alternative.

LASSO selection uses “shrinkage,” which is a statistical technique to shrink data values, producing a more parsimonious model. LASSO regression has been used in both complex genetic studies (Fontanarosa & Dai, 2011) and more traditional chronic disease investigations

(Lupton-Smith et al., 2022). Strengths of regression using LASSO selection include its ability to handle multicollinearity and its ability to produce a simple model that is easier to implement and to understand than a complex random forest, neural network or other machine learning model. LASSO models can also be more generalizable than models produced using other methodology, because they may prevent overfitting to the test data that can occur with complex models that use many predictors (Collins et al., 2024; Gunn et al., 2023). Limitations of LASSO selection include its sensitivity to outlying observations, which are frequently seen in healthcare data and should be capped or transformed. LASSO models are also impacted by predictor variables which have very different scales. This study transformed continuous variables to align scales and control outliers, skewness and kurtosis. Because of shrinkage, a LASSO model will retain a single variable among a parameterized set, which may or may not be the most important variable in the set from an etiologic perspective and should be addressed manually.

This inquiry tested three different types of LASSO regression modeling: linear, negative binomial and logistic. All three modeling approaches were used to train a preliminary model within the Training partition. Based on the observed calibration, discrimination, parsimony and explainability of these models, a top modeling approach was chosen to build a single final model. This final model was then deployed with the Validation partition to obtain Validation performance.

Linear regression models a linear relationship between a set of predictors and a continuous outcome. It requires stricter assumptions of residual normality than logistic regression, although these are less concerning for very large sample sizes such as those seen in this study. In this study, the preliminary linear model used 2022 data to predict the number of adverse 2023 outcomes (as measured in the adverse outcome score) among 2022 adolescent primary care patients.

Negative binomial regression uses a log link to model the relationship between predictors and a discrete count response variable. Because it is used to model count data, it can

be useful in healthcare contexts. Negative binomial distributions are used instead of Poisson distributions for over-dispersed data, in which the standard deviation of the outcome exceeds the mean, such as the 2023 outcome score. The preliminary negative binomial model used 2022 data to predict the count of 2023 adverse outcomes using a log link.

Logistic regression models the log of the odds of a binary event or outcome based on predictor variable values. Logistic models are also termed binary classification models. Logistic regression requires fewer assumptions than linear regression, and clinician audiences and risk score stakeholders may be more familiar with logistic regression than negative binomial regression. However, information is lost when continuous outcome data are binarized for logistic regression, which can be a limitation. For this study, the threshold of interest was the top 5% of adverse outcomes (“true high risk”), to align with the H+H adult high risk flag (top 5%). Flagging the top 5% is also replicated in the literature as a common a-priori risk threshold absent other clear considerations (Lewis et al., 2011). The logistic model predicted the log odds of a 2022 primary care patient being in the top 5% of observed 2023 adverse outcomes, based on 2022 data. Although the model is predicting log odds of being in the top 5% of adverse outcomes, implementation of a risk model can modify alerting thresholds based on use case.

After data preparation, the first modeling step was to model the crude association of each candidate predictor and the outcome of interest for each preliminary modeling approach. Crude associations for variables selected by the final model are reported in Results. Crude associations for the preliminary models are not presented for parsimony but are available upon request.

The next step was preliminary LASSO variable selection and model fitting across linear, negative binomial and logistic models. LASSO regression requires a criteria to select the optimal shrinkage parameter, which determines how many variables enter the model. For the preliminary model testing, Schwarz's Bayesian Criterion (SBC, also referred to as Bayesian Information Criterion or BIC) was used as this criteria (Sarkar, 2015). The LASSO modeling step

at which the SBC criteria was minimized was selected as optimal. Because of computational bandwidth restrictions in the H+H environment, SBC selection was used to select a modeling modality from the preliminary models, and the more intensive five-fold cross-validation was employed to train the final model, as described below.

LASSO model selection details are presented in table 4.3, which includes selected lambda, SBC and lasso steps for the SBC-selected preliminary models. Figures 4.2-4.4 provide LASSO selection plots which show change in SBC at each modeling step. Where LASSO selection retained two overlapping versions of a variable, a single version was prioritized for retention to improve model parsimony and ease of data maintenance and flag generation in future implementation at Health+Hospitals. For example, if the preliminary LASSO model selected both the Alcohol+Substance Use flag and the separate Alcohol flag, the variable with the higher beta was retained. The preliminary linear, negative binomial and logistic models were then re-fit with SAS to generate fit statistics and parameter estimates.

The linear LASSO model was fitted using Proc GLMSelect, with a seed of 32 and SBC lambda selection. SBC was minimized at 62 LASSO steps. The linear LASSO model selected multiple overlapping variables which were manually reduced, as follows: The model selected both the Abuse or Foster flag and Abuse or Neglect Diagnosis, a sub-component of that flag. The broader Abuse or Foster flag was selected. It also selected both the Serious Mental Illness visit variable and the sub-component individual visit diagnosis flags for OCD, conduct disorder and psychosis. The individual flags had more influential beta values and were thus retained. Similarly, the model selected the developmental disorders flag and its sub-component visit diagnosis variables of chromosomal abnormalities, developmental delay and pervasive developmental disorders. The individual flags had more influential beta values and were thus retained. The preliminary model also selected both the Risky Sex flag and the STI sub-component of that flag (Risky Sex flag retained); binned ED visits, continuous ED visits and a flag for 3+ non-psych ED visits (continuous ED visits retained); binned outpatient visits and 5+

outpatient visit flag (5+ outpatient visit flag retained due to larger beta); and the School Problems flag and educational difficulties sub-component (retained School Problems flag). Finally, the linear LASSO selected both the Anemia Diagnosis variable and the Sickle Cell diagnosis sub-component. The sickle cell sub-component was retained due to its much higher beta value. The 2+ Somatic Symptoms flag and the sleep disorders diagnosis variable were selected, but both were retained because they measured overlapping but distinct domains. With these adjustments, the linear LASSO variables were used to re-fit the linear model using Proc GLM to generate fit and performance statistics.

The negative binomial LASSO model was fitted using Proc HPGenselect with the SAS default of a LASSO RHO of 0.80 and SBC lambda selection. SBC was minimized at 22 LASSO steps. Three manual adjustments were made to the variables selected via LASSO. The negative binomial LASSO selected both Depression, and the Anxiety or Depression flag. Anxiety was subbed for this flag. The Serious Mental Illness predictor was swapped for the Serious Mental Illness or ADHD predictor because the ADHD visit variable was already in the model. Lastly, the negative binomial LASSO selected two overlapping ED predictors (EDcat and EDNonsychCat), of which EDcat was retained as the more comprehensive variable. After variable adjustments, the model was re-fit using Proc Genmod to generate fit statistics as outlined below.

The logistic LASSO model was fitted using Proc HPGenselect, using the SAS default of a LASSO RHO of 0.80. SBC was minimized at 22 LASSO steps. Four adjustments were made during manual review to the variables selected via LASSO. Because depression was already retained in the model, the Anxiety variable was swapped for the Anxiety or Depression variable. Because the Drug Use or Alcohol variable was already retained in the model, the Drug Use variable was dropped. Because the ADHD visit variable was retained by the model, the Serious Mental Illness flag was substituted for the Series mental Illness or ADHD flag. Finally, the logistic LASSO selected both continuous and censored 0-5+ ED visit variables. The censored

ED variable was retained due to a higher beta value. After variable selection, the model was re-fit using Proc Logistic to generate fit and performance statistics as outlined below.

Assessing Global Model Fit: Next, a modeling approach was chosen for the final model by assessing model fit and performance. Model fit should indicate that the relationship between the predictors and outcome is well specified by the chosen modeling distribution. For the preliminary linear LASSO regression model, fit was assessed visually with an expected vs. observed outcomes plot and a residual vs. outcome plot to identify deviations from normality. Fit statistics included root mean squared error (RMSE) and adjusted R squared, to quantify the residual noise in the data and the percent of the data explained by the model respectively, as well as the mean residual, residual skewness and residual kurtosis. These statistics and visualizations were generated via Proc Univariate, Proc GLM and Proc SGPlot.

For the preliminary negative binomial model, fit was assessed visually with a model fit plot, a predicted value vs. residual plot, and an expected vs. observed outcome plot. Fit statistics included root mean squared error (RMSE), adjusted R squared and deviance value, or the ratio of the deviance to degrees of freedom (UCLA Advanced Research Computing, n.d.). The dispersion parameter was examined to determine appropriateness of the negative binomial model over the poisson distribution.

For logistic models, model fit is often termed calibration. Per Efthimiou et al, “[c]calibration refers to the agreement between observed outcomes and the model's predictions” across the spectrum of risk (Efthimiou et al., 2023). In this study, calibration was examined using a calibration plot and quantified with the calibration slope (or alpha calibration) and intercept (or beta calibration) for the preliminary and final logistic models, with slopes closer to 1 and intercepts closer to 0 indicating good calibration (Stevens & Poppe, 2020). Calibration plots and fit statistics are presented in Results.

Assessing Global Model Performance: As defined by Efthimiou et al, “Discrimination is the ability of the model to correctly rank-order patients with respect to their outcomes. For

example…among two randomly chosen patients, the one with the higher predicted outcome will also have the higher observed outcome" (Efthimiou et al., 2023). For the linear and negative binomial models, discrimination was assessed using Spearman's correlation coefficient, a rank-derived correlation measurement. For the logistic model, discrimination was assessed using the AUROC (also known as the c-statistic), which quantifies area under the receiver operating curve, aka true positive vs. false positive rates across all risk score thresholds. An AUROC of 0.5 would indicate the model is no better than random chance, while an AUROC of 1.0 would indicate perfect prediction.

Assessing Threshold-Specific Performance For NYC Health+Hospitals, the ability to determine whether or not an adolescent is at *high* risk is far more important than differentiating gradations of *low* risk. Among our adult patients, we flag the top 5% of patients as high risk using a home-grown risk algorithm. For adolescent patients, need will likely outstrip intervention availability, and there is programmatic value in starting with a smaller group of patients while intervention workflows are developed, piloted and iterated upon after the conclusion of this study (see Discussion for further implementation details). For a parallel 0-4 risk prediction study, leadership selected the top 5% as the target risk algorithm prediction threshold. This study therefore aimed to predict risk for the top 5% of high risk adolescents.

This was operationalized by labeling the top 5% of observed outcomes as "true high risk," and flagging the top 5% of predictions across each model as "predicted high risk." The distribution of the outcome of interest (0-25) did not map exactly onto percentiles. The closest 2023 outcome score threshold to the 95th percentile was 6+, which in the Training set captured the top 5.55% of outcomes. Patients with an outcome score $\geq$ 6 were defined as a case and referred to as "true high risk." To classify predictions, model predictions were rank-ordered within each model, and the top 5% of predictions were labeled as "predicted high risk." This was used instead of a predicted 6+ threshold for the linear and negative binomial models because the goal was to understand how each approach would perform under implementation constraints.

The threshold-specific performance assessment then evaluated predicted vs. observed performance for the top 5% across each modeling modality.

To calculate threshold-specific performance, true high risk patients who were also predicted high risk were classified as true positives (TP). True high risk patients who were *not* predicted high risk were false positives (FP). Patients who were *not* true high risk and *not* predicted high risk were true negatives (TN). Patients who were *not* true high risk but were predicted high risk were false negatives (FN). Threshold-specific performance was assessed via sensitivity (TP/TP+FN), specificity (TN/TN+FP), positive predictive value (TP/TP+FP), negative predictive value (TN/TN+FN), and accuracy (TN+TP/TN+TP+FN+FP). Alert rate (% predicted high risk) and event rate (% true high risk) were also reported.

Final Model After reviewing fit and performance, logistic LASSO regression was assessed as providing the best balance of parsimony, discrimination and explainability while offering moderate to strong calibration (additional details reported in Results). Logistic LASSO was selected to train the final prediction model. A five-fold cross-validation was undertaken to select model variables, following the approach outlined by Hastie et al (Hastie et al., 2009). Five-fold cross-validation produces five separate models, holding out 20% of the data each time, within the Training partition (Sarkar, 2015). The average squared error (ASE) for each fold's model is generated using that fold's hold-out data (sometimes called the 'testing' data) at each value of lambda produced by a given model step. Heterogeneity of values and performance was not evaluated across clusters in this study.

In this study, the Training data were assigned at random to five hold-out samples. Five folds were then constructed using 80% of the Training sample, with an associated 20% hold-out. A LASSO model was trained within each fold using Proc HPGenselect and then deployed on that fold's hold-out ("test") to identify that model's AIC, SBC, ASE and lambda at each model step, using the partition function. Across the five folds, test AIC, SBC, ASE, ASE standard error (SE) and lambda were averaged for each lasso step.

To select the best LASSO model, the one standard error rule was used, which aims to balance parsimony and accuracy by selecting the most parsimonious model within one standard error of the minimum ASE (Krstajic et al., 2014). Across the five folds, ASE was minimized at Step 28. To derive the final lambda, this study used the lambda from the most parsimonious model whose ASE was within minimum ASE + 1 SE. This occurred across the folds at LASSO step 21. This choice was supported graphically by the LASSO selection plot, which showed a decrease in ASE improvement after Step 21.

Because this form of LASSO modeling produces five models, selected parameters vary across folds. For this study, any variable selected by at least two folds at Step 21 was tested in the final logistic regression, yielding 22 candidate variables. Final model variables selected via 5-fold cross-validation were compared to variables selected in the preliminary SBC-based model (crosswalk reported in Results). During manual review, three cross-validation variables were removed because multiple versions of the same variable were retained across folds. The anxiety or Depression flag was dropped because anxiety and depression variables were separately retained. Drug use was dropped because it was a sub-component of the Alcohol or Substance Use flag, which was retained. Primary care visits (0-5+) was dropped because primary care visits (0-20+) was already retained. Because ADHD visit diagnosis was already in the model, Serious Mental Illness was subbed for Serious Mental Illness or ADHD flag. Eighteen of the 22 variables selected across 2+ folds were retained without modification, including the Abuse or foster care or psychosocial difficulties flag; alcohol or drug use; anxiety; serious mental illness visit flag; behavioral health visit flag; depression; developmental disorders flag; ADHD visit, asthma visit, obesity visit, ED visits, outpatient visits, primary care visits, LGBTQ+ flag, pediatric comorbidity index count, SDOH needs, sex at birth, smoking and suicidality. In addition to these variables, two etiologically meaningful variables flagged by the key informants (sickle cell visit diagnosis, eating disorder visit diagnosis) were tested to improve clinical buy-in from rendering adolescent providers. Sickle cell had a high beta value and was significant. Eating

disorder diagnosis failed to achieve significance, despite a beta higher than other retained variables, but improved AUROC and was thus retained. The final model parameters and associated odds ratios are presented in Results with the final model calibration, discrimination and threshold-specific performance.

Sensitivity Analysis: Within the Training partition, performance was reported among patients who were not lost to outpatient follow-up to evaluate model performance among patients with a lower likelihood of unmeasured missing outcome data (Rijnbeek & Reps, 2021). Among adolescents, it is not uncommon to utilize care more episodically. True loss to follow-up might realistically be defined as no primary care within 18 months. However, this sensitivity analysis was seeking to understand the impact of no patient visits during the outcome year on measurement of that outcome. Loss to follow up is therefore aligned with the outcome year (2023) rather than a clinical 18 month window. Patient baseline characteristics were compared between patients lost to outpatient follow-up in 2023 and those who remained engaged in care. Threshold-specific performance was then compared between the full Training sample and those who remained engaged in outpatient care in 2023.

Bias Evaluation: This study evaluated fairness across race/ethnicity, insurance, primary language and gender. LGBTQ+ status was determined to be of too low quality for inclusion in bias evaluation for 2022, with the hope that data are better collected in future. The fairness metrics of interest were equal opportunity difference (EOD, comparison of False Negative Rate by subgroup) and predictive parity difference (PPD, comparison of positive predictive value by subgroup), reflecting disparity in ability to be flagged as positive by the model (EOD) or to be positive if flagged (PPD), per Huang and Xu (Huang et al., 2022; J. Xu et al., 2022). Sensitivity, specificity, accuracy, event rate and alert rate were also reported by subgroup. For each sociodemographic class, the largest subgroup acted as the referent. The referents were therefore Hispanic/Latinx race/ethnicity, Medicaid insurance, female sex and English language.

The threshold of >10 percentage point difference from the referent was used to flag the presence of moderate bias for EOD and PPD.

Model Validation To produce the internal Type 2b validation, the final model was run on the partition data using SAS stored model scoring. The top 5% high risk prediction threshold obtained from the Training data (0.281) was applied to predictions within the Validation data to flag predicted high risk. Model calibration plot, calibration slope and calibration intercept; AUROC; sensitivity, specificity, positive predictive value, negative predictive value and accuracy were compared between the Training and data to assess internal validity.

Result Reporting Results were reported following the Transparent Reporting of a multivariable prediction model for Individual Prognosis or Diagnosis +AI (TRIPOD+AI) statement, which was published across 11 prominent medical journals to unify risk prediction literature in 2015 and then updated to reflect AI considerations in 2024 (Collins et al., 2015, 2024). The TRIPOD+AI questionnaire is provided in the Appendix. The APA Publication Manual (7th Edition) guidelines were used for formatting throughout, and Zotero reference management software was used for all references. No large language models (Chat GPT, Gemini, etc) were used for writing or editing this dissertation.

Chapter 4: Results

Introduction

Chapter 4 presents in depth the results of this study to train a score to identify adolescents at high risk of adverse socio-medical outcomes within primary care at NYC Health+Hospitals. This section describes the study population; outlines preliminary models tested across three modeling approaches, and describes the final model variable models, performance and bias. First, the study population is described across both Training and Validation samples using the 2022 baseline characteristics and 2023 outcome scores. Subsequently, preliminary LASSO modeling fit and performance are described for linear, negative binomial and logistic models. Next, the development of the final logistic model is described using five-fold cross-validation to select the optimal LASSO lambda. Final model parameters and performance statistics within the Training partition are described. Within the Training partition, a comparison of 2022 characteristics and 2023 outcome distribution is presented by predicted high risk status, and performance is assessed by lost to follow-up status. Bias is also evaluated by insurance, preferred language, race/ethnicity and sex at birth. Last, final model calibration and discrimination within the Validation partition are presented.

The goals of model selection within the safety net context are to craft a model that is 1) parsimonious, meaning the model is simple to run and maintain; 2) high performing, with an AUROC of $\geq$70%, sensitivity of $\geq$20%, PPV $\geq$30% and accuracy of $\geq$90% to surpass the null hypothesis; and 3) explainable, such that front line providers can understand why a patient has been identified as high risk and how a model is broadly working. Towards these goals, this inquiry tested LASSO selection across three modeling approaches: linear, negative binomial and logistic. Final model selection prioritized performance and model fit or calibration. As described in Results, the preliminary linear model was not parsimonious and showed evidence

of model assumption violation. The negative binomial and logistic models selected similar variables and showed similar performance. However, calibration/fit was slightly stronger in the logistic model, which also had higher levels of familiarity and explainability than the negative binomial model for analytic staff and providers. Ultimately, the logistic model was selected, and final model performance of this model is described in detail.

Study Population

Training and Validation Sample Characteristics

After model cleaning and imputation of missing variables, the final sample for this study consisted of 37,126 adolescents with ≥1 primary care visit to a NYC Health+Hospitals clinic in 2022. The Validation sample partition included only adolescents whose most recent primary care visit was in the Bronx (n = 10,764). The Training sample partition included adolescents whose most recent primary care visit was in any H+H clinic outside the Bronx (n = 26,362). Validation data was not examined during variable cleaning or transformation stages of modeling, and Validation sample characteristics were calculated only after modeling was complete.

As would be expected with a Validation sample chosen to mimic real-world population variation, the Training and Validation samples differed significantly across demographic, clinical and utilization criteria (Table 4.1: Demographic, Clinical and Utilization Characteristics of Training and Validation Samples (2022)). The Training sample was evenly distributed by age (range 17.1% aged 12 to 16.4% aged 17). The Validation sample skewed slightly but significantly younger, with 19.1% of the sample aged 12, and 15.0% of the sample aged 17. Sex at birth was almost perfectly even in both samples and did not differ significantly. However, the distribution of race/ethnicity was significantly different across the two samples. In the Training sample, 45.1% of adolescents were recorded as Hispanic/Latinx, followed by Black (36.2%), Other (11.8%), Asian (5.4%) and White (1.5%). In the Validation sample, more than half of adolescents were Hispanic/Latinx (57.6%), followed by Black (27.4%), Other (8.0%),

Asian (4.3%) and White (2.8%). While still significant, the differences in preferred language were smaller, with 61.3% (Training) and 59.3% (validation) of patients preferring to speak English, and 34.3% (Training) and 37.0% (validation) preferring to speak Spanish. More Training patients were commercially insured (19.9% vs. 16.2%) or uninsured (6.1% vs. 2.5%), while Medicaid was highest in the Validation sample (74.0% vs. 81.3%). Although significant, differences in LGBTQ+ status were minor, with 4.6% (Training) to 4.8% (validation) of adolescents having an LGBTQ+ identity recorded in the EMR.

Social needs differed sharply between the Training and Validation samples. Only 3.8% of the Training sample had one or more social needs documented in a diagnosis or social screening in 2022, as compared to 10.5% of the Validation sample. Homelessness was slightly lower in the Training sample (2.1% vs. 3.5%, $p < 0.001$) and the percent living in public housing was slightly higher (10.3% vs. 9.1%, $p < 0.001$). Percent living in a poverty census tract was > 25 percentage points lower in the Training sample (51.2%) compared to the Validation sample (78.7%). Based on available social determinants data, 2022 social needs were markedly higher in the Validation sample.

Utilization was also distinct across the two samples. Within the Training sample, adolescent patients had 1.9 primary care visits on average (standard deviation: 1.57), with a median of 1 and a maximum of 145. (The maximum of 145 visits was verified via chart review. The patient with that outlier was retained in the data, but the visit variable was censored in candidate predictors.) Training patients also had 1.3 average specialty visits (SD 4.01). Validation sample adolescents had fewer primary care visits (mean 1.6, SD 1.03) and equivalent specialty visits. Although inpatient visits did not differ significantly between the two samples (mean 0, SD 0.19-0.20), Training patients had significantly fewer average ED visits than Validation patients (mean 0.3 vs. 0.4, SD 0.83 vs. 0.93).

The two samples also differed significantly by chronic comorbidity status. The Pediatric Comorbidity Index count measures the number of comorbid conditions recorded via visit or

billing diagnosis in a given period. On average, patients in the Training sample had a 1.7 PCI comorbidity score documented in 2022 (SD: 1.57), but scores ranged from 0-15. In the Validation sample, patients had an average of 2.0 comorbidities (SD 1.70), and a range of 0-13. Differences in PCI distribution between Training and Validation samples were significant ($p < 0.001$). Differences in 2+ somatic symptoms were not significant across samples.

Training and Validation samples illustrated similar mental health needs. Although significant ($p=0.04$), differences in 1+ Serious Mental Illness (SMI) diagnoses were minor (5.4% vs. 4.8%). No significant difference was observed across the depression or anxiety flag (17.0% - 17.1%), suicidality (4.4% - 4.0%), or evidence of abuse or trauma (6.8% - 7.2%). However, behavioral health needs were greater across multiple domains in the Validation sample. Training patients had fewer documented pregnancies or STIs in 2022 (1.8% vs. 2.6%). Alcohol or substance use was more than five times higher in the Validation sample (3.8% vs. 21.2%). Documented school-related needs were similar across samples, with no significant difference in developmental-delay related disorders (9.1% - 9.5%) or ADHD diagnosis (5.4% - 5.2%). Documented school challenge differences were significant but minor (4.0% vs. 4.6%).

Distribution of Outcome of Interest in Training and Validation Samples

The 2023 outcome score was materially lower among Training sample patients than Validation sample patients (Table 4.2: Distribution of Adverse Outcomes among Training and Validation Samples (2023)). Training sample patients had an average 2023 adverse outcome score of 1.8 (SD 2.05), with a median of 1 and a range of 0-25. Validation sample patients had an average 2023 adverse outcome score of 2.2 (SD 2.17), with a median of 1 and a range of 0-25. The cutoff of 6+ on the adverse outcome score was chosen as the closest threshold to capture the top 5% of risk within the Training data. Within the Training sample, 5.55% of patients had a 2023 adverse outcome score of 6+, as compared to 7.19% of the Validation sample ($p < 0.001$). Additionally, 61.1% of the Training sample had a low outcome score of 0 or 1, as compared to

50.6% of the Validation sample. Across all measurements, the adverse outcome score was lower in the Training than the Validation sample.

Preliminary LASSO Modeling

Preliminary LASSO Model Selection

As outlined in Methods, the preliminary LASSO models were trained using linear, negative binomial and logistic modeling distributions, with the optimal LASSO model selected for the preliminary modeling step by minimizing the Schwarz Bayesian Criterion (SBC, also referred to as Bayesian Information Criterion or BIC). A summary of model selection steps and criteria is presented in Table 4.3 (Preliminary LASSO Model Selection Details), and Figures 4.2 - 4.4 (LASSO Model Selection Plots) illustrate the progression of SBC across LASSO modeling steps.

The linear LASSO model was the least parsimonious of the preliminary LASSO models. The linear LASSO model minimized the SBC at 62 LASSO selection steps with a lambda of 0.0115 and an SBC of 21,807. Although SBC was minimized at step 62, the LASSO selection plot for the linear model illustrates that SBC gains after step 59 were miniscule, and slope of SBC improvement decreased after step 31. The negative binomial LASSO model minimized the SBC at 22 LASSO selection steps with a lambda of 0.0074 and an SBC of 89,909. Although SBC was minimized at step 22, the LASSO selection plot for the negative binomial model illustrates that SBC gains after step 20 were minimal. The preliminary logistic LASSO model minimized the SBC at 22 LASSO selection steps with a lambda of 0.0074 and an SBC of 7720. The logistic LASSO plot in Figure 4.4 shows that after Step 21 improvement was minimal. Information criteria such as SBC (reported in Table 4.3) are used to compare nested models, with lower numbers indicating stronger models. This study did not produce nested models. It is interesting to note that SBC was dramatically smaller in the logistic model than the linear or negative binomial models. However, SBC cannot be directly compared across non-nested models or modeling distributions.

Preliminary LASSO Model Parameters

Table 4.4 (Preliminary Model Parameters for Linear, Negative Binomial and Logistic Models) presents the parameter estimates for each retained 2022 predictor variable across the three modeling distributions for the preliminary LASSO models after manual adjustment as described in Methods. Table 4.5 summarizes variable agreement across all three approaches. In the interest of parsimony, this section primarily highlights parameter commonalities and differences across modeling approaches.

A core set of predictor/domains were retained in some format by all LASSO modeling modalities. A total of 51 of 88 candidate predictors were retained by at least one of the models (Table 4.4). Twelve of the variables were retained by every model, including anxiety and depression; diagnosis of ADHD, asthma or obesity; LGBTQ+ status; age; sex, social needs flag; smoking status; primary care visits (0-10+); and PCI score (Table 4.4-4.5). Very similar variables were also retained by all models for abuse, foster care or psychosocial difficulties; alcohol and substance use, ED visits and outpatient visits. The domains of serious mental illness and developmental disorders were retained across all models, but the negative binomial and logistic models retained the parsimonious SMI and developmental disorder flags, whereas the linear model retained individual components of these flags. Two of three modeling modalities also retained neighborhood-level limited English proficiency; race/ethnicity; suicidality; behavioral health visit flag; and well child visit flag. All other predictors referenced in the parameter table were retained by only one modeling modality.

The ten most influential predictors in the preliminary linear LASSO model were (from most to least) chromosomal abnormalities, ADHD, OCD, pervasive developmental disabilities, sickle cell, outpatient visits, conduct disorder, psychosis, eating disorder diagnoses, and 1+ behavioral health visits. These top ten parameter estimates ranged from 0.970 (chromosomal abnormalities) to 0.641 (1+ behavioral health visits). This indicates that holding all other variables constant, a visit diagnosis of chromosomal abnormalities in 2022 increased the

predicted 2023 adverse outcome score by 0.970. The continuous nature of ED visits and PCI visits meant that these variables were also quite influential for patients with multiple comorbidities or acute visits. All parameters were significant except white vs. Latinx race, trauma or PTSD diagnosis and Commercial vs. Medicaid insurance.

The ten most influential predictors in the preliminary negative binomial LASSO model were (from most to least) ADHD visit, developmental disorders flag, alcohol or drug use, depression, smoking status, PCI count, abuse or foster care or psychosocial challenges, serious mental illness visit, asthma visit, and anxiety. These parameters ranged from 0.288 (ADHD) to 0.122 (anxiety). This indicates that holding all other variables constant, a 2022 ADHD visit diagnosis increased the predicted log count of adverse outcomes in 2023 by 0.288. The continuous nature of the PCI score means that this variable is also quite influential for patients with multiple comorbidities. All parameters were significant except neighborhood-level limited English proficiency, neighborhood-level single parent households, and white vs. Latinx race.

The ten most influential predictors in the preliminary logistic regression model were (in descending order) an ADHD visit, developmental disorders flag, alcohol or substance use, LGBTQ+ flag (LGBTQ+ vs. Not LGBTQ+), PCI, smoking status (Cur/Form. vs Never), depression, binned ED visits, sex at birth (female vs. male), a serious mental illness visit and LGBTQ+ flag (Missing vs. Not LGBTQ+, protective). These parameters ranged from 0.484 (ADHD) to -0.554 (Missing vs. Not LGBTQ+). This indicates that holding all other variables constant, a 2022 visit for ADHD increased the log odds of having a 2023 adverse outcome score in the top 5% (“high risk”) by 0.484. All parameters were significant except neighborhood-level limited English, age, obesity diagnosis and smoking status (missing vs. never smoker).

Linear Preliminary Model Performance

Fit and performance for the preliminary linear LASSO model is presented in Table 4.6 (Linear LASSO Model Fit and Performance Statistics), Figure 4.5 (Linear Expected vs.

Observed Outcomes, Preliminary Model) and Figure 4.6 (Residual Distribution by Predicted Outcome, Linear LASSO Model).

Linear model fit was middling. Residuals were unequally distributed across observed (Figure 4.5) and predicted (Figure 4.6) outcomes, suggesting meaningful divergence from normality. Although residual mean was 0.00, residuals showed a skewness of 2.38 and kurtosis of 14.00 (Table 4.6). This model's RMSE was 1.50.

Linear model performance was moderately strong (Table 4.6). Predicted and observed outcomes showed a Spearman correlation coefficient of 58.60% ($p < 0.001$), indicating moderate correlation of ranked outcomes (Akoglu, 2018). Mean observed and predicted outcomes were equivalent to two decimal places (1.79), although standard deviation was substantially larger for observed vs. predicted outcomes (2.05 vs. 1.39). The model had an adjusted R squared of 0.463, indicating that it explained approximately 46% of the variability observed in the outcome data.

Negative Binomial Preliminary Model Performance

Fit and performance for the preliminary negative binomial LASSO model is presented in table Table 4.7 (Negative Binomial LASSO Model Fit and Performance Statistics), Figure 4.7 (Negative Binomial Expected vs. Observed Outcomes, Preliminary Model), Figure 4.8 (Negative Binomial LASSO Model Fit Plot) and Figure 4.9 (Negative Binomial LASSO Model Residual Distribution by Predicted Value).

Negative binomial model fit was mixed. Observed deviance value per degrees of freedom was 0.9360, which approaches the ideal of 1 (Table 4.7). A fit plot showing predicted vs. observed outcome by 0-10+ outpatient visits while holding all other variables at their mean showed a relatively linear relationship (Figure 4.8). However, the spread of predicted outcomes was twice that of observed outcomes for the negative binomial model, with mean and median predicted values falling below observed values for most outcomes (Figure 4.7). Residual

distribution also meaningfully differed by predicted outcome level (4.9). This model's RMSE was 1.71.

Negative binomial model performance was moderate (Table 4.7). Predicted and observed outcomes showed a Spearman correlation coefficient of 58.12%, indicating moderate correlation of ranked outcomes ($p < 0.001$) that resembled the linear model (Akoglu, 2018). Mean observed and predicted outcomes were similar at 1.79 (SD: 1.80) vs. 1.81 (SD: 2.05). Standard deviations were closer than those observed in the linear model, suggesting that the negative binomial distribution may better capture the dispersion of the outcome count data than a linear distribution. However, the model had an adjusted R squared of 10.28, indicating that only 10% of the variability observed in the outcome data was explained by the model.

Logistic Preliminary Model Performance

Fit and performance for the preliminary logistic LASSO model is presented in Table 4.8 (Logistic LASSO Preliminary Model Fit and Performance Statistics), Figure 4.10 (Figure 4.10 - Preliminary Logistic LASSO Calibration Curve), and Figure 4.11 (Preliminary Logistic LASSO ROC Curve).

The preliminary logistic LASSO model showed good calibration (Table 4.8). A line of ideal calibration would have an intercept of 0 and a slope of 1. The calibration slope of the preliminary logistic LASSO model was 0.96 (SE: 0.02), and the calibration intercept was 0.00 (SE: 0.00), approaching ideal calibration. In a calibration curve visualizing 20 partitions of data based on predicted probability of the outcome (Figure 4.10), the bottom 90% of points appear well calibrated. The 91-95th percentile is modestly but significantly underestimated by the model. The top five percent (aka model target for “High Risk” label) is slightly overestimated by the model, but the 95% confidence interval overlaps the line of perfect calibration. The preliminary logistic LASSO model also showed good discrimination, with an AUROC value (90.66%, CI: 89.85%-91.48%) in the “excellent” range of 90-100 (Nahm, 2022).

Threshold-Specific LASSO Model Performance

Preliminary model prediction performance at the high risk threshold is presented in Table 4.9 (Threshold-Specific Performance Across Preliminary Models, Top 5%). For all models, event rate was 5.55% (capturing “true high risk” or who had a 2023 adverse outcome score of 6+) and the alert rate was 5% (top 5% of predictions across each model marked as “predicted high risk”). Logistic LASSO showed the highest sensitivity (47.74%), specificity (97.51%), positive predictive value (PPV, 52.96%) and accuracy (94.75%). Linear and negative binomial model sensitivity and positive predictive values were approximately a percentage point below that of the logistic model, and specificity and accuracy were all within a percentage point of the logistic model.

Choice of LASSO Modeling Approach

The logistic LASSO was chosen as the modeling approach for this classification task of identifying the top 5% of high risk adolescents in primary care. Both linear and negative binomial models illustrated residual patterns that suggested potential misspecification of the model. The logistic model showed strong calibration and discrimination. All three models showed comparable sensitivity, specificity, accuracy and positive predictive value at the threshold of interest (top 5% of predictions), with the logistic model performance slightly outpacing the other two models. The logistic model achieved this performance far more parsimoniously than the linear model, but was largely comparable to the negative binomial model.

Additionally, an important consideration for clinical models is their “explainability” to front line clinicians. Primary care providers need to understand and trust a model which is recommending enhanced assessment, referrals or care. Linear models are extremely explainable - “this model predicts number of days in the inpatient setting” - but the linear model performed poorly within these data. Between broadly comparable approaches, logistic modeling is a more familiar modeling approach than negative binomial modeling for both clinicians and analysts. Across fit, discrimination and explainability, the logistic model was determined to be

the strongest option with maximum explainability for clinicians and rank-and-file analysts at H+H.

Final Logistic LASSO Model

Logistic LASSO Five Fold Cross-Validation

As outlined in Methods, to train the final logistic model, five-fold cross-validation was used to select the optimal lambda. Across the five folds, the average squared error was minimized at LASSO step 28 with a cross-fold lambda of 0.00193 (Table 4.10 - Model Selection Details for Minimum ASE and Minimum ASE + 1 SE Models Using Five Fold Cross-Validation). The model which minimized cross-validation ASE retained 53 features, which is only slightly less parsimonious than the preliminary linear regression model.

The most parsimonious model within one standard error of the minimum ASE was produced across folds at model step 21, which retained an average of 23 features with a cross-fold lambda of 0.00922. This modeling step and lambda were selected to produce the final model. The LASSO selection plot showing average ASE and 95% confidence intervals at each LASSO selection step is presented in Figure 4.12 (Logistic LASSO Model Selection Plot: Five Fold Cross Validation using Average Squared Error). This plot shows that the ASE improvement rate decreased after step 21, supporting the selection of this model.

Comparison of Cross-Validation and SBC Minimization for Logistic LASSO Selection

Five fold cross-validation produces five separate models, holding out 20% of the data each time. Variables which were selected by at least two folds of the minimum ASE + 1 SE model were tested for inclusion in the final logistic model. A crosswalk of variables retained by 2+ folds during cross-validation versus SBC selection is presented in Table 4.11 (Comparison of Variables Chosen by SBC vs. Cross-Validation LASSO Selection for Logistic Model). The variables presented in this table do not align exactly with those in the preliminary logistic model (Table 4.4) and the final logistic model (Table 4.12) because this table presents the variables selected by the LASSO step, before manual refinement. These two selection approaches

yielded similar models. SBC selection alone retained age and neighborhood-level limited English proficiency. Cross-validation alone retained 0-5+ primary care visits and anxiety. All other variables were retained by both LASSO selection methods.

Final Logistic LASSO Model Odds Ratios

The variables selected via cross-validation were reviewed and adjusted for collinearity. Two etiologically valuable variables (sickle cell visit diagnosis, eating disorder visit diagnosis) were manually added. A detailed description of this process is available in Methods. Crude and adjusted final model parameters are presented in Table 4.12. Crude and adjusted odds ratios (ORs) for final model predictors are presented in Table 4.13. For the sake of parsimony, this section discusses odds ratios.

All predictor variables retained by the final logistic LASSO model showed significant crude associations with the outcome. The strongest associations were observed for the behavioral health visit flag (OR 14.73, 95% Confidence Interval 12.77-16.99), anxiety (OR 8.687, CI 7.744-9.745), ADHD visit (OR 8.284, CI 7.263-9.449), Depression (OR 8.153, CI 7.302-9.102), disordered eating (OR 8.125, CI 5.86-11.267) and suicidality (OR 7.99, CI 6.935-9.205). Serious mental illness; abuse, foster care or psychosocial needs; alcohol or drug use; and current/former smoker status all had crude ORs >5. Among the continuous variables, the strongest associations were seen with PCI (OR 2.128, CI 2.062-2.196) and binned ED visits (OR 1.986, CI 1.901-2.075). Two variables capturing missing screenings (unknown vs. never smoker, missing LGBTQ+ screening vs. not LBTQ+) were protective.

The range of ORs was much narrower in the adjusted final model. The most influential categorical parameters of the final model were sickle cell (aOR 3.06, CI 1.50 - 6.23), ADHD visit diagnosis (aOR 2.68, CI 2.23 - 3.22), developmental disorders flag (aOR 2.21, 1.85 - 2.63), and alcohol or drug use (aOR 2.12, CI 1.66-2.72). The most influential continuous parameters of the final model were PCI score (aOR for each 1 unit change 1.38, CI 1.30 - 1.44) and 0 to 5+ ED

visits (aOR for each 1 unit change 1.28, CI 1.21-1.36). In descending order of magnitude, depression, current smoker status, female sex, serious mental illness diagnosis, anxiety, asthma, the abuse or psychosocial or foster care flag, and eating disorder diagnosis all had an OR ≥ 1.5 (aOR range 1.78 (Depression) to 1.50 (eating disorder diagnosis)). The model also retained social determinant need, suicidality, behavioral health visit flag, unknown smoking status, LGBTQ+ status, outpatient visits (0-10+) and primary care visits (0-20+), with ORs ranging from 1.45 (SDOH need) to 1.09 (aOR for each 1 unit increase in primary care visits).

After adjusting for other variables, LGBTQ+ status was no longer significant (aOR 1.13, CI 0.91 - 1.39). The only significant protective factor was a missing LGBTQ+ status (aOR as compared to Not LGBTQ+ (0.46, CI 0.40 - 0.54), indicating that a patient was not screened in 2023 or responded with “unknown/declined.’ This may represent differential screening patterns by risk status of adolescents within primary care, and may also be predictive of no-show status for 2023 which would produce an artificially low risk score due to unmeasured outcome data. This is further explored in Discussion. Obesity diagnosis was also slightly protective (aOR 0.92, CI 0.80 - 1.05) but failed to achieve significance, indicating that this variable, while selected by all models and therefore strengthening model performance, is not particularly predictive of adolescent risk in this sample after adjusting for other covariates.

Final Logistic Model Fit and Calibration within the Training Partition

Fit and performance for the final logistic model within the Training partition are presented in Table 4.14 (Final Logistic Model Fit and Performance Statistics, Training Sample), Figure 4.13 (Final logistic Model Calibration Curve, Training Sample), and Figure 4.14 (Final logistic Model ROC Curve, Training Sample).

The final logistic model showed good calibration that was comparable to the preliminary logistic LASSO model (Tables 4.14, 4.8). The calibration slope of the final logistic model within the Training data was 0.96 (SE 0.015) and the calibration intercept was 0.00 (SE 0.002), which aligned with the preliminary logistic LASSO calibration fit and slope to two decimal places. The

final calibration curve showed modest but significant under-estimation in the 91-95th percentiles. In the target top 5% bucket, the final model showed slight but not significant over-estimation. This overestimation was smaller in the final than the preliminary model. The final model showed excellent global discrimination, with an AUROC of 90.70 (CI 89.89-91.51).

Final Logistic Model Threshold-Specific Performance

Threshold-specific performance for predicted vs. actual top 5% high risk is presented in Table 4.15 (Threshold-specific Performance of Final Logistic Model across Training and Validation Samples). For the top 5% threshold, the final model showed comparable sensitivity (47.61%), specificity (97.50%), accuracy (94.73%) and PPV (52.81%) to the preliminary logistic LASSO model, within the Training partition. To translate to clinically meaningful terms, this sensitivity of the final model means that the final model correctly identified 47-48 of every 100 adolescents who were truly in the top 5% of 2023 adverse outcome scores (6+). A PPV of 51.81 means that if a predicted high risk flag produced an intervention list of 100 patients, 52 of those patients would go on to have a 2023 outcome score of 6+, while 48 patients would have a score <6. This does not take into account loss-to-follow-up, whereby the true adverse outcome score of a patient is not being measured.

Comparison of 2022 Characteristics and 2023 Outcome Score among Predicted High Risk Patients

Patient characteristics by predicted high risk status are presented in Table 4.16 (Patient Characteristics by Predicted High Risk Status (Training)), Table 4.17 (Outcome Distribution by Predicted High Risk Status (Training)) and Figure 4.15 (Outcome Distribution by Predicted High Risk Status (Training)). Age, gender and race/ethnicity distribution differed significantly by predicted high risk status ($p < 0.001$). While patients not predicted to be high risk resembled the full population distributions, predicted high risk patients were older, more Hispanic/Latinx and female. Predicted high risk patients were more likely to be insured by Medicaid, and less likely to be uninsured or commercially insured ($p<0.001$). Predicted high risk patients were more likely

to have an LGBTQ+ identity (21.1% vs. 3.8%, $p < 0.001$). Predicted high risk patients were more than twice as likely to have a social need (8.0% vs. 3.6%, $p < 0.001$). Homelessness and living in public housing or a high-poverty census tract were all significantly higher among predicted high risk patients ($p < 0.001$ public housing and census poverty, $p = 0.05$ homelessness). Predicted high risk patients had significantly higher utilization ($p < 0.001$) across all settings, including primary care (3.0 vs. 1.8 avg. visits), specialty care (8.8 vs. 0.9), ED (1.5 vs. 0.3) and inpatient (0.2 vs. 0.0). Predicted high risk patients had an average of 3.7 additional PCI comorbidities in 2022 ($p < 0.001$) with a higher range of comorbidities (1-15) than those not predicted high risk (0-8).

Behavioral and mental health needs were also markedly higher among predicted high risk patients. Predicted high risk patients had higher documented rates of Depression or Anxiety (78.0% vs 13.7%), suicidality (32.2% vs. 3%) and serious mental illness (33.1% vs. 3.9%). Documented abuse or trauma was significantly higher among predicted high risk patients (40.9 vs. 5.0%), as were alcohol or substance use (20.3% vs. 2.9%) and tobacco use (13.6% vs. 1.7% current/former). Documented pregnancy or STIs were low for both groups, but higher among predicted high risk (5.2% vs. 1.6%). Rates of diagnosed ADHD (37.8% vs. 3.7%) and documented school difficulties (11.2% vs. 3.6%) were also more than three times higher among predicted high risk patients. These differences were significant at $p<0.001$.

Distribution of 2023 adverse outcome scores differed significantly by predicted high risk status (Table 4.17). Fully 52.8% of predicted high risk patients went on to be top 5% high risk, as compared to 3.1% of those not predicted to be high risk ($p < 0.001$). Predicted high risk patients had an average 2023 outcome score of 6.3 (SD: 3.48), with an observed range of 1-25. Patients not predicted to be high risk had an average 2023 outcome score of 1.6 (SD: 1.63), with scores ranging from 0-20. Only 1.4% of high risk patients had a 2023 outcome score of 0-1, as compared to 64.2% of those not predicted to be high risk.

Lost to Outpatient Care Sensitivity Analysis

As mentioned in Methods, it is important to retain lost to follow-up individuals in model Training data, but performance can be further quantified by a sensitivity analysis by lost to follow-up status. Within the primary care context, adolescents that fall out of outpatient care are a particular concern. Table 4.18 (Patient Characteristics by Lost to Outpatient Care Status (Training)) provides a comparison of 2022 characteristics by lost to follow-up status, and Table 4.19 (Sensitivity Analysis for Top 5% Performance by Lost to Outpatient Care Status (Training)) provides a comparison of top 5% performance. In the Training set, 26.5% of 2022 adolescent primary care patients received no primary or specialty care in 2023 (n = 6,990). These patients differed significantly by age, sex at birth, race/ethnicity, preferred language, LGBTQ+ status, homelessness, all utilization *except* inpatient visits, and PCI comorbidities ($p < 0.001$ except for sex at birth, $p = 0.003$). No significant difference was observed by social needs, percent living in public housing, or percent living in a high poverty census tract. Particularly marked differences among those who were lost to outpatient care included older age (20.7% aged 17, versus 14.9% among those not lost to outpatient care), English language preference (66.1% vs. 59.6%), missing LBGTQ+ screening data (67.7% vs. 51.6%) and chronic comorbidities (1.3 vs. 1.8 average PCI score). Note that lost to outpatient care was defined as no outpatient care within the system in 2023, which does not differentiate true loss to follow-up, patients who sought care in the 12-18 month window rather than within the year, or patients whose care transferred to another system or to their university. Lost follow-up patients had fewer 2022 average primary care visits (1.6 vs. 2.0), specialty visits (0.5 vs. 1.6) and ED visits (0.3 vs. 0.4).

Among patients lost to outpatient care, 2.3% were flagged by the final model as high risk and 1.16% went on to have an observed top 5% high risk outcome. Among those who remained in outpatient care, 6.0% were flagged as high risk, and 7.1% went on to be in the top 5% of outcomes. For those who remained in care, sensitivity, specificity and NPV were broadly comparable to the full Training sample performance, while PPV was four percentage points

higher (57.01%) and accuracy was a percentage point lower (93.72%). Among those lost to outpatient care, specificity, accuracy and NPV remained high but sensitivity dropped by 7 percentage points (40.74%) and PPV dropped by more than half (21.02%). Put in clinically relevant terms, at the top 5% prediction threshold, the final model was materially better at identifying high risk outcomes among “sticky” patients who remained in outpatient care than it was for patients who missed outpatient care in 2023. An exploration of these differences and their implications for implementation is offered in Discussion.

Final Logistic Model Fairness

Model performance and bias across insurance, language, race/ethnicity and sex at birth is presented in Table 4.20 (Sociodemographic Bias Evaluation for Top 5% High Risk (Training)). Event rate (true top 5%) and alert rate (predicted top 5%) varied by sociodemographic group. Event rate was highest among patients with Medicaid; English and Spanish speakers; Black/African American or Hispanic/Latinx individuals; and females. Alert rates were also highest among these groups (Table 4.20). Using predictive parity, bias was seen by insurance (uninsured) and language (other). Using equal opportunity, bias was seen by insurance (uninsured) and race/ethnicity (AAPI).

As outlined in Methods, this study used a threshold of a 10 percentage point difference in fairness metric to flag bias. The fairness metrics were predictive parity (PPV difference of sub-group vs. referent) and equal opportunity (false negative rate (FNR) difference of sub-group vs. referent). The largest sub-group in each category was used as the referent, including Medicaid insurance, English language, Latinx race/ethnicity, and female sex.

Across both predictive parity and equal opportunity, the most severe bias was seen among uninsured patients. Uninsured patients had a PPV 14.5 percentage points below the referent and an FNR 25.3 percentage points above the referent. This means that true high risk uninsured patients were less likely than Medicaid patients to be flagged as high risk by the algorithm. It also means that uninsured patients appearing on a high risk intervention list would

be less likely than Medicaid patients to be truly in the top 5%. In brief, the biases observed in this study indicate that uninsured patients from the 2022 training sample were less likely to be flagged *and* less likely to be flagged correctly by this risk score. This may translate to decreased risk score performance among newly arrived asylum seeking patients or other patients whose parents have not enrolled them in CHIP Medicaid insurance, for which all children in New York were eligible in 2022. Likely sources of bias and approaches to mitigate bias before implementation are discussed in detail in Chapter 5.

Final Model Performance on Validation Sample: Assessing Internal Validity

Validation Calibration and Discrimination

To approximate real-world performance of a risk score on naive data in which it was not trained, a Validation partition is withheld during model Training. The final model is then deployed within the Validation partition to assess validity. Although an ideal Validation sample would be external, this study used geographic partition to approximate a pseudo-external validation (type 2B). As outlined in Study Population and Tables 4.1-4.2, the Validation sample differed across most 2022 characteristics and showed a higher burden of true high risk (7.2%) than the Training sample (5.6%). This suggests that the Validation sample successfully captured meaningful population variation.

The calibration and discrimination seen within the Validation sample are presented in Table 4.21 (Validation Calibration and Discrimination), Figure 4.16 (Validation Calibration Curve), and Figure 4.17 (Validation ROC Curve). The Validation calibration was strong. Within the Validation partition, the final model showed a slightly stronger calibration slope (1.03, SE 0.02) and comparable calibration intercept to the Training partition (Tables 4.21, 4.14). The Validation calibration curve (Figure 4.16) showed significant under-estimation in the 86th-90th percentiles. However, model underestimation was not significant in the top ten percent. The final model showed good global discrimination in the Validation partition, with an AUROC of 89.30% (CI 88.09% - 90.52%). This was a decrease of 1.40 percentage points from the AUROC

observed in the Training sample, but Training and Validation AUROC confidence intervals overlapped.

Validation Threshold-Specific Performance

Threshold-specific performance for predicted vs. actual top 5% high risk is presented in Table 4.15 (Threshold-specific Performance of Final Logistic Model across Training and Validation Samples). Observed 2023 outcome scores were higher in the Validation than the Training sample, with 7.2% versus 5.6% of patients meeting the true high risk definition (“event rate”). Using the top 5% prediction threshold from model Training to flag predicted high risk in the Validation sample, the alert rate for the Validation sample was 6.0% (Table 4.15). At this top 5% threshold, the final model had a Validation sample sensitivity of 46.64%, a specificity of 97.12% and accuracy of 93.49%. This represents a very modest decrease in performance from the Training dataset (0.38 - 1.24 percentage points). Positive predictive value improved within the Validation sample (55.62% vs. 52.81% Training). To translate to clinically meaningful terms, while an intervention list of 100 patients within the Training sample would contain 52 true positives, an intervention list within the Validation sample would contain 55 true positives. This may reflect the increased prevalence of the high risk outcome (6+ days) in the Validation sample, as positive predictive value is influenced by prevalence. Within the Training sample, the final model correctly categorized 94-95 of every 100 people. Within the Validation sample, the final model correctly categorized 93-94 of every 100 people, a very small decrease in performance within naive data from the same system.

Chapter 5: Discussion

Introduction

Chapter 5 provides a detailed discussion of the results, contributions and limitations of this study. Summary of Results reviews preliminary and final model performance, outlines Training and Validation cohorts, provides key statistics and highlights model bias. This section also answers the research questions posed in Chapter 1. The What Does This Study Contribute? Key Findings in Context section identifies final model strengths, outlines how outcome and predictor variables aligned with domains in the literature, explores variables which were *not* retained by the final model, and establishes potential sources of algorithmic bias and bias remedies. Real-World Implications explores model implementation at H+H. This section identifies potential workflows, highlights past risk score implementation, provides detailed implementation steps and explores a high-level logic model for how a primary care adolescent risk score could improve health and wellbeing at H+H. Limitations details the study limits and shortcomings, with a particular focus on generalizability, choice of a locally defined outcome of interest, analytic approach and missing data. Further Research explores the next phases of work across bias mitigation, temporal validation, implementation, patient outcomes and evolution of the field with the emergence of AI. Finally, the Conclusion section summarizes the study as a whole and offers key study takeaways.

Summary of Results

This section of Chapter 5 provides an overview of model results and answers the research questions posed in Chapter 1. This section outlines the modeling approach, final model parameters, observed model biases, and performance within the Training and Validation partitions. It then answers the study research questions, identifying the demographic, SDOH, utilization and clinical variables that predict risk among adolescents in primary care at H+H.

Overview

This study sought to test whether the risk of adverse health outcomes could be predicted among adolescents in primary care at a large urban safety net system, using the prior year's demographic, socioeconomic, utilization and clinical data. The null hypothesis was that a risk model could not be developed to predict high risk status (top 5% high risk, defined as ≥6 Adverse Outcome Score in 2023) with an ≥70% overall AUROC and a ≥90% accuracy, ≥20% sensitivity and ≥30% positive predictive value.

After exploring linear and negative binomial regression modalities, this study trained a binary logistic regression model using 5-fold cross-validated LASSO selection. The model was trained on geographically partitioned data (Bronx primary care vs. all other boroughs) with 71.0% of adolescents in the Training set and 29.0% in the Validation set. Predictors selected by LASSO in two of the five folds were tested in the final logistic model. Where overlapping predictors were retained across folds, the strongest predictor was retained. Finally, sickle cell and eating disorders were tested in the model and retained based on key informant feedback that these variables were important for inclusion to maximize provider buy-in.

The final logistic model retained 21 variables across 23 parameters. Within the Training data, the AUROC of the final model was 90.7 (0.89.9 - 91.5), and calibration slope and intercept were 0.959 and 0.002 respectively. At a top 5% prediction threshold, the final model showed a 47.6% sensitivity, 97.5% specificity and 52.8% positive predictive value within the Training sample. Top 5% performance was better for patients *not* lost to follow-up across sensitivity and positive predictive value. The model showed evidence of bias among uninsured patients (PPD, EOD), Asian patients (EOD only) and those who preferred a language other than English or Spanish (PPD only). These biases will need to be investigated and mitigated before implementation.

The model showed minor performance decreases within the Validation cohort. AUROC in this cohort was 89.3 (88.1-90.5), with a calibration slope and intercept of 1.030 and 0.002

respectively. At the top 5% risk threshold, the model showed a 46.6% sensitivity, 97.1% specificity and a 55.6% positive predictive value. Using the Validation statistics, this model exceeded performance thresholds established by the null hypothesis. This study suggests that risk of adverse events as defined by NYC Health+Hospitals can successfully be predicted for the top 5% of adolescents in primary care within a large urban safety net system using EHR and publicly available data, assuming acceptable post-processing bias mitigation.

Answering the Research Questions: Which Demographic, SDOH, Utilization and Clinical Variables Predict Risk among Adolescents in Primary Care at H+H?

This section delineates what predictors were retained by the final risk score, organized by research question. The Introduction posed the following research question for this study: "How accurately can the risk of adverse health outcomes for adolescents in primary care at a large urban safety net system be predicted by the prior year's electronic medical record and public data across demographic, SDOH, utilization and clinical domains?" It then established sub-questions, to identify significant predictors across 1) demographic, 2) SDOH, 3) utilization and 4) clinical variable domains. The final risk model retained 3 demographic parameters, 1 SDOH parameter, 4 utilization parameters and 15 clinical parameters.

The final model retained two demographic variables disaggregated into three parameters (LGBTQ+ vs. Not LGBTQ+; Missing Screening vs. Not LGBTQ+; sex at birth). The model retained one composite SDOH variable over individual SDOH factors such as homelessness or food insecurity. This composite variable incorporates a positive social determinants of health screening, a SSHADESS food insecurity screening, a homelessness indicator, and a NYCHA public housing address. This composite was likely retained over individual flags because LASSO selection is penalized to create a more parsimonious model.

The final model also retained four utilization variables, including ED visits (0-5+); outpatient visits (0-10+); BH visit flag (0 vs. 1+) and primary care visits (0-20+). Outside the behavioral health visit flag, these indicators are continuous variables that were censored to

improve skewness and kurtosis, at clinically meaningful or population-derived cutpoints. These variables are heavily weighted for patients who frequent primary care or outpatient care, thereby capturing patient clinical complexity, a key domain of concern per key informants. The model retained 0-5+ categorical ED visits, which will capture, by proxy, patients using the ED for primary care, or patients with emergent behavioral health needs. The model did not retain any inpatient or well child visit indicators.

Of the 23 parameters retained by the model, 15 were primarily clinical. Behavioral health was a large contributor to the model, across mental health (anxiety, depression, suicidality visit diagnoses or screenings; serious mental illness visit diagnosis); smoking, alcohol or substance use (screening or visit diagnosis); and healthy eating (eating disorder or obesity visit diagnosis). Although only one composite abuse, neglect and trauma variable was retained, it was highly influential. The model retained two school challenge variables (developmental delay-related visit diagnosis, ADHD) which could also be categorized under behavioral health. Within chronic disease, the model retained asthma and the PCI comorbidity count, which is highly influential for multi-morbid patients. Sickle cell and eating disorder diagnosis variables were manually added outside of the LASSO selection process to address key informant concerns.

What Does this Study Contribute? Key Findings in Context

As outlined in Chapter 2, this study draws from deep bodies of literature on adolescent development and risk, social determinants of health, adverse childhood experiences and the role of primary care in adolescent health. This section situates the findings of the study within those literatures. It delineates key study strengths. It outlines how the 2023 adverse outcome score and 2022 model predictors align with the domains of adolescent risk identified in the literature. It also explores the variables that were not selected by the model, placing omissions into clinical context. Finally, it surveys the sociodemographic biases observed in this study and identifies remedies to address them during implementation.

Final Model Strengths

This section outlines the final predictive model's key strengths. Model weaknesses are detailed in Limitations. Top model strengths included the use of an outcome customized to clinical practice at H+H; consistency of results observed across modeling approaches; differentiation of Training and Validation samples; minimal Training to Validation performance degradation, strong Validation performance overall and at the threshold of interest, and use of the PRISMA+AI guidelines.

There were several important strengths observed in this modeling approach. First, this model predicted a locally tailored outcome that was based on ACEs literature but refined locally by providers and patients, maximizing its applicability to clinical practice at H+H. Duke's Framework for Algorithm-Based Clinical Decision Support cites the importance of building clinical practice algorithms locally, to avoid the siloing of "commercial model developers and academic researchers on one side and clinical teams and patients on the other" within development (Bedoya et al., 2022). Including clinicians can strengthen algorithm construct validity, pointing developers towards complete, reliable, meaningful data (Ehrenstein et al., 2024). Patients also want to participate in problem identification, as Adus et al found in a series of AI and algorithm focus groups (Adus et al., 2023). This study anchored outcome development on provider and patient feedback. This approach allowed the study to craft an EHR-based outcome which included social, behavioral and medical risk despite the paucity of a validated, holistic adolescent risk score in the literature. Bringing these voices back to the table to design implementation will be important in the next phase of research.

Second, a reasonable consistency of results was observed across modeling approaches. This study tested logistic, linear and negative binomial approaches to arrive at the final prediction methodology. Negative binomial and binary LASSO preliminary models selected a collective 24 variables across 29 parameters, using SBC selection. Of these, all but two (negative binomial) to three (binary) variables were shared across both models. Within logistic

LASSO modeling, 5-fold cross-validation and SBC minimization also yielded similar results. SBC was minimized at model step 22 (lambda=0.0074), while five-fold cross validation identified the optimal ASE + 1SE at model step 21 (lambda=0.0092). This increases confidence that the lambda selected for LASSO variable selection was reasonable.

Third, this study's Type 2B pseudo-external validation using geographic partition of H+H patients by borough captured distinct populations in the Training and Validation samples. The goal of geographic partition is to capture distinct populations, to test model performance on data it has not yet met (Collins et al., 2015, 2024). This is distinct from a true external validation using data from another system, which is the gold standard for external validity. In this sample, sharp differences were observed between Training and Validation samples across social needs (3.8% vs. 10.5%), percent living in poverty (51.2% vs. 78.7%), chronic disease burden (1.7 vs. 2.0 PCI count) and primary care utilization (1.9 vs. 1.6 avg visits). Training and Validation samples also differed sharply by documented alcohol or substance use, although other behavioral health factors were similar. On average, Validation patients were sicker but had fewer visits. Average 2023 adverse outcome scores were also lower in Training than Validation (1.8 vs. 2.2 avg. score). This suggests the Training and Validation samples captured distinct populations within the safety net and lends strength to the pseudo-external type 2B validation done in this study. External generalizability is still limited by both the lack of a validation population outside of NYC Health+Hospitals and the tailoring of the outcome of interest to NYC Health+Hospitals' context via key informant interviews.

Fourth, performance degradation from Training to Validation samples was minor. Validation calibration slope improved over Training calibration slope, and intercepts were almost equivalent. AUROC measures performance across all possible risk thresholds. AUROC decreased modestly (1.4pp), and Training and Validation AUROC confidence intervals overlapped. This indicates broadly commensurate performance between Training and Validation samples across all risk thresholds.

Sensitivity and specificity are measured at specific thresholds, and indicate the performance of the algorithm as it would be implemented. Sensitivity at the 5% high risk threshold decreased by a single percentage point, from 47.6% in the Training sample to 46.6% in the Validation sample. Specificity was almost unchanged (97.5% vs. 97.1%). Positive predictive value was higher in the Validation than the Training sample. This is unsurprising given that outcome scores were higher in Validation and PPV is impacted by outcome prevalence (Monaghan et al., 2021). The consistency of performance from Training to Validation may indicate that the algorithm will yield reasonable performance when implemented on data it has not yet met, although regular model evaluation and tuning for temporal drift will be essential.

Fifth, observed performance was high for the clinical context of deployment at NYC Health+Hospitals. The choice of risk threshold (top 5% of predicted vs. observed risk within Training) was highly tailored to NYC Health+Hospitals. H+H only has resources to identify and treat the top 5% of high risk patients across pediatrics and adult medicine. The adult high risk algorithm trained by NYC Health+Hospitals also predicts top 5% high risk. Validation performance of this algorithm (30.4% sensitivity) was lower than that observed in this study (Li et al., 2023). This suggests that observed performance meets clinical deployment needs at H+H.

This study's algorithm performance should be interpreted conservatively (see Analytic Approach in Limitations for further discussion). However, observed performance within the Validation cohort of this study was strong. The AUROC observed in this study was very good to excellent (Nahm, 2022), and it exceeded reported AUROC ranges for common predictive algorithms such as LACE (Rajaguru et al., 2022; Struja et al., 2020). It is difficult to compare the threshold-specific performance of this adolescent primary care algorithm to external literature. Some algorithm literature reports algorithm performance against reference standards, which were unavailable for this study (Ehrenstein et al., 2024). The sensitivity of this algorithm was higher than a Framingham study which predicted top 5% risk for 10-year cardiovascular risk

(16.9% sensitivity at ≥40% Framingham score) (Brindle et al., 2005). However, cardiovascular risk events are not directly comparable to adolescent risk. The model was well-calibrated within the Validation sample for the population of interest (top 5% risk).

Finally, this study followed TRIPOD+AI reporting guidelines (Collins et al., 2015, 2024), which were published across 11 journals in 2014 and 2024 to unify algorithmic reporting. The TRIPOD checklist in the appendix identifies the pages on which each TRIPOD criterion is addressed. The use of a standard protocol to organize both this research and its reporting is important to ensure that study results are robust, meaningful and replicable.

Outcome Domains

As outlined in Chapter 2 (Literature Review), there is no standard definition of risk for adolescents in primary care. Expert bodies and leaders in the field enumerate sources of risk across development, behaviors, mental health and physical wellbeing. Because there is no one standard definition of risk, this study pursued a hyper-local and tailored approach by starting from a defined evidence base (adverse childhood outcomes associated with ACEs exposure) and asking key informants to tailor outcomes of concern to an urban safety net adolescent population. This approach may provide a template for other systems looking to craft a definition of adolescent risk that is tailored to local clinical realities but encompasses social, behavioral and physical health.

The final 2023 adverse outcomes used to train the risk score covered many of the principle domains identified in the literature. From the WHO's 2024 inaugural set of global adolescent health indicators, the 2023 adverse outcome score directly captured alcohol, cannabis and nicotine use; HIV and STI prevalence; food insecurity; health services use; depression and anxiety; and suicidality (The Adolescent Health Indicators Recommended by the Global Action for Measurement of Adolescent Health, n.d.). It also captured proxy indicators of WHO's healthy weight, diet and exercise; early sexual initiation and use of contraceptives; and exposure to violence. From the domains of self-reported risk surfaced on the NYC Youth Risk

Behavior Survey, the 2023 adverse outcome score in this study captured sexual health; vaping, alcohol and substance use; violence and neglect; mental health; food insecurity; asthma; insomnia; and diet and physical activity (NYC Department of Health and Mental Hygiene, n.d.). The score also retained variables across each of the "sexual risk-taking, substance use [and] …mental health risk" domains identified by the IOM Committee on the Science of Adolescence (National Research Council, 2011).

Some risk domains could not be captured in the EHR. IOM's "risky driving" domain was not recorded in the clinical data, although this factor may be less relevant in New York City. EHR data also proved a limitation in capturing bullying (cited by IOM, WHO & YRBS). Per YRBS, 17% of NYC adolescents reported being bullied at school in 2019, while 14% reported being bullied online (NYC Department of Health and Mental Hygiene, n.d.). This is almost certainly an issue for H+H patients. Without routine screening or enhanced data matching with the Department of Education, bullying will remain poorly captured in the EHR. The 2023 adverse outcome score also failed to capture WHO's poverty and educational opportunities, due to poor EHR data quality in these domains, although included variables like SDOH needs were likely correlated with household poverty.

Incarceration or justice involvement (WHO, YRBS, IOM) was also not included in the 2023 adverse outcome score. H+H captures short-term incarceration for adults via a data match with Rikers, a city jail. Recent efforts to move adolescents off Rikers mean that this data match did not identify any patients in this cohort. It may be possible to establish data exchanges with state and city partners that better capture adolescent justice involvement, which impacts lifetime wellbeing (Boch et al., 2019; Folk et al., 2021; Graf et al., 2021).

Building a good risk score requires a meaningful outcome. As outlined here, there was no gold standard definition of generalized risk for the primary care adolescent available in the literature or the EHR. Because a global, validated measure was not available, this study crafted its definition of adverse outcomes using safety net patient and provider perspectives and a

national body of ACEs literature. While the resulting score may have more limited generalizability outside the safety net context, the tailored outcome selected by this study reflects risk domains identified by national and international bodies, using EHR and available public data.

Predictor Domains

As outlined above, the final risk score retained 3 demographic parameters, 1 SDOH parameter, 4 utilization parameters and 15 clinical parameters, spread across medical and behavioral sources of risk. These parameters aligned with many of the major domains of risk outlined in Outcome Domains. Five final model parameters measured mental health and traumatic stress (IOM, WHO, YRBS). Two parameters addressed the alcohol, substance use and tobacco domain (IOM, WHO, YRBS).Two parameters aligned with the healthy weight, diet and exercise domains, and one composite SDOH need flag captured the food insecurity domain (WHO, YRBS). Four parameters also covered the WHO health services use domain. The PCI comorbidity count did not align neatly with any of the identified risk domains across IOM, WHO or YRBS (but did align with H+H key informant request to include chronic complexity). Asthma was included in the YRBS, while sickle cell did not align with the identified literature domains. While no bullying variables were available in the EMR, the developmental delay and ADHD diagnosis variables that were retained are likely correlated with school difficulty (IOM, WHO, YRBS). LGBTQ+ flag and sex captured gender and sexual orientation domains of identity (Adolescence, 2011; Alderman et al., 2019; Allred, 2024; BALOCCHINI et al., 2013; Saewyc, 2011).

This study did not include an ACE screening. H+H does not screen for ACEs across pediatric primary care. There is a paucity of evidence supporting effective interventions to pair with ACE screening at the individual level in pediatric primary care, and routine ACE screening is not recommended by the American Academy of Pediatrics (Stirling et al., 2024). Screening for trauma without offering meaningful, evidence-based resources to patients and families risks

patient trust and exposes providers to moral injury. However, the final model did select variables which align with many ACE-related outcomes cited in the literature. The model was heavily weighted towards mental and behavioral health outcomes such as ADHD (Bomysoad & Francis, 2020), anxiety (Balistreri & Alvira-Hammond, 2016; Bomysoad & Francis, 2020; Elmore & Crouch, 2020), depression (Anderson et al., 2022; Blum et al., 2019; Bomysoad & Francis, 2020; Brockie et al., 2015; Elmore et al., 2020; Elmore & Crouch, 2020), substance use (Bomysoad & Francis, 2020; Brockie et al., 2015; Folk, Kemp, et al., 2021), suicidality (Anderson et al., 2022; Brockie et al., 2015; Folk, Kemp, et al., 2021), learning or behavioral problems (Burke et al., 2011) and overall poor mental health (Baldwin et al., 2021; C. Bethell et al., 2019; Boch et al., 2019; Folk, Ramos, et al., 2021). It also included ACE-associated chronic disease outcomes, such as asthma (Bellis et al., 2018; Oh et al., 2018), chronic comorbidity (Kerker et al., 2015) and obesity (Oh et al., 2018). As mentioned in the previous section, sexual health variables such as STI prevalence or pregnancy were a surprising omission, particularly given the strong association of sexual risk taking and ACE exposure (Brown et al., 2015; Hillis et al., 2001; Lin et al., 2011; Richter et al., 2014; Folk, Kemp, et al., 2021; Oh et al., 2018). The SDOH composite flag, including housing, food insecurity and screened social needs, imperfectly captures some of the sources of intergenerational SDOH-mediated ACEs, as outlined in this study's conceptual framework from the Institute for Health Equity (Allen, Matilda & Donkin, Angela, 2015). In summary, the final model includes variables which measure major sources of risk from the ACES and SDOH literature, as well as local, national and international governing bodies.

What Fell Out? Variables Not Retained by the Final Risk Model

The final risk model did not retain age, language, insurance status or race/ethnicity. At first glance, this is surprising. These variables may be etiologically meaningful. For example, older age may yield increased risk-related behaviors, and race/ethnicity or preferred language membership may expose patients to increased racism or micro-aggressions, which increase

traumatic stress (Shonkoff, Slopen, et al., 2021). However, predictive models are selecting for predictive signal, and not causality. LASSO selection will remove collinear or less predictive variables, regardless of etiology. Thus, etiologically meaningful variables like age were dropped in favor of other collinear variables with more powerful signals.

Within this study, directly measured EMR variables appeared to be more powerful predictors of adolescent risk than geographic SDOH proxy variables. This study tested Social Vulnerability Index (SVI) proxy neighborhood measures of poverty, income, education, housing quality, single parent household and languages spoken, matched to patient census tracts. The SVI data were included to approximate missing SES data from the EMR, and to approximate neighborhood-level factors impacting adolescent development within the Community level of the socio-ecological model. However, the final model did not retain any of these geographically matched variables. It is possible that data were not accurately representing patient SES. For example, many patients attending Bellevue live in public housing complexes in Murray Hill, an otherwise affluent neighborhood. These patients may appear higher income on SVI than they truly are. It may also be that these variables were ultimately not influential enough to stay in the model because of the relative homogeneity of income within a safety net population, as compared to a commercial population. These findings align with a cardiovascular risk prediction study by UMass which found heterogeneous effects of neighborhood-level SDOH variables on algorithmic performance and bias (Liu, 2025).

Of all the chronic diseases included in the PCI (chronic GI issues, diabetes, cardiovascular disease, epilepsy, asthma), the model retained only asthma. Other chronic disorders are still included in the PCI comorbidity count. The model similarly retained ADHD but dropped component flags for important behavioral, mental and developmental health conditions such as autism (aka pervasive developmental disorders), OCD, psychosis or conduct disorders in favor of the composite developmental delay and serious mental illness diagnosis flags. This approach makes the model easier to run and routinely operationalize, but may blunt its ability to

target specific subgroups. Subsequent segmentation or “phenotyping” of the high risk population by disease status may be fruitful to guide programming and identify levers of intervention.

The model also did not retain predictors measuring risky sexual behavior. In this study, available variables included pregnancy and gonorrhea, chlamydia or syphilis lab, diagnosis or medication. The study excluded HIV due to sensitivity of the data and trichomoniasis due to lack of a validated case definition within H+H. SSHADESS sexual health screenings were also too incomplete to include. Sexual risk-taking is a major source of risk among adolescents, with important consequences for reproductive health (Alderman et al., 2019; Klein et al., 2020). Chlamydia and gonorrhea are prevalent nationally, with 2023 rates of 1,711.7 and 406.2 respectively per 100,000 adolescents aged 15-19 (Centers for Disease Control and Prevention (CDC), 2024). The LASSO algorithm may have excluded sexual health variables because of the relatively small N of patients with a positive response. However, given the importance of sexual health in adolescent risk, this exclusion may also suggest that indicators of risky sexual decision-making are poorly captured in the EHR as of 2022/2023.

It is also notable that the model did not retain the composite resilience variable, which included SSHADESS screenings for exercise and social support. Sources of resilience such as trusted relationships with adults, close social connections, supportive school and team-based environments and community support are essential in mitigating the impact of ACEs exposure on lifetime morbidity and mortality (Bellis et al., 2018; Giovanelli et al., 2020; Hall et al., 2021; Shonkoff, Boyce, et al., 2021; Sosnowski et al., 2021). These details may be documented in free text (or not at all). As AI increasingly changes the healthcare landscape and digital scribes improve our ability to document the encounter, it will be important to explore whether more comprehensive versions of resilience can be measured using EH data, without adding to provider screening burdens.

The model also did not retain the composite indicator for school difficulty, including SSHADESS screening for challenges in school or poor school performance or a diagnosis of

educational difficulties. As noted elsewhere, school performance is poorly documented in the EMR. The WHO and IOM both cited bullying as an important construct of adolescent health, but there were few EHR variables available to capture this construct. This area may similarly benefit from improved natural language processing and encounter voice capture technology advances, to harness informal provider-patient conversations already happening about school.

It is also interesting to note that the model did not retain the composite somatic symptoms predictor, nor its individual component flags (disordered sleep, nausea/vomiting, migraine/headache). The connection between somatic symptoms and exposure to trauma is well established (Herzog & Schmahl, 2018; Oh et al., 2018; Pfaltz et al., 2022). Some adolescents experiencing depression or anxiety may describe somatic symptoms before expressing sadness, particularly adolescents from Asian communities (Kim et al., 2019). While this variable may have been too non-specific to provide predictive power in this model, monitoring for somatic symptoms among high risk adolescents is still important for culturally competent clinical practice.

Addressing Algorithmic Bias

Algorithmic bias is ubiquitous, reflecting upstream inequities in healthcare access and community investment (Obermeyer & Topol, 2021). Although Validation calibration and performance was strong in this study, bias above the 10 percentage point threshold was observed for uninsured patients across both predictive parity (PPD) and equal opportunity (EOD). Bias was also observed for patients speaking a language other than English or Spanish (PPD only) and patients with Asian race/ethnicity (EOD only). Bias must be addressed and minimized prior to implementation.

The field of algorithmic fairness research is fast evolving. The TRIPOD 2015 statement was updated to TRIPOD+AI in 2024 and now requires that predictive models report their performance variation across sensitive attributes (Collins et al., 2024). There is increasing

regulatory consensus that bias must be addressed (Mackin et al., 2025a). Bias can be addressed in pre-processing, in-processing or post-processing stages (Huang et al., 2022)

Among bias mitigation techniques, the family of threshold adjustment approaches may be the easiest to implement (Mackin et al., 2025b). With threshold adjustment, bias is assessed by comparing a metric of interest (often equal opportunity or predictive parity) between a reference group and all other subgroups within a sociodemographic class. For sub-groups showing significant bias, the “high risk” threshold at which an intervention or flag is triggered is adjusted to bring the fairness metric closer to the reference group, making performance fairer.

In this study, bias was defined as >10 percentage point difference in either equal opportunity (using false negative rate) or predictive parity (using positive predictive value). EOD disparity indicates that true positives within a given sub-group have a worse chance of being flagged as high risk than those in the reference group. The impact of this kind of disparity is that patients who truly need an intervention have a lower likelihood of being flagged for it, based on sociodemographic group. PPD disparity indicates that individuals in the sub-group who are flagged as high risk by the algorithm have a lower likelihood of going on to be a true positive than those flagged within the reference group. The impact of this kind of disparity is that patients who are flagged have less chance of benefiting from an intervention based on their sociodemographic group. Because these two measures have distinct statistical properties, it may be impossible for a given algorithm to achieve both predictive parity and equal opportunity, and many studies choose a primary metric based on the implementation use case (Xu et al., 2022).

This study examined both predictive parity and equal opportunity. Equal opportunity is a common metric for use cases in which an intervention is low cost and it is most important to catch all true positives (Paulus & Kent, 2020). For use cases in which an intervention is extremely high cost or resource intensive, it is very important that every patient offered the intervention be able to benefit from it. In those use cases, predictive parity may be a better

measure. As detailed below in Implementation, the primary use case for this risk score is to flag patients within primary care so that primary care providers can be sure to do a full range of screenings or connect patients with a warm handoff to a social worker or community health worker. These use cases span high and low cost interventions, and thus both fairness metrics were calculated.

The most severe bias observed was by insurance status, with uninsured patients having the lowest PPV (39.13%) and highest FNR (75.68%) of any sociodemographic group. Uninsured patients also had the lowest event rate (2.31% observed) and flag rate (1.4% predicted). It is not surprising that a risk score built on clinical data would under-capture risk among uninsured patients. Being uninsured is a barrier to care. According to the 2018 Community Health Survey, 18% of uninsured New Yorkers didn't get needed medical care, as compared to 9% of people with private insurance (NYC Department of Health and Mental Hygiene, n.d.). New Yorkers from immigrant communities cite insurance as a key barrier to care (Lent & Hickson, 2023; Pavilon & Virgin, 2022). As of June 2025, all children under the age of 18 are still eligible for New York State's CHIP Medicaid coverage, regardless of immigration status (National Academy for State Health Policy, 2019). However, families may not know that this coverage exists or how to enroll in it.

The finding of bias by insurance status also likely reflects the dynamics of recent asylum seeker arrivals in New York City. From 2022-2024, 210,000 people seeking asylum arrived in New York City. NYC Health+Hospitals saw at least 15,000 patients seeking asylum in this period (Newton-Dame et al., 2025). Patients in these families are new to H+H, and did not have a chance to accrue documentation in the EMR before a 2022 arrival. These patients would appear less risky than they actually are based on EMR data. However, as patients become established and potentially receive needed outpatient services, more data is accrued. These patients might indeed have a more accurate EMR-based picture of risk in 2024-2025 than 2022-2023. It will also be important to accurately capture risk for this population, because migration itself confers

risk. Adolescents who have recently arrived in New York City may have special healthcare needs based on gaps in care or trauma accrued during the migration experience (Kroening & Dawson-Hahn, 2019; P. Iqbal et al., 2022; Robertshaw et al., 2017). Meeting the physical, mental and behavioral health needs of these patients will be critical.

One upstream solution is coordinated insurance enrollment at the point of care for uninsured families. Assuming that CHIP eligibility remains unchanged in the state after final adjudication of 2025 federal Medicaid cuts, NYC Health+Hospitals can expand current insurance navigator services to reduce adolescent uninsurance. Because insurance enrollment may take place in the financial office and not in the clinic, it may be helpful to staff volunteer hospital navigators who can walk parents and adolescents to the right office for enrollment. Ideally, these volunteers will match the communities they serve, particularly linguistically. This will allow families to engage in the full spectrum of care to which they are entitled, ultimately yielding more complete data and tailored predictions.

A near-term remedy is also necessary before deploying this algorithm into production. In the 2023-2024 Bias in Algorithms project, the Office of Population Health established an algorithmic bias mitigation playbook detailing how to develop custom thresholds to minimize algorithmic bias within the deploying system or population (Mackin et al., 2025). As the next stage of this work, it will be essential to tailor thresholds to reduce false negative rate among uninsured patients to improve insurance parity.

It will also be important to analyze these biases to identify any underlying system patterns which require targeted quality improvement. Potentially intersectional bias was observed for uninsured, Asian and “other” language speaking patients. Further investigation using more granular race/ethnicity, language and facility sub-groups is warranted to understand this bias. For example, if the bias is particularly observed for uninsured, Chinese speaking adolescents, there may be linguistic competency, healthcare access and workflow issues to identify and address.

Finally, it is important to identify unmeasured bias lurking in the outcome itself (Obermeyer et al., 2021). For example, an outcome of cost may under-capture risk for uninsured patients or patients of color as compared to an outcome of objectively measured mortality, because some patient populations have more access to the care they need (Obermeyer et al., 2019). Future research should measure performance and assess outcome correlation against a more objective measure of global risk such as school absenteeism to understand the impact of unmeasured missing data in this study's outcome. This will require a data use agreement with the NYC Department of Education that demonstrates direct research benefit to students or schools according to FERPA requirements.

Real-World Implications

This study set out with a practical goal: to train an adolescent risk score for primary care that was tailored to NYC Health+Hospitals patients, clinical practice, and real-world data. This section outlines the practical next steps for this work. These include identifying how the risk score could impact H+H system workflows and program planning for adolescent care; providing practical implementation examples from other H+H risk score portfolios; outlining an implementation plan for this score; and providing a logic model of potential inputs, activities, outputs and outcomes.

Implementing the Adolescent Risk Score at H+H

Although risk score implementation is outside the scope of this study, it is important to provide a contextual overview that may inform dissemination, socialization and future validation of this score. At the conclusion of this study, the refinement and implementation phases of work will begin. This section provides a brief overview of practical steps on that implementation pathway.

First, the algorithm should be re-tuned on 2024-2025 data to account for algorithmic drift and changes in data collection patterns, with imputation repeated for missing data. Next, bias should be re-evaluated. As outlined in Addressing Algorithmic Bias, before implementation

bias must be mitigated via an effective strategy such as threshold adjustment. The model should be socialized with key informants, Pediatrics and Adolescent Councils, the Predictive Analytics Governance Workgroup and the AI Governance Committee, with clinical insight driving implementation threshold selection. Potential auto-qualifiers should be reviewed during this process. Pediatrics staff will then work with OPH analysts to design proposed adolescent risk score workflows, which will be refined by facility primary care providers in a listening tour modeled on young child risk.

On the technical side, Office of Population Health (OPH) analysts will automate the collection of model variables and risk score calculation, after bias mitigation and re-tuning are complete. Pediatrics leadership will submit a demand to push the risk score into the Epic EMR. Epic build modifications to support finalized workflows will be requested by Pediatrics leadership and supported with tailored Training, building on lessons learned in the young child risk project. Once live, model drift and bias will be monitored at least quarterly via the Epic Trust and Assurance Suite or a commensurate tool. Bias or drift that exceeds system thresholds will trigger review, leading to redevelopment or deprecation.

Resources and Workflows for High Risk Adolescents at H+H

This section surveys the potential H+H resources and programming that might benefit from an embedded adolescent risk score. These include enhanced screening, specialty care referrals, CHW placement, community referrals, benefits navigation, and program planning.

An adolescent risk score embedded thoughtfully in the EMR should prompt more robust assessment of high risk adolescents. A 20-minute primary care visit precludes universal screening for all patients across all domains. A pilot to place tablets for self-screening into pediatric waiting rooms improved screening rates but gaps remain. Many providers experience alert fatigue, and best practice alerts are in no way a panacea when providers are not given adequate time to address patient needs. Risk-adjusted scheduling which increases well child visit length for high risk patients could facilitate more in depth screening and clinical

assessment. An order set of core screenings and assessments, chosen by adolescent providers, is worth testing to understand uptake and interest. This order set or alert could also prompt providers to consider whether additional specialty referrals or more frequent primary care follow-up were warranted.

An adolescent risk score can also complement clinician judgement to improve specialty scheduling. Providers often refer adolescents to enhanced care within the H+H network, but capacity limitations within specialty and behavioral healthcare can make scheduling challenging. A high risk indicator for adolescent patients could complement clinical assessment to prioritize patients whose referrals are awaiting scheduling, or whose providers have ordered an e-consult.

The risk score could also identify patients who might benefit from connection to a Community Health Worker (CHW). H+H has developed the largest primary care-based CHW program in the nation, with 250 CHWs deployed into adult and pediatric primary care, women's health, and behavioral health services. CHWs help patients and their families with a spectrum of medical and social challenges for a three-month deployment, including medication management, care coordination, connection to community support, benefits navigation, lifestyle support and health coaching. CHWs address what is most important to the patient or family in their own health journey. In many cases, a CHW will assist a client family with upstream challenges that stand in the way of wellbeing and health management, such as assistance in housing applications. The CHWs which have been deployed in pediatric primary care have been primarily focused on the 0-4 age range as part of the Healthy Steps and 3-2-1 Impact models. Although risk scores should never preclude provider decision-making, an adolescent risk score which is featured in pre-visit planning materials could complement clinical judgement to guide CHW enrollment where indicated and desired by the family.

In sites that do not have a CHW program for adolescents, patients flagged as high risk would still be able to benefit from community resources. NYC Health+Hospitals integrated the FindHelp community referral platform into its Epic EMR. With FindHelp patients can be referred

to verified community CBOs, including food pantries, violence prevention programs, housing supports, behavioral health treatment centers and other community organizations which cater to social needs. Through FindHelp, providers can refer a patient to a community resource, who can then send back enrollment and service details for that patient. A high risk flag for adolescent patients would indicate to providers that they should consider leveraging FindHelp to refer the patient to community support, in consultation with social screening results.

H+H has also developed a social medicine program that can benefit higher risk adolescents and their families. H+H has created a medical-legal partnership with NYLAG to provide no-cost legal assistance on-site at most facilities. High risk children exposed to ACEs are more likely to have experienced familial justice involvement or housing insecurity (Folk et al., 2021; Graf et al., 2021). NYLAG assists patients with eviction prevention, immigration status, domestic violence and other prevalent health related legal issues. H+H also offers on-site WIC enrollment for pregnant women and mothers of small children, which may include some pregnant or parenting adolescents that would be flagged as high risk under the model of this inquiry. A high risk alert could prompt practices to connect patients and their families to these social medicine resources with a warm hand-off, to improve engagement and increase trust.

Finally, an adolescent risk score can inform system goals, policy targets and funding strategies. H+H has begun to risk-adjust panel sizes for primary care providers in adult medicine and pediatrics. Risk-adjusting primary care provider panel size allows systems to better match capacity and demand (Shekelle et al., 2019). Currently, pediatric panel adjustment uses PCI data as a crude measure of patient need. Incorporating a risk score tuned specifically to H+H adolescents would more accurately reflect provider experience of care delivery to patients with social, behavioral and medical needs, improving systemwide staffing models and matching capacity to demand. Population and facility-level risk snapshots will also be important to argue for enhanced CHW staffing models Finally, population-level risk snapshots can identify key

drivers of risk within our patient populations, allowing facilities to hone programming or argue for resources with empirical data.

Past Risk Score Implementation at H+H

H+H has already implemented a home-grown risk score into clinical practice. The Office of Population Health trained an adult acute utilization risk score (Li et al., 2023; Ziring et al, 2018) which was incorporated into Epic. The adult risk score is used during outpatient pre-visit planning and flags patients who should be explicitly evaluated for a referral to the CHW program. Schedulers use the risk score to prioritize follow-up outreach after inpatient discharge. The risk score acts as an eligibility criterion, allowing patients to access enhanced care management resources who would otherwise be ineligible. It is also used to risk-adjust primary care provider panels. The adult risk score established system workflows and implementation pathways that adolescent risk can leverage.

The adult risk score implementation also offered some lessons learned. This risk score was largely designed as a top-down implementation. While a limited listening tour was undertaken within primary care to socialize the risk score, providers were not robustly included during implementation planning or risk score roll-out. Feedback from primary care providers that the adult risk score was not capturing the most primary care-amenable sources of risk were received after the score was implemented, and uptake within the primary care setting was uneven. The adult risk score alert had an extremely low response rate (<5%) because providers were not sure what resources to offer these patients, and was ultimately deprecated. The adult risk score implementation also did not include patients in design or implementation. Ultimately, a process of iterative implementation and socialization identified that the adult risk score was particularly valuable for workflows that cross care settings (e.g. inpatient discharge planning, primary care acute follow-up or care management outreach). The risk score is now used regularly, but this discovery process unfolded across four years. Continued and frequent

partnership with patients and providers during workflow design, implementation, training and assessment stages is a critical lesson for the adolescent risk score.

H+H is also actively implementing a young child risk score (0-4). This risk score uses data to approximate the manual risk tiering conducted in the Healthy Steps and 3-2-1 Impact programs, with the goal of risk tiering every young child. There are several relevant innovations within this project that may benefit the adolescent risk score. The proposed implementation of the young child risk score will include a manual override, allowing providers to adjust risk tiers that they feel are inappropriate. The risk score implementation will also provide a score breakdown to front line providers, enhancing explainability and transparency (Economou-Zavlanos et al., 2023). While this is in pilot stage, requiring the OPH team to push data into the Epic Clarity environment, such an approach would allow individual providers to target drivers of risk (or their proxy variables) more effectively within the clinical encounter, tailoring assessment and referrals to patient needs. Developed as lessons learned from the adult implementation, these approaches improve provider buy-in and facilitate human-in-the-loop decision making (Bedoya et al., 2022). They should be replicated for adolescents.

Another innovation from the young child risk score project is the use of “auto-qualifiers” which immediately flag a child as high risk when identified, independent of the risk score. This strategy is modeled on Denver Health, a safety net which augmented its adult risk score with both physician judgement and auto-qualifiers (Johnson et al., 2015). The auto-qualifiers, sourced from clinician leaders within the system, incorporate critical low-prevalence indicators which fell out of the young child LASSO model. These include factors such as maternal mental illness, maternal substance use disorder or child foster care status. Incorporating auto-qualifiers selected by key informants such as early adolescent pregnancy (12-15y) or foster care status would enrich the adolescent risk score and minimize the impact of missing data for patients with sparse records.

The young child risk score project is currently conducting a listening tour with primary care sites to review and revise proposed high risk workflows. While the workflows will be distinct, this socialization primes the pump for adolescent risk score implementation. These networks should be leveraged when H+H is ready to implement, to ensure that proposed workflows are valuable across our sites.

Logic Model: Resources, Outputs and Outcomes

Finally, this study offers a logic model to describe how this risk score might function to improve care and wellbeing for adolescent primary care patients at NYC Health+Hospitals (Table 5.1). After model development is finished, the adolescent risk score will require some routine system resources. In addition to modeling data, these include OPH analyst time to run the model monthly and evaluate model drift and bias quarterly; IT staff time to build Epic workflows, Pediatrics time to train, and provider time to learn workflows. We anticipate these resources would yield a monthly or weekly refresh of an adolescent risk score in Epic, including an editable risk tier in the patient's chart, a workbench report of high risk patients who are overdue for a visit, a patient list of upcoming high risk scheduled visits for pre-visit planning, a high risk order set of suggested screenings and referrals, and a best practice alert to prompt recommended primary care delivery.

We expect certain measurable outcomes due to risk score implementation. In the short term, the percent of high risk patients with completed screenings across sexual health, substance use, mental health,and social needs screenings would increase, as would adolescent CHW enrollments, collaborative care engagements, behavioral health referrals and specialty referrals. Referrals would also increase to internal social medicine supports such as benefits enrollment or legal services and to external community support via the FindHelp platform. No show rate would also decrease among high risk patients as a result of enhanced visit outreach, and well child and follow-up visit rates would increase due to coordinated scheduling.

In the mid-term, CHW engagements with high risk patients and families would increase, and more eligible families would be enrolled in SNAP, WIC or appropriate insurance. More families with legal needs would receive NYLAG consultations, and more families with food insecurity would visit local food pantries identified by FindHelp. Patients with mild to moderate depression, anxiety or ADHD would receive coaching and medication management through Collaborative Care and other integrated behavioral health models, while patients with more severe needs would be seen in Behavioral Health.

In the long term, we would hope that using the high risk score to connect high need patients and families to H+H and community resources would improve health and wellbeing. Among high risk patients, this would translate to improved school attendance and performance; higher self-rated overall health; improved symptom control and management of mental and behavioral health conditions; and increased family financial, food, housing and legal security. This risk score leverages a suite of programming developed by dedicated, hard working staff across H+H and the City, and any success of this program would be directly attributable to their efforts.

We also hope that the provision of an adolescent risk score in primary care will improve targeting of resources as the Medicaid policy landscape shifts. The passage of H.R.1 One Big Beautiful Bill Act (OBBB) in 2025 portends massive Medicaid cuts, with a potential loss of insurance for 1.7 million New Yorkers without additional action at the State or City level. As of publication, CHIP does not carry immigration status eligibility requirements in New York State. However, families who do not meet the 5 year legal residency requirements for Medicaid or who were previously covered under the Essential plan will lose insurance and may not know that their children are still eligible. These insurance eligibility changes will result in budget shortfalls at the State level (projected 13.5 billion) and within safety net hospitals like NYC H+H. Although the OBBB includes 50 billion in gap funding for rural hospitals to buffer these changes, no additional funding has been earmarked for urban hospitals, and Disproportionate Share Hospital

(DSH) payments are a likely future target of additional budgetary cuts. More frequent algorithmic tuning and bias evaluation may be needed to address the likely shifts in healthcare data that these seismic changes portend. Implementation workflows will also have to be updated frequently, as H+H programming is impacted by budgetary cuts. In this environment, using tools like this risk score to support clinician assessment in identifying patients who *most* need our services will be critical.

Limitations

This study includes key limitations. Broadly, these limitations can be categorized under generalizability and outcome definition; analytic approach; and missing data. Using data from a single system to predict a highly customized outcome limited generalizability, as did the use of a type 2b pseudo-external Validation sample. LASSO regression may have reduced predictive power compared to more intensive AI and machine learning methods. Residual leakage between prediction and outcome data may have been present despite minimization efforts. Missing data also presented challenges, including the lack of key variables such as resilience, bullying, or social isolation, and unmeasured missingness within EMR variables themselves. This section outlines these limitations and identifies some potential future research directions to understand their impact.

Generalizability and Outcome Selection

This study's adolescent risk score may not generalize to other healthcare care settings. This score was developed on primary care patients, who have more complete data than patients visiting H+H for specialty and acute care only. Performance must be evaluated before deploying this score outside primary care. Generalizability may also be limited for other geographic areas. This study leveraged data from NYC Health+Hospitals and public datasets collected in NYC. Health information exchange data was not uniformly integrated into the H+H EHR during the study period, and thus utilization and diagnosis data from other institutions were not included in Training data. Findings from this study may have limited generalizability to hospital systems

serving suburban or rural adolescents and middle or high-income patients, whose healthcare seeking behaviors and sources of risk may differ.The risk score should be retrained using data from multiple systems and regions before deployment outside H+H.

The validation strategy employed in this study may also have important implications for generalizability. This study employed a Type 2B pseudo-external validation using a geographic split within NYC Health+Hospitals data. Gold standard validations test risk scores on data from other systems to understand generalizability and performance degradation across settings (Collins et al., 2015, 2024). Without a true external validation, generalizability outside of NYC Health+Hospitals cannot be ascertained.

Another factor impacting generalizability is this study's use of a locally defined outcome of interest. A review of the literature failed to identify a gold standard outcome incorporating medical, social and behavioral risk for adolescent patients. Identified sources provided narrow risk scores for a single outcome such as suicide (King et al., 2019; Su et al., 2020), enumerated a lengthy list of global indicators (The Adolescent Health Indicators Recommended by the Global Action for Measurement of Adolescent Health, n.d.) or introduced comprehensive research instruments which are not feasible to deploy within primary care (Bethell et al., 2021). As a result, this study aimed to craft an outcome of interest specifically tailored to the concerns of NYC Health+Hospitals providers and adolescent patients, via key informant interview. The strength of this approach is a score which is highly tailored to NYC Health+Hospitals' unique urban safety net population. However, using a customized outcome over a validated instrument likely decreased generalizability outside NYC Health+Hospitals or similar safety net healthcare systems. Future opportunities to compare this outcome of interest to validated instruments such as Bethell et al,'s Integrated Child Risk Index would be valuable to establish construct validity and quantify generalizability.

The customized outcome approach offers some additional limitations worth noting. Clinician key informants were identified via snowball recruitment with a system leadership seed.

Patient key informants were recruited from an existing adolescent patient advisory group. These perspectives may not be representative of all Health+Hospitals clinicians or patients. Key informants included primary care providers, system leaders, behavioral health providers and social workers but did not include nurses, an important clinical team member. Nursing leadership did not consent to nurse recruitment during key informant enrollment. Included providers and patients were demographically diverse, but no Hispanic/Latinx participants were included. Incorporating Hispanic/Latinx and nursing participants will be essential for the implementation stage. Another potential limitation of the outcome of interest is that it weighted all adverse outcomes equally. Indices such as the Pediatric Comorbidity Index used in this study assign different weights to different variables to reflect clinical importance. However, reviewed literature did not identify any ranking schema to determine the relative importance of each risk factor (e.g. trauma vs. substance use vs. sexual risk taking) for adolescent health and wellbeing, and the key informant process did not yield objective weights or rankings for adverse outcomes. The adverse outcome score used to train this risk model therefore weights all factors equally, although the resulting risk model provides population-derived weights which can be implemented.

Finally, this study's outcome likely underestimates true adverse experience risk among adolescents. The study defined risk by identifying adolescent outcomes associated with Adverse Childhood Experiences and tailoring those outcomes to the NYC Health+Hospitals population. However, the outcome period for this study is limited to one year, whereas the production of poor health due to toxic stress manifests across a lifetime. Adolescent stressors may not fit neatly into a prediction year, or may be reported to a clinician after they are first experienced. This study may underestimate longitudinal risk among developing adolescents.

Analytic Approach

There are several analytic limitations that merit discussion. These include choice of LASSO regression over AI and machine learning approaches, lack of clinical utility estimation,

and potential data leakage between prediction and outcome years. This section discusses these analytic limitations and offers ideas for future research and development.

This analysis did not use AI or deep machine learning, which may represent a performance limitation. Healthcare systems are increasingly deploying machine learning, neural networks and AI-based applications to tackle prediction problems in healthcare (Alowais et al., 2023). AI systems promise a high return on investment, but implementation requires significant expenditure on software, systems, governance and staff skill development (Bharadwaj et al., 2024). A 2024 survey of 40 major health systems found that while 34 were using AI and predictive analytics for clinical risk stratification, only 12 felt it had been highly successful (Poon et al., 2025). After lack of AI maturity, financial concerns were reported as the biggest barrier to robust healthcare AI implementation.

Pairing lower lift methods to address more straightforward problems may make sense for safety net systems like H+H. This study selected LASSO regression using a logistic link over more sophisticated methods primarily because of the computational limitations of the NYC Health+Hospitals system during development. An advantage of LASSO is that it prioritizes parsimony, making the model easier to run and maintain long-term for the H+H ambulatory care data team. LASSO's relative simplicity may decrease overfitting and thereby improve temporal stability and in-network generalizability as documentation and systems change (Collins et al., 2024; Gunn et al., 2023). However, model performance would likely have been stronger using a more intensive deep learning technique that leverages hundreds of variables and can learn over time. Future research should compare these approaches head-to-head for risk prediction within adolescent primary care to understand how best to identify kids and families needing intervention and support.

The adolescent risk model should not be interpreted causally. LASSO modeling selects at random a single predictor from a set of equally predictive and collinear predictors, without

regard to causal pathway or clinical importance. Variables selected via LASSO for this score should be interpreted as meaningful prediction signals without inferring a causal relationship.

Another analytic limitation is that this study did not cover implementation or measure clinical utility of the model. This study followed TRIPOD+AI reporting requirements (Collins et al., 2024). However, some sources also recommend including measures of clinical utility such as a decision curve analysis or net benefit statistic (Steyerberg & Vergouwe, 2014). These clinical utility measures require a numeric estimate of treatment benefits and risks, which was unavailable as of study publication. As outlined below in Implementation and Impact, patients flagged as high risk will leverage multiple intervention pathways, which are likely to vary in availability and implementation by facility. In the implementation phase of this research, clinical utility assessment for each action pathway should be paired with rapid QI RCT methodology to understand what intervention levers are most appropriate for and beneficial to the high risk adolescents identified by this score (Horwitz & Krelle, 2023). Of note, many implementation-focused evaluations of algorithmic performance also evaluate facility-specific statistics when deciding whether to implement across the whole system or deploy on a site-by-site basis. This was not undertaken within this study but will be an important step in the implementation process.

Another analytic limitation of this study is the potential for leakage within EHR data. Leakage is defined as “the unintended use of known information as unknown” (Luo et al., 2016). Validation leakage is the contamination of the Validation data with Training data. Examples include using the same patients in both Validation and Training datasets or allowing knowledge about the Validation dataset to shape model building in the Training partition. Outcome leakage occurs when information from the prediction period is inadvertently captured in the outcome of interest. This is a common issue within healthcare data, where information from past encounters may be inadvertently copied to future notes. The result of leakage is an overoptimistic model, which performs better within the study data than it does in real life (Luo et al., 2016). Results

presented in this study may suffer from this optimism. Longitudinal performance monitoring before and after implementation of the adolescent risk model should quantify performance optimism.

Validation leakage is likely minimal within this study. This study partitioned the Validation data before any data cleaning or descriptive statistics. Each patient identifier contributed to only one dataset. It is possible that a patient might have been accidentally registered under two patient identifiers at separate primary care visits across two facilities within the study period. However, given the lag between the enrollment window (2022) and data abstraction (2024), most patients with more than one identifier would have already been merged in Epic before the study by the H+H Health Information Management team.

A bigger threat is unintended outcome leakage. The pediatric comorbidity index (PCI) measures chronic conditions per patient, which are likely to be stable from year to year. To minimize outcome leakage, this study did not include problem list diagnosis, which may be carried forward automatically from visit to visit. All diagnosis variables were derived from visit or billing diagnoses added at a specific encounter within the prediction or outcome window. Diagnosis-related variables in this study therefore measure care sought for a given condition, rather than prevalence of a condition itself. Although the full 2023 PCI count was examined for possible inclusion in the outcome score, it was found to be too correlated with the 2022 PCI count. Instead, a categorized PCI variable identifying adolescents in the top 75th and 95th percentiles of adolescent comorbidity was added to the outcome of interest, contributing no more than 2 points to a patients' adverse outcome score. However, it is likely that residual PCI-related leakage is present in this score.

Because of the potential for leakage within two years of adolescent healthcare data, this score is likely flagging risk prevalence rather than new risk incidence. It is likely a diagnostic risk model rather than a true prognostic risk model. Per the TRIPOD statement, diagnostic models "estimate the probability of the presence of a particular health condition," while prognostic

models “estimate…whether a particular outcome will occur in the future” (Collins et al, 2024). This reflects the reality of adolescent development and risk. Risk in adolescence is not a sudden outcome, but rather an accumulation of challenges, successes, life experiences, behaviors and physical conditions. Toxic stress and its impact on metabolic, immunologic and neurological development are cumulative (Danese & McEwen, 2012). A diagnostic risk algorithm is valuable for primary care. A prognostic algorithm which predicts sepsis best among patients who are already septic is fatally flawed. However, a diagnostic algorithm which correctly flags current *or* future high risk adolescents allows clinicians to connect them to services that address adolescent or family needs. In addition to quantifying model optimism, it will be important to collect qualitative provider feedback on the accuracy and usefulness of seeing a diagnostic high risk flag in the primary care encounter.

Missing Data

This study used EMR and geographic data to describe both risk predictors and outcomes. However, true adolescent risk and resilience factors are complex, multi-faceted, and often poorly or incompletely measured in available data. Missing data present important limitations for the adolescent predictive risk score developed in this study.

Individual and family resilience can buffer the effects of toxic stress on childhood health and development development (Bellis et al., 2018; Giovanelli et al., 2020; Hall et al., 2021; Shonkoff, Boyce, et al., 2021; Sosnowski et al., 2021). However, resilience-related variables are not collected routinely in the Epic EMR system. Community-level resilience is also poorly measured in datasets like the Social Vulnerability Index, and zip or census tract level matches of resilience factors may not translate reliably to the individual patient. This study was unable to measure protective factors, outside two poorly collected exercise and social support screening questions in SSHADESS - a significant limitation. AI applications like Ambient Scribe which transcribe and summarize the full clinical encounter may increase our ability to identify patient resilience in the future. This is explored more fully in Further Research.

Some risk factors are also under-captured or not captured at all within this study's EMR data. These include experience of bullying, missing school often, feelings of social isolation and problematic smartphone or social media use. Patients who disclose these kinds of risks to their provider without screening positive for school difficulties in SSHADESS should receive a manual risk tier adjustment, as outlined in Real-World Implications.

Another problem in prediction research is missing data that cannot be measured, or "partially observed" data (Pham et al., 2024). Chronic comorbidity indices such as the PCI that use ICD-10 codes fail to capture undiagnosed disease. Patients may not disclose social needs or substance use because of stigma or fear of repercussions (Earnshaw, 2020; Hare et al., 2023). Patients may miss needed care entirely because of scheduling or access barriers. These patients receive less care than an otherwise equivalent peer, and they will therefore look less risky to an algorithm using EMR data (Obermeyer et al., 2019). Exogenous events may also impact data capture. This study used data from 2022, a transitional year in which healthcare, schools and families were all emerging from the COVID-19 pandemic's unprecedented upheavals and social stressors. This may have impacted what data were recorded during the clinical encounter. Future research should periodically re-evaluate variable inclusion as part of routine model retuning, to ensure that the score uses the best available data to flag adolescent risk.

The rate of loss to outpatient follow-up observed in this study among the Training cohort (26.5%) suggests that unmeasured missingness is a common problem at H+H. Algorithm performance was significantly worse among patients who were lost to outpatient follow-up, and these patients differed significantly in care utilization, demographics and chronic disease burden from patients not lost to follow-up. A predictive algorithm will work poorly for these patients. Workflows to re-engage primary care patients who have fallen out of care should be established in parallel to high risk-oriented workflows, to ensure that adolescents in the lost to follow-up population are not underserved.

Absent upstream investments in comprehensive care, studies can deploy several statistical techniques to address missing data. List-wise deletion, or the omission of patients with missing data, is not recommended for missing data unless all data are missing completely at random - an assumption seldom upheld by healthcare data (Kang, 2013; Pham et al., 2024). Imputation is commonly leveraged to address missingness. MICE and K nearest neighbor imputation strategies preserve data variability and are the gold standard in prediction imputation (Azur et al., 2011; Beretta & Santaniello, 2016; Pham et al., 2024). However, both MICE and K nearest neighbor failed to run within the H+H SAS environment. Hot deck imputation is a method that preserves sample size while maintaining a group mean, but it underestimates sample variance (Andridge & Little, 2010). For categorical variables with high levels of missingness that are not missing at random, modeling a “missing” category is another frequently leveraged approach (Pham et al., 2024). Although it was a weaker method, hot deck imputation was used for lower missingness variables (<30%) because it was able to run within the H+H environment. For variables with >30% missingness in which missing values were likely to be correlated with other patient characteristics, a “missing” category was explicitly modeled.

Two variables with modeled missing data categories were ultimately retained in the model: smoking status, and LGBTQ+ status. In univariate regression, both unknown smoking and unknown LGBTQ+ status were protective. After adjustment, only unknown LGBTQ+ status was protective. It is likely that two things are happening clinically when a patient is missing an LGBTQ+ status. First, that patient may have come to fewer appointments overall or have missed their well child visit, and the provider may not have had time to screen them. These patients may be more likely to be lost to follow-up, in which case their 2023 adverse event score would appear artificially low. Second, some clinics may be differentially launching screening when they suspect that a patient may want to discuss their sexuality or gender identity. This may mean that screening is undertaken more often for non-cis/hetero presenting youth. In turn, non-cis/hetero youth are at increased risk of attempting suicide, elevating risk (Allred, 2024).

Thus, a missing screening may represent poor EHR data capture *and* differential screening patterns based on provider perception of LGBTQ+ status. In 2025, primary care screening rates among adolescents hover around 40% for gender identity and 50% for sexual orientation, an improvement from 2022 but falling short of the H+H 80% screening goal. It will be important to reevaluate this variable as data capture continues to improve.

This study also attempted to enrich missing EHR data with available NYC proxy data. SDOH screening data were augmented with a homelessness and NYCHA public housing address match, creating a composite flag that was retained in the final model. The study also attempted to supplement unmeasured SDOH data such as income with neighborhood-level Social Vulnerability Index (SVI) data, available from the CDC for every NYC census tract. SVI variables fell out of the final model, but this strategy should be investigated for other gaps.

As mentioned in Real-World Implications, “auto-qualifiers” may augment some missing data by using sparse but critical data to automatically mark a child as high risk. Auto-qualifiers are most important for patients who are missing or underutilizing care who are truly high risk but who might not be scored highly by the algorithm due to under-documentation in the record. Auto-qualifiers allow us to flag patients who have *any* data in a high concern area.

Foster care data provide an illustrative example of the potential of auto-qualifiers. In 2024-2025, H+H established automated data exchanges with several City agencies to provide a full census of children visiting H+H who have been placed in foster care. This foster care flag is far more robust than the incomplete SSHADESS screening data included in this study, triggering a tailored z-code in the EMR. For children with a very chaotic home life, needed care may be missed more frequently, and these children may be less likely to be flagged because of their unmeasured missing data, but *more* likely to need services. Automatically marking adolescents with foster care history as high risk may help level the playing field for these patients within the healthcare system. Future research should explore the impact of using select auto-qualifiers such as this new foster care flag on algorithmic bias and performance.

Future Research

Further research will be necessary to ensure that the adolescent risk algorithm produced by this study is effective, fair and trustworthy when implemented. This dissertation is focused on the definition, design and preliminary assessment stages of research (Weitzman Institute and AcademyHealth et al., 2025). Additional research and monitoring will be necessary as the algorithm progresses in the implementation lifecycle (Coalition for Health AI, 2024). In the next phase of work, it will be essential to design and assess real-world implementation, leveraging approaches from both data science and implementation science to collect stakeholder and patient feedback, improve bias performance, assess the model for temporal drift, establish longitudinal monitoring, and evaluate the impact of model-driven workflows on the provision of care and downstream adolescent health outcomes.

Bias, Fairness and Digital Equity

As outlined in Addressing Algorithmic Bias, mitigating observed bias to improve algorithmic fairness is the most pressing next phase of work. Threshold mitigation may improve observed predictive parity and equal opportunity disparities. Mitigation performance will need to meet the established H+H criteria for success before implementation, outlined in the recent Bias in Algorithms project as “1) absolute subgroup EODs <5 percentage points [subsequently adjusted to 10 percentage points by H+H], 2) accuracy reduction <10%, and 3) alert rate change <20%” (Mackin et al., 2025b). Due to the intersectional nature of observed bias, threshold mitigation by insurance may ameliorate observed insurance and language biases.

While implementation will be conditional on bias mitigation, algorithmic fairness is only one aspect of digital equity. The Agency for Healthcare Research and Quality and the National Institute for Minority Health and Health Disparities have identified five key principles to address algorithmic bias, including imbedding health equity in each stage of the life cycle, prioritizing transparency and explainability, engaging patients in algorithm development, naming fairness challenges and trade-offs, and being accountable for fair and equitable outcomes (Chin et al.,

2023). This dissertation has touched on the first four principles, but work remains to be done to ensure that implementation of an adolescent risk algorithm supports equitable, high quality primary care. It will be important to study the impact of deploying this algorithm within clinical practice on equitable delivery of screenings, referrals and follow-up care. It will also be important to deepen patient partnership, and to partner with patients *and* parents to design and study equitable algorithm implementation. As the National Academy of Medicine has flagged, parent and family research partnerships improve outcomes and reduce stigma (Weitzman Institute and AcademyHealth et al., 2025). Although patients were consulted during algorithmic development, expanding patient involvement and incorporating parent perspectives will strengthen implementation and subsequent research. Avoiding patient stigma and protecting adolescent patient privacy are essential for fostering trust in the adolescent clinical relationship.

Ultimately, algorithms are often biased because they use data collected in biased real-world systems, where different groups of patients have differential access to high-quality, comprehensive care. To truly address bias, upstream changes are required to improve access and care within a fragmented US healthcare system. Increasing private and public primary care reimbursement would facilitate universal screening, more frequent healthcare touches, enhanced access for patients who need more intensive management and improved outreach for patients lost to follow-up. Enhanced funding for CHW models could also improve data capture by addressing care coordination and social needs outside the clinical encounter, allowing providers to operate at top of license (Ignoffo et al., 2024). Providing comprehensive care to all patients, regardless of sociodemographic group or income level, will improve unmeasured and measured missing data and enhance algorithmic performance.

Temporal Validation

Another major phase of research for this risk score will be temporal validation to monitor and adjust for algorithmic drift. Input data drift is typical as utilization patterns and data documentation workflows change over time (Sahiner et al., 2023). The prediction period of this

study overlapped with the tail of the Omicron COVID-19 surge, which may have impacted both healthcare seeking behavior and documentation in ambulatory care. Since 2024, H+H has improved its screening and coding for social determinants of health, which may change the prevalence of key study predictors. The ongoing Medicaid 1115 Waiver in New York State may also impact future collection of SDOH data within the medical encounter as screening expands for Medicaid patients within regional Social Care Networks (New York State Department of Health, n.d.; Office of Health Insurance Programs, 2021). As these two examples illustrate, there are many sources of algorithmic drift. In the next phase of work, it will be essential to re-tune the algorithm using the most complete two years of data available, and these results should be transparently published. Open-source code repositories like Aequitas, AIF-360 or Epic's Trust and Assurance Suite (currently being implemented at H+H) offer packages to facilitate longitudinal performance measurement. These data can be leveraged for research while meeting regulatory requirements.

Implementation and Patient Outcomes

Ultimately, this algorithm will only be successful if it improves the provision of healthcare services within adolescent primary care. An implementation-focused body of research will be important to establish the value of this risk prediction approach. The NYU Rapid Randomized Controlled Trial Lab has established a playbook for rapid RCTs for quality improvement (Horwitz & Krelle, 2023). Rapid RCTs can refine implementation by quickly adjusting messaging, Training and delivery, improving uptake and acceptability. As outlined in the Duke framework, garnering provider input at each stage is essential for algorithmic buy-in (Bedoya et al., 2022). H+H is currently implementing a tailored version of Duke's ABCDs algorithm implementation model, which prioritizes real-world effectiveness and efficacy measurement. Pairing quantitative measures of uptake such as fire rate, alert action rate and order set open rate with qualitative provider acceptability data will be important to understand implementation success.

In an ideal world, the risk score would be rolled out using prospective randomization at a provider or clinic level to understand the impact of the score on patient care. However, this kind of randomization can be challenging from a logistical and technical perspective in a safety net with limited IT resources and provider time. Pseudo-randomization compares service provision and outcomes for patients just above and below a risk threshold to understand the impact of a risk score on care (Apathy, 2025). The next phase of study can leverage pseudo-randomization to track algorithmic impact across established system metrics like no-show rate, well child visit rate, referral rate for specialty and behavioral health care, completion rate for referred visits, CHW referral and engagement rate, and FindHelp referral rate. As outlined in Implementation, we hope that a risk score will trigger providers to screen high risk patients thoroughly. This can be evaluated using SSHADESS screenings such as PHQ-9, GAD-7, SDOH, tobacco use, alcohol use and substance use, as well as lab-based screenings such as chlamydia and gonorrhea.

As identified in Outcome Definition, bounding generalized adverse outcomes among adolescents can be challenging. Some outcomes which are currently collected by the system which could be studied include depression severity as measured by PHQ-9, anxiety symptoms as measured by GAD-7 score, rate of ED visits for asthma or mental health needs, and BMI trajectory. While these measures are useful barometers of adolescent wellbeing, the next phase of research should consult with clinicians, patients and families to identify the most important adolescent outcomes which align with system quality goals and best reflect patient wellbeing.

Evolution of the Field

Finally, as the healthcare ecosystem evolves, it will be important to investigate how predictive algorithms like this one complement emerging AI. Standards are evolving quickly. This study reported model development and results using the TRIPOD+AI guidelines (Collins et al., 2024). However, collaboratives such as the Coalition for Health AI (CHAI) are currently working to design new certifications and standards for a rapidly evolving landscape (Coalition

for Health AI, 2024). Regulatory requirements are also evolving (Mackin et al., 2025a). It will be essential to evaluate the adolescent risk score using these more complex frameworks as they are finalized.

It will also be important to evaluate the ability of new AI tools to enrich adolescent risk prediction as they become widely available. AI and machine learning are revolutionizing healthcare across cancer detection, laboratory interpretation, healthcare staffing, genomic discovery and many more domains (Alowais et al., 2023). These innovations may enrich risk domains in this risk score. For example, large language models are increasingly being leveraged to identify social determinants of health documented within clinical notes, with lower cost and higher sensitivity than previous natural language processing methods (Mazurenko et al., 2025; Vest, 2025). As H+H implements more large language model functionality within its firewall and its EMR, it will be important to evaluate the potential of AI to enrich missing screening data on topics like sexual orientation, substance use and sexual health and to establish new case definitions for data we don't systematically collect such as bullying and smartphone use. Frequent landscape surveys will enable H+H to compare the adolescent risk score to emerging AI solutions, weighing performance gains against implementation costs.

Conclusion

NYC Health+Hospitals provides high-quality care across 11 hospitals and 6 ambulatory care centers to help all New Yorkers live their healthiest lives. Adolescents are an important patient population, and 40 to 50,000 adolescents seek primary care at H+H each year. The COVID-19 pandemic exacerbated existing trends in unmet mental and behavioral health needs among teens (Abrams, 2023). However, at H+H, provider time to surface and address those needs has remained static or decreased as the system attempts to meet increased patient demand (Carbajal, 2024). Access limitations extend beyond the primary care encounter to enhanced programming. While H+H's social medicine programming is robust, it is unable to service all adolescents in primary care. Risk models can help systems like NYC

Health+Hospitals connect high risk patients to services, resources and comprehensive care without adding new screenings.

Primary care providers are tasked with caring for the whole adolescent. This means supporting resilience and addressing a spectrum of risk during this crucial window of brain development, identity formation and social exploration. Primary care providers must identify mental health conditions like depression or anxiety, manage chronic diseases like asthma and diabetes, identify signs of more serious issues like substance use disorder or physical abuse, and serve as a trusted adult as adolescents gain independence and make choices about their bodies. While there are a wide array of generalized risk scores for adults, predicting outcomes such as medical cost, cardiovascular morbidity or acute utilization, pediatric risk scores are often trained to a single specific outcome. A literature review failed to surface a global, validated measure of adolescent risk which could be deployed in primary care. This represents a gap in both literature and practice.

This study sought to address that gap by training an adolescent risk score which used available EHR and public data to predict risk in 2023, as measured by a customized adverse outcome score. H+H patients and providers were asked to craft a definition of risk at H+H from a menu of health outcomes associated with ACE exposure in the literature. The resulting 2023 Adverse Outcome score covered major domains of mental, behavioral, social and physical risk identified by IOM and WHO. A high risk target of the top 5% was chosen to align programmatically with adult and young child high risk score implementation.

The final risk model retained 21 predictors across 23 parameters which covered demographics (3), social determinants (1), utilization (4) and behavioral and physical health (15). This model showed strong Validation performance, as measured by AUROC, calibration slope and calibration intercept. At a top 5% threshold, the model correctly identified 46% of true positives, and more than half of patients flagged by this model would go on to be in the top 5% of high risk adolescents. This study's Type 2B validation strategy used internal data, which limits

the generalizability of these findings to other healthcare systems. However, observed performance surpasses existing programmatic thresholds at H+H, suggesting implementation potential. Significant sociodemographic variation was observed for model performance, and this was particularly marked for insurance. Bias mitigation via threshold adjustment or a commensurate method will be essential prior to implementation.

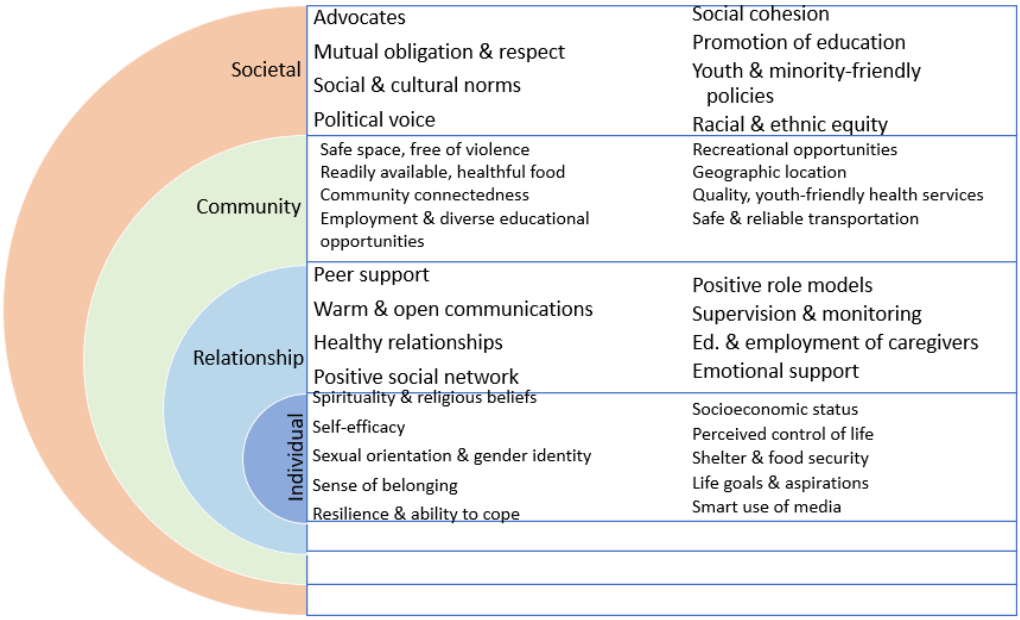
Although this dissertation represents significant work, the true potential of this algorithm will be revealed during implementation. H+H has invested in social medicine programming such as embedded CHWs and social workers in primary care, on-site insurance and benefits enrollment, embedded legal services and community referrals. This risk score has tremendous potential to streamline referral and enrollment for these services, within and outside of the primary care encounter. The adolescent risk score also has the potential to improve screening and referrals among high risk adolescents, so that NYC Health+Hospitals is better able to identify and meet patient needs. These clinical pathways, if implemented, have the potential to improve care delivery and adolescent health outcomes.

This dissertation represents the first phase of work to improve adolescent care targeting. The next phase of research must evaluate the success of implementation, providing real-time data on what works and what doesn't. It must document the success or failure of bias mitigation, to improve equity in practice. It must also identify how robust the score is to changes in data documentation, patient populations, insurance and immigration policy, and AI innovations over time, to ensure that the score continuously reflects reality on the ground. This study successfully trained an EHR-based algorithm to identify adolescents in primary care at high risk of multiple adverse outcomes, within a large urban safety net system. Thoughtful, data-driven implementation in close partnership with patients and providers will be essential to translate the performance observed in this study into improved care delivery for New York City.

Tables and Figures

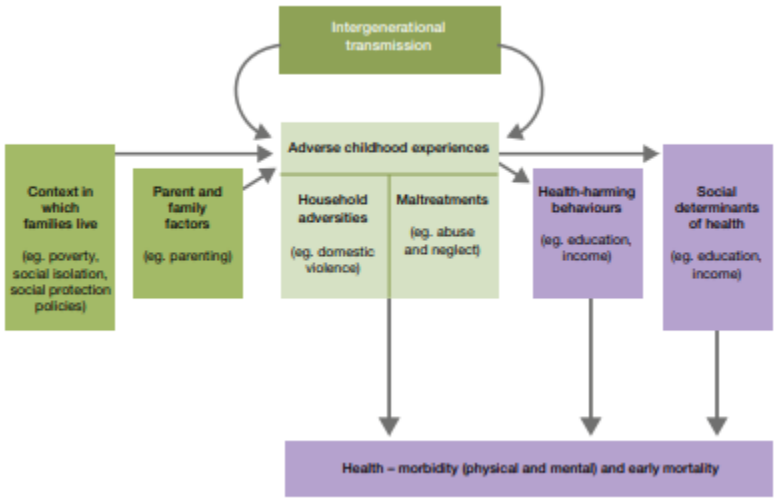
Introduction

Figure 1.1: Theoretical framework: Social-Ecological Model of Adolescent Development



Source: Youth 360°: How & Where Youth Live, Learn, & Play Matters (Presentation). (n.d.). Healthy Teen Network. Retrieved November 5, 2022, from https://www.healthyteennetwork.org/resources/youth360-presentation/

Figure 1.2: Impact of Adverse Experiences in the Home on Children and Young People



Source: Allen, Matilda & Donkin, Angela. (2015). The Impact of Adverse Experiences in the Home on Children and Young People (9; pp. 1–31). Institute for Health Equity. https://www.instituteofhealthequity.org/resources-reports/the-impact-of-adverse-experiences-in-the-home-on-children-and-young-people

Methods

Table 3.1: 2023 Final Outcome Score Components

Domain	Item
Abuse, Trauma, Neglect	Abuse or Neglect (Dx or Screen)
Abuse, Trauma, Neglect	Trauma Dx
Abuse, Trauma, Neglect	Enuresis/Encopresis Dx
Abuse, Trauma, Neglect	Psychosocial Difficulties Dx
Mental Health	Depression (Dx or Phq)
Mental Health	Anxiety (Dx or Gad-7)
Mental Health	Suicidality (Cssrs, Phq, Dx)
Mental Health	Serious Mental Illness Dx Flag (Ocd, Specified Personality Disorders, Conduct Disorder, Psychosis Pci Categories, or Witness or Perpetration Of Violence (Ccsr))
Alcohol & Substance Use	Alcohol Use (Dx or Screen)
Alcohol & Substance Use	Substance Use (Screener or Dx, Incl. Marijuana)
Alcohol & Substance Use	Overdose Dx
Alcohol & Substance Use	Tobacco Use (Smoking, Dip or Vaping)
Healthy Eating	Bmi Change 2022-2023 (Top/Bottom 5%)
Healthy Eating	Eating Disorder Dx
Healthy Eating	Ed Visit: Weight Loss
Mental Health	2+ Somatic Symptoms
Social Determinants	Sdoh Need (Dx or Screen)
Social Determinants	Lifestyle Dx
School Challenges	Problems In School (Dx or Screen)
School Challenges	Adhd (Dx)
School Challenges	Developmental Challenges Flag (Dx Of Developmental Delay, Chromosomal Abnormalities (Including Downs Syndrome) And Pervasive Developmental Disorders (Including Autism Spectrum Disorder)
Sexual Health	Pregnancy, Abortion, Delivery or Miscarriage
Sexual Health	1+ Positive Sti
Chronic Disease & Complexity	Count Of Chronic Conditions Per The Pci, Grouped As 0-2, 3-4 And 5+. Cutpoints Sourced From 75th, 95th Percentiles In Training Population.
Chronic Disease & Complexity	Sickle Cell Dx
Chronic Disease & Complexity	Pediatric Comorbidity Index (Pci) Categories For Cardiovascular Disease, Gi Chronic Conditions, Diabetes, Epilepsy
Concerning Utilization	3+ Ed Visits

Concerning Utilization	1+ Ip Psych Visits
Concerning Utilization	2+ Ip Visits
Concerning Utilization	5+ Op Visits
Concerning Utilization	No Well Child Visit
Concerning Utilization	Ed Visit: Diabetes
Uncontrolled Chronic Disease	Ed Visit: Cardiovascular Disease
Uncontrolled Chronic Disease	Ed Visit: Epilepsy
Uncontrolled Chronic Disease	Ed Visit: Chronic Gi Illnesses
Uncontrolled Chronic Disease	Ed Visit: Asthma
Unmet Bh Needs	Ed Visit: Anxiety
Unmet Bh Needs	Ed Visit: Depression
Unmet Bh Needs	Ed Visit: Conduct Disorder
Unmet Bh Needs	Ed Visit: Ocd
Unmet Bh Needs	Ed Visit: Psychosis
Unmet Bh Needs	Ed Visit: Spec. Personality Disorders
Unmet Developmental Needs	Ed Visit: Chromosomal Malform.
Unmet Developmental Needs	Ed Visit: Congenital Malformation
Unmet Developmental Needs	Ed Visit: Dev Delay
Unmet Developmental Needs	Ed Visit: Pervasive Dev Disorders
Unmet Bh Needs	Ed Visit: Adhd

Results

Table 4.1: Demographic, Clinical and Utilization Characteristics of Training and Validation Samples (2022)

Characteristic	Training (N=26362)	Validation (N=10764)	Total (N=37126)	P-value
Age, n (%)				<.0001[1]
12	4508 (17.1%)	2058 (19.1%)	6566 (17.7%)	
13	4437 (16.8%)	1829 (17.0%)	6266 (16.9%)	
14	4396 (16.7%)	1837 (17.1%)	6233 (16.8%)	
15	4355 (16.5%)	1702 (15.8%)	6057 (16.3%)	
16	4337 (16.5%)	1722 (16.0%)	6059 (16.3%)	
17	4329 (16.4%)	1616 (15.0%)	5945 (16.0%)	
Sex at Birth, n (%)				0.7456[1]
Female	13208 (50.1%)	5413 (50.3%)	18621 (50.2%)	
Male	13154 (49.9%)	5351 (49.7%)	18505 (49.8%)	

Race/Ethnicity, n (%)				<.0001[1]
Asian	1412 (5.4%)	466 (4.3%)	1878 (5.1%)	
Black	9541 (36.2%)	2944 (27.4%)	12485 (33.6%)	
Latinx	11889 (45.1%)	6199 (57.6%)	18088 (48.7%)	
Other	3123 (11.8%)	857 (8.0%)	3980 (10.7%)	
White	397 (1.5%)	298 (2.8%)	695 (1.9%)	
Preferred Language, n (%)				<.0001[1]
English	16161 (61.3%)	6378 (59.3%)	22539 (60.7%)	
Other	1164 (4.4%)	404 (3.8%)	1568 (4.2%)	
Spanish	9037 (34.3%)	3982 (37.0%)	13019 (35.1%)	
Insurance, n (%)				<.0001[1]
Commercial	5252 (19.9%)	1743 (16.2%)	6995 (18.8%)	
Medicaid	19508 (74.0%)	8754 (81.3%)	28262 (76.1%)	
Uninsured	1603 (6.1%)	267 (2.5%)	1869 (5.0%)	
LGBTQ+ Status, n (%)				<.0001[1]
LGBTQ+	1221 (4.6%)	512 (4.8%)	1733 (4.7%)	
Missing Data	14736 (55.9%)	5691 (52.9%)	20427 (55.0%)	
Not LGBTQ+	10405 (39.5%)	4561 (42.4%)	14966 (40.3%)	
1+ Social Needs, n (%)				<.0001[1]
1+ SDOH Needs	1011 (3.8%)	1130 (10.5%)	2141 (5.8%)	
No SDOH Needs Recorded	25351 (96.2%)	9634 (89.5%)	34985 (94.2%)	
Homelessness, n (%)				<.0001[1]
Experienced Homelessness (Last 12m)	560 (2.1%)	374 (3.5%)	934 (2.5%)	
Not Homeless	25802 (97.9%)	10390 (96.5%)	36192 (97.5%)	
Public Housing, n (%)				0.0006[1]
Lives in Public Housing	2716 (10.3%)	983 (9.1%)	3699 (10.0%)	
Not in Public Housing	23646 (89.7%)	9781 (90.9%)	33427 (90.0%)	
Census Tract Poverty Level, n (%)				<.0001[1]
High Poverty (>=30%)	13492 (51.2%)	8470 (78.7%)	21962 (59.2%)	
Medium to Low Poverty (0-29%)	12870 (48.8%)	2294 (21.3%)	15164 (40.8%)	
Primary Care Visits				<.0001[2]
Mean (SD)	1.9 (1.57)	1.6 (1.03)	1.8 (1.44)	
Median	1.0	1.0	1.0	
Range	1.0, 145.0	1.0, 17.0	1.0, 145.0	
Specialty Visits				<.0001[2]
Mean (SD)	1.3 (4.01)	1.3 (3.62)	1.3 (3.90)	
Median	0.0	0.0	0.0	
Range	0.0, 180.0	0.0, 64.0	0.0, 180.0	
ED Visits				<.0001[2]
Mean (SD)	0.3 (0.83)	0.4 (0.93)	0.4 (0.86)	
Median	0.0	0.0	0.0	
Range	0.0, 22.0	0.0, 28.0	0.0, 28.0	
Inpatient Visits				0.6890[2]
Mean (SD)	0.0 (0.19)	0.0 (0.20)	0.0 (0.19)	
Median	0.0	0.0	0.0	
Range	0.0, 10.0	0.0, 10.0	0.0, 10.0	
Chronic Comorbidities (PCI)				<.0001[2]
Mean (SD)	1.7 (1.57)	2.0 (1.71)	1.8 (1.62)	
Median	1.0	2.0	1.0	

Range	0.0, 15.0	0.0, 13.0	0.0, 15.0	
2+ Somatic Conditions				0.1110[1]
No	26189 (99.3%)	10677 (99.2%)	36866 (99.3%)	
Yes	173 (0.7%)	87 (0.8%)	260 (0.7%)	
Serious Mental Illness Dx				0.0409[1]
No	24949 (94.6%)	10243 (95.2%)	35192 (94.8%)	
Yes	1413 (5.4%)	521 (4.8%)	1934 (5.2%)	
Depression or Anxiety Dx				0.8049[1]
No	21891 (83.0%)	8927 (82.9%)	30818 (83.0%)	
Yes	4471 (17.0%)	1837 (17.1%)	6308 (17.0%)	
Suicidality				0.0605[1]
No	25194 (95.6%)	10334 (96.0%)	35528 (95.7%)	
Yes	1168 (4.4%)	430 (4.0%)	1598 (4.3%)	
Pregnancy or STI				<.0001[1]
No	25897 (98.2%)	10483 (97.4%)	36380 (98.0%)	
Yes	465 (1.8%)	281 (2.6%)	746 (2.0%)	
Alcohol or Substance Use				<.0001[1]
No	25361 (96.2%)	8479 (78.8%)	33840 (91.1%)	
Yes	1001 (3.8%)	2285 (21.2%)	3286 (8.9%)	
Tobacco Use				<.0001[1]
Current or Former	605 (2.3%)	296 (2.7%)	901 (2.4%)	
Never	12693 (48.1%)	7214 (67.0%)	19907 (53.6%)	
Unknown	13064 (49.6%)	3254 (30.2%)	16318 (44.0%)	
Abuse or Trauma				0.1294[1]
No	24575 (93.2%)	9987 (92.8%)	34562 (93.1%)	
Yes	1787 (6.8%)	777 (7.2%)	2564 (6.9%)	
Developmental Delay-Related Disorder				0.2698[1]
No	23965 (90.9%)	9746 (90.5%)	33711 (90.8%)	
Yes	2397 (9.1%)	1018 (9.5%)	3415 (9.2%)	
School Challenges				0.0083[1]
No	25302 (96.0%)	10266 (95.4%)	35568 (95.8%)	
Yes	1060 (4.0%)	498 (4.6%)	1558 (4.2%)	
ADHD Dx				0.3106[1]
No	24929 (94.6%)	10207 (94.8%)	35136 (94.6%)	
Yes	1433 (5.4%)	557 (5.2%)	1990 (5.4%)	

[1]Chi-Square p-value; [2]Kruskal-Wallis p-value;

Table 4.2: Distribution of Adverse Outcomes among Training and Validation Samples (2023)

Characteristic	Train (N=26362)	Validation (N=10764)	Total (N=37126)	P-value
2023 Adverse Outcome Score				<.0001[1]
Mean (SD)	1.8 (2.05)	2.2 (2.17)	1.9 (2.09)	
Median	1.0	1.0	1.0	
Range	0.0, 25.0	0.0, 25.0	0.0, 25.0	
Outcome Score Distribution, n (%)				<.0001[2]
0	5370 (20.4%)	1640 (15.2%)	7010 (18.9%)	
1	10738 (40.7%)	3810 (35.4%)	14548 (39.2%)	
2	4568 (17.3%)	2017 (18.7%)	6585 (17.7%)	
3	2209 (8.4%)	1278 (11.9%)	3487 (9.4%)	
4	1257 (4.8%)	745 (6.9%)	2002 (5.4%)	
5	758 (2.9%)	500 (4.6%)	1258 (3.4%)	
6+ (true high risk)	1462 (5.5%)	774 (7.2%)	2236 (6.0%)	

[1]Kruskal-Wallis p-value; [2]Chi-Square p-value;

Figure 4.1: Distribution of Adverse Outcomes among Training and Validation Samples (2023)

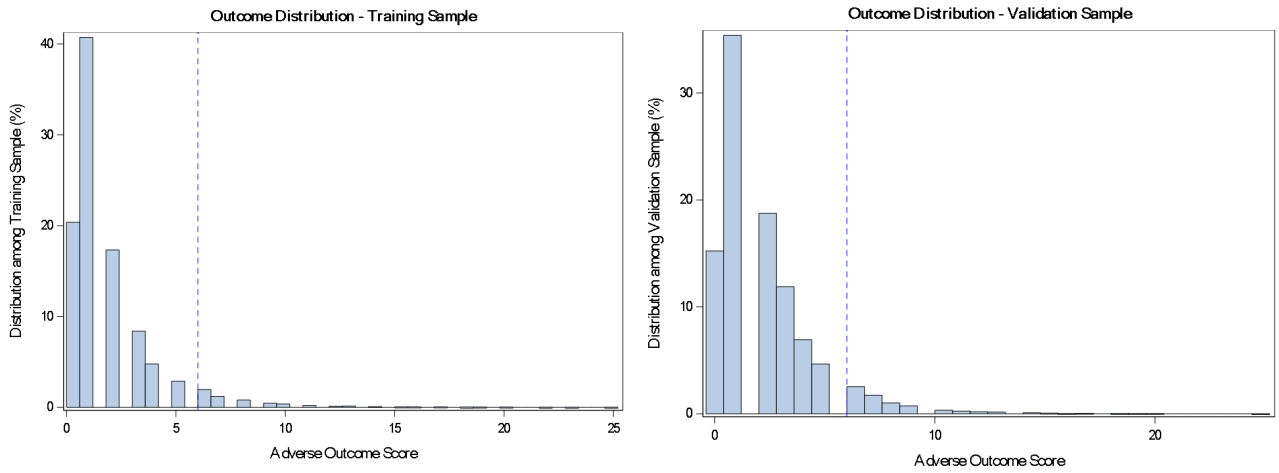


Table 4.3: Preliminary LASSO Model Selection Details

Initial LASSO Model Fit Statistics at Minimized SBC Model

Criterion	Preliminary Linear LASSO	Preliminary Negative Binomial LASSO	Preliminary Logistic LASSO
Lambda	0.0115	0.0074	0.0074
SBC	21807	89909	7720
# LASSO Steps	62	22	22
AIC	47,722	89549	7376

*SAS includes n-1 continuous *and* categorical variable levels when tabulating effects in model for lasso selection.

Figure 4.2: Linear LASSO Model Selection Plot (SBC vs. Model Step)

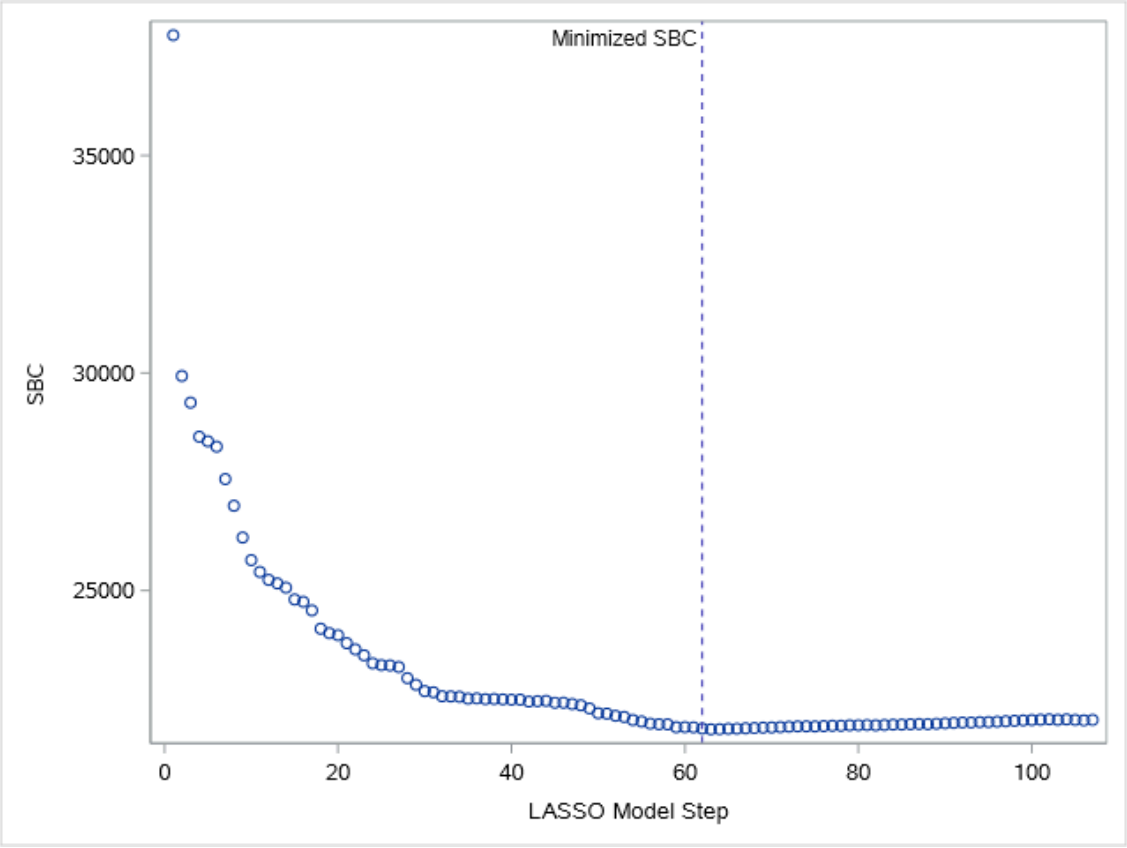


Figure 4.3: Negative Binomial LASSO Model Selection Plot (SBC vs. Model Step)

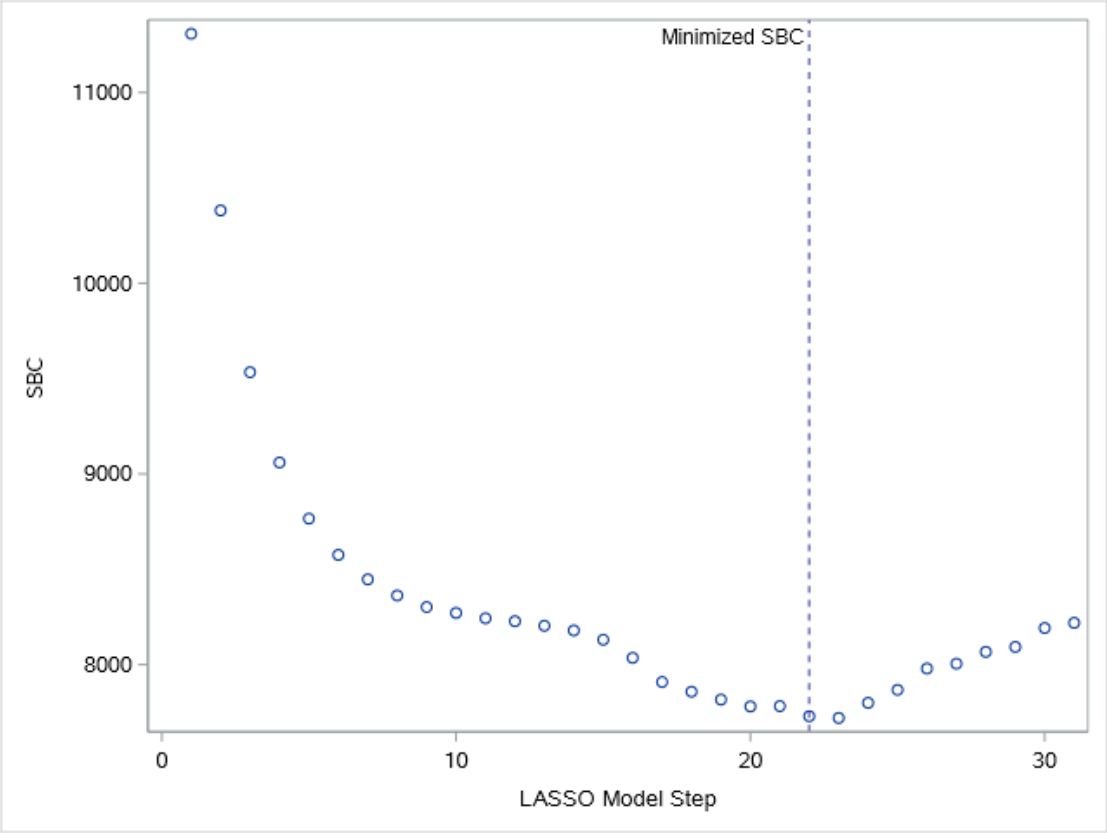


Figure 4.4: Logistic LASSO Model Selection Plot (SBC vs. Model Step)

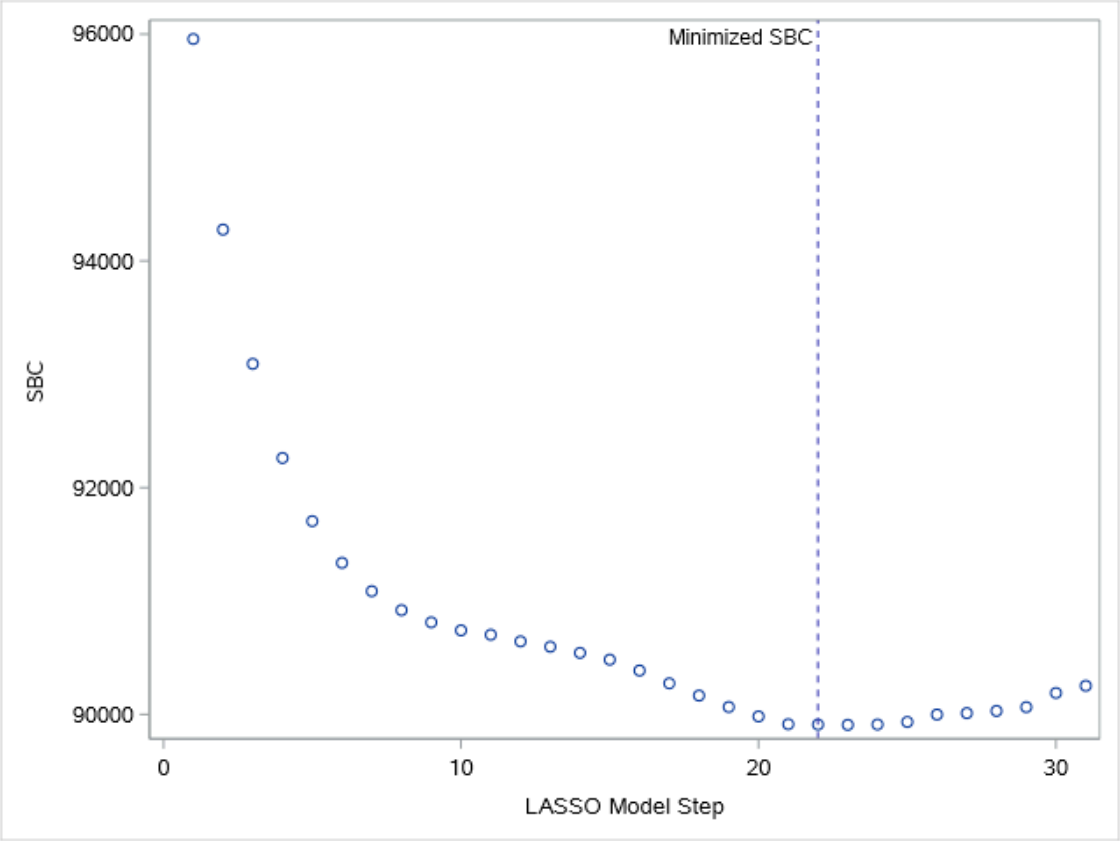


Table 4.4: Preliminary Model Parameters for Linear, Negative Binomial and Logistic LASSO Models

	Linear			Negative Binomial			Logistic		
Parameter	**Est.**	**Stand. Error**	**Wald P Value**	**Est.**	**Stand. Error**	**Wald P Value**	**Est.**	**Stand. Error**	**Wald P Value**
Intercept	0.2805	0.0937	0.0028	-0.3133	0.053	<.0001	-1.9694	0.3412	<.0001
Abuse/Foster Psychosocial				0.1596	0.0143	<.0001	0.2061	0.0383	<.0001
Abuse/Foster	0.236	0.0687	0.0006						
Age	0.035	0.0059	<.0001	0.0246	0.0033	<.0001	-0.0085	0.0211	0.6886
Alcoholdrug				0.2361	0.024	<.0001	0.3802	0.0636	<.0001
Alcohol	0.4926	0.1087	<.0001						
Anxiety	0.4818	0.0391	<.0001	0.1222	0.017	<.0001	0.2229	0.0428	<.0001
Smidx				0.1413	0.0187	<.0001	0.225	0.0464	<.0001
Bhvisitflag	0.6413	0.0636	<.0001				0.1386	0.0574	0.0159
Depression	0.4665	0.0346	<.0001	0.2224	0.0152	<.0001	0.2842	0.0423	<.0001
Devdelayflag				0.282	0.0166	<.0001	0.3825	0.0455	<.0001
Druguse_all	0.6308	0.0574	<.0001						
Dx_adhd_pci	0.9382	0.046	<.0001	0.2877	0.0188	<.0001	0.4844	0.0471	<.0001
Dx_asthma_pci	0.212	0.0286	<.0001	0.1239	0.0137	<.0001	0.2104	0.0385	<.0001
Dx_chromosomal_pci	0.97	0.1342	<.0001						
Dx_conduct_pci	0.7837	0.0975	<.0001						

	Linear			Negative Binomial			Logistic		
Dx_devdelay_pci	0.4115	0.0402	<.0001						
Dx_diabetes_pci	0.5074	0.1037	<.0001						
Dx_eatingdisorder_pc	0.6443	0.1184	<.0001						
Dx_eneuresis	0.4639	0.0781	<.0001						
Dx_epilepsy_pci	0.2846	0.0836	0.0007						
Dx_injuries_pci	-0.0993	0.0335	0.003						
Dx_lifestyle	0.4122	0.0813	<.0001						
Dx_obesity_pci	-0.2668	0.0225	<.0001	-0.0931	0.0116	<.0001	-0.0539	0.0356	0.1297
Dx_ocd_pci	0.9313	0.1505	<.0001						
Dx_perdevdis_pci	0.8799	0.061	<.0001						
Dx_psychosocial	0.284	0.0406	<.0001						
Dx_psychotic_pci	0.7672	0.1427	<.0001						
Dx_sicklecell	0.8669	0.1669	<.0001						
Dx_sleep_pci	0.2388	0.0586	<.0001						
Dx_trauma_ptsd	0.0459	0.0814	0.573						
Ed	0.2819	0.0127	<.0001						
Ed_cat				0.0902	0.0059	<.0001	0.2466	0.0311	<.0001

	Linear			Negative Binomial			Logistic		
Insurance (Commercial Vs. Medicaid)	0.0035	0.0236	0.8816						
Insurance (Uninsured Vs. Medicaid)	0.1206	0.0403	0.0028						
Ip	0.1501	0.053	0.0046						
Lgbtq (Lgbtq+ Vs. Not Lgbtq+)	0.3084	0.0466	<.0001	0.105	0.021	<.0001	0.339	0.0708	<.0001
Lgbtq (Missing Vs. Not Lgbtq+)	-0.226	0.0202	<.0001	-0.1384	0.0113	<.0001	-0.5542	0.055	<.0001
Log_ep_limeng				0.0006	0.0036	0.8703	-0.0062	0.0225	0.7846
Log_ep_sngpnt				-0.0038	0.0029	0.1861			
Opvisitcat				0.0324	0.0021	<.0001	0.1088	0.0119	<.0001
Opvisitflag	0.8067	0.0713	<.0001						
Pci	0.3808	0.0117	<.0001	0.163	0.0046	<.0001	0.3254	0.0253	<.0001
Pcvisitcat	0.0552	0.0076	<.0001	0.0251	0.0036	<.0001	0.0895	0.0194	<.0001
Race (AAPI Vs. Hisp/Latinx)	-0.1271	0.0425	0.0028	-0.0731	0.0245	0.0029			
Race (Black/AA Vs. Hisp/Latinx)	-0.1379	0.0216	<.0001	-0.0796	0.0127	<.0001			
Race (Other Vs. Hisp/Latinx)	-0.1086	0.0304	0.0004	-0.0758	0.0173	<.0001			
Race (White Vs. Hisp/Latinx)	-0.0762	0.0769	0.3215	-0.0233	0.0439	0.5965			
Risky_sex	0.5067	0.073	<.0001						
Schoolprob	0.1188	0.0524	0.0233						
Sdohall	0.0795	0.0293	0.0067	0.1041	0.0131	<.0001	0.186	0.0376	<.0001

	Linear			Negative Binomial			Logistic		
Sdohneed	0.379	0.0556	<.0001						
Sex (Female Vs. Male)	0.1074	0.0194	<.0001	0.067	0.0109	<.0001	0.2297	0.0363	<.0001
Smoking (Cur/Form. Vs Never)	0.5913	0.0679	<.0001	0.1919	0.0293	<.0001	0.2992	0.101	0.003
Smoking (Missing Vs Never)	0.0696	0.0201	0.0005	0.0299	0.0113	0.008	-0.0697	0.0645	0.2801
Somatic	0.4577	0.1211	0.0002						
Suicidality	0.3614	0.0509	<.0001				0.1512	0.0522	0.0038
Wcv_flag	-0.1394	0.0226	<.0001	-0.0861	0.0124	<.0001			
Dispersion				0.0914	0.0043				

Note: Model parameters selected with LASSO using minimum SBC criteria in SAS, manually adjusted where applicable as outlined in Methods, and re-run as a generalized linear model to generate parameter estimates, confidence intervals, fit plots and statistics.

Table 4.5: Preliminary LASSO Model Parameter Comparisons

Variable Selection Overview

Variable Selected by:	Exact Variable
1 Model	23
2 Models	11
3 Models	12
Any Model	51

Table 4.6: Linear LASSO Model Fit and Performance Statistics (Training)

Statistic	Value
Root MSE	1.4996
R-Square	0.4630
Adj R-Sq	0.4630
Spearman Correlation Coefficient (p-value)	0.5860 (<.0001)
Mean Predicted Outcome (SD)	1.7884 (1.3925)
Mean Observed Outcome (SD)	1.7884 (2.0464)
Mean Residual (SD)	0.00 (1.4996)
Residual Skewness	2.3794
Residual Kurtosis	13.9964

Figure 4.5: Linear Expected vs. Observed Outcomes, Preliminary Model

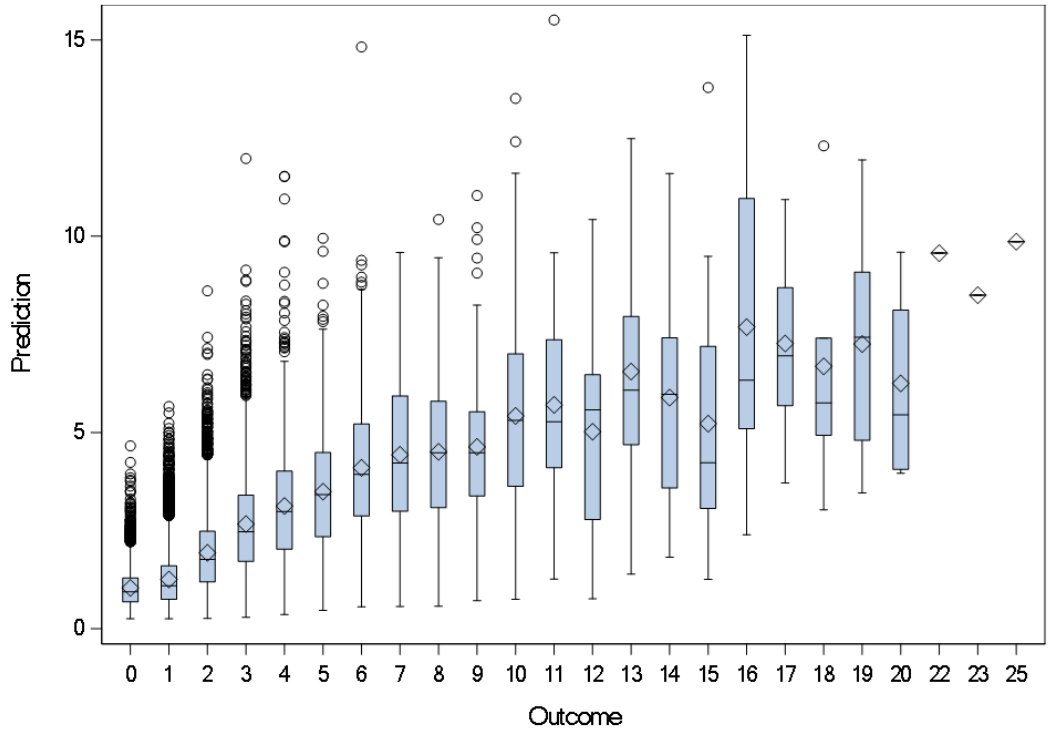


Figure 4.6: Residual Distribution by Predicted Outcome, Linear LASSO Model

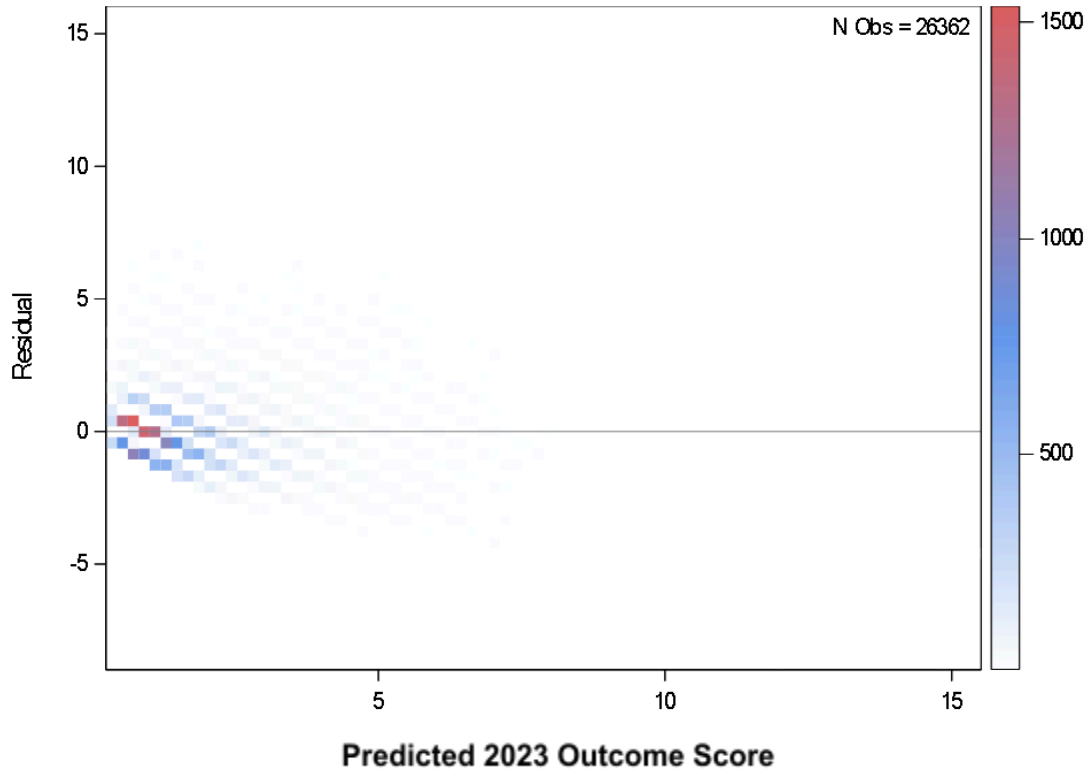


Table 4.7: Negative Binomial LASSO Model Fit and Performance Statistics (Training)

Statistic	Value
Deviance Value (Value/DF)	24649 (0.9360)
Root MSE	1.7053
R-Square	0.1029
Adj R-Sq	0.1028
Spearman Correlation Coefficient (p-value)	0.5812 <.0001
Mean Predicted Outcome (SD)	1.8136 (1.8004)
Mean Observed Outcome (SD)	1.7884 (2.0464)
Dispersion Parameter (SE)	0.0914 (0.0043)

Figure 4.7: Negative Binomial Expected vs. Observed Outcomes, Preliminary Model

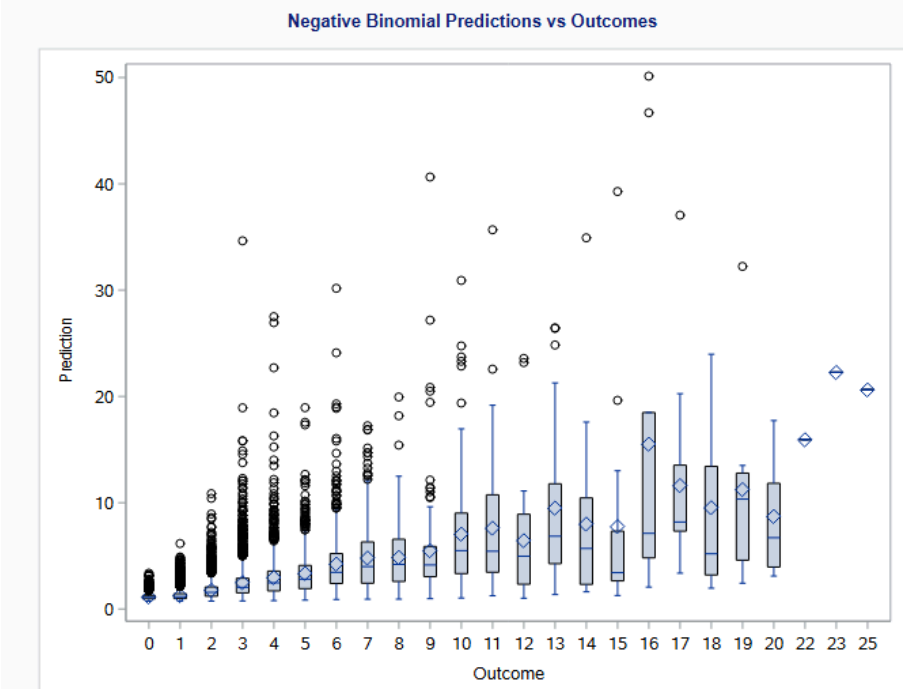


Figure 4.8: Negative Binomial LASSO Model Fit Plot

Fit Plot of Outcome vs. Outpatient Visits, Holding All Other Variables Constant, with 95% Confidence Limits

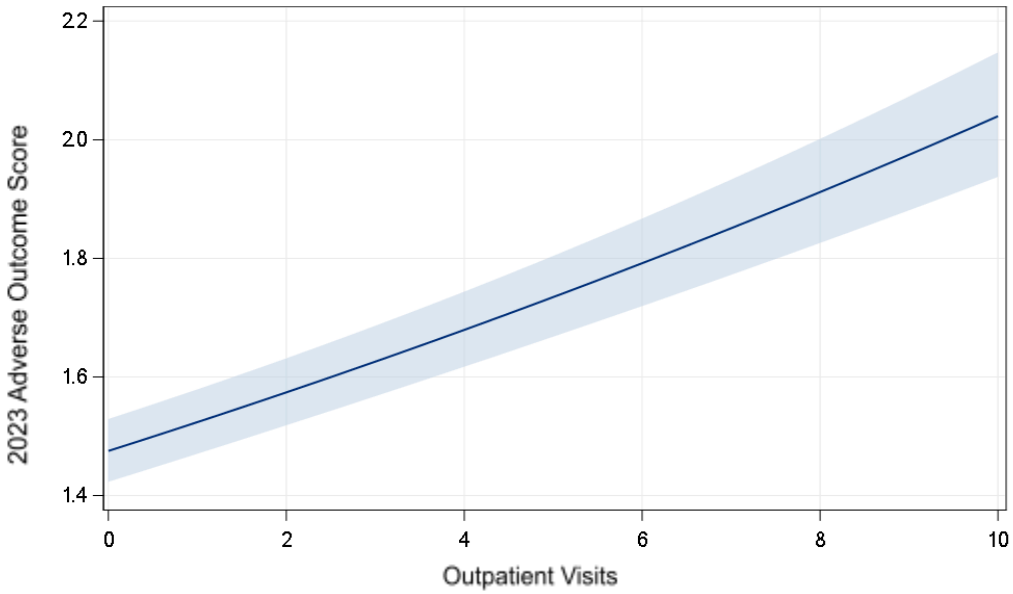


Figure 4.9: Negative Binomial LASSO Model Residual Distribution by Predicted Value

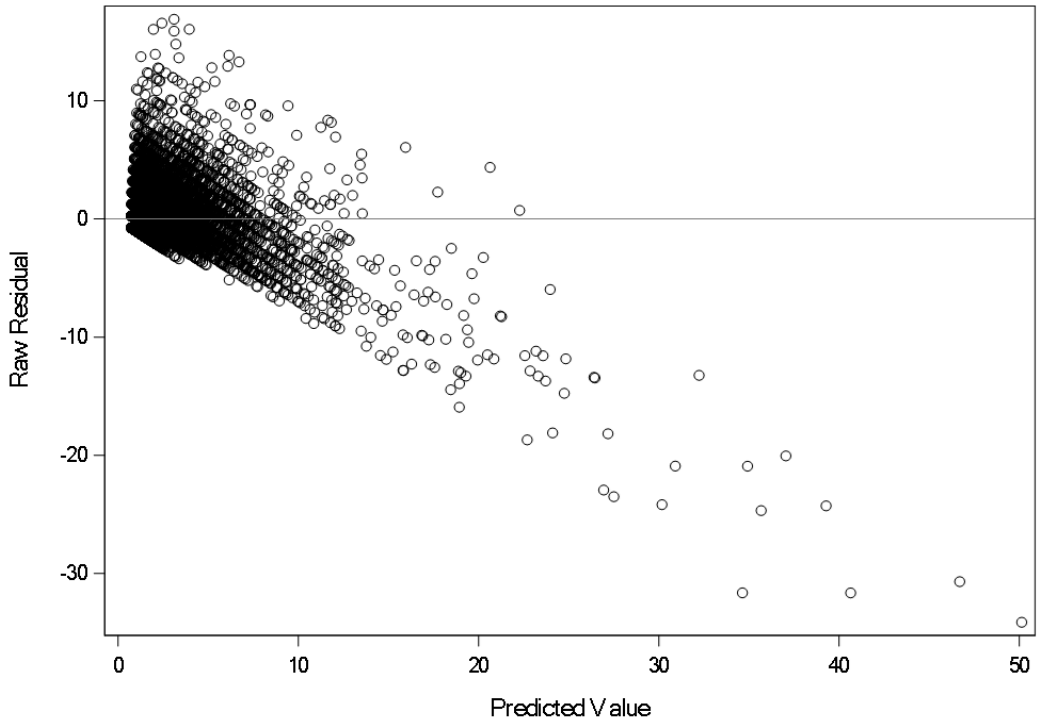


Table 4.8: Preliminary Logistic LASSO Model Fit and Performance Statistics (Training)

Statistic	Value
Calibration Slope (SE)	0.9598 (0.0153)
Calibration Intercept (SE)	0.0022 (0.0021)
AUROC (95% CI - Wald)	0.9066 (0.8985-0.9148)

Figure 4.10: Preliminary Logistic LASSO Calibration Curve (Training)

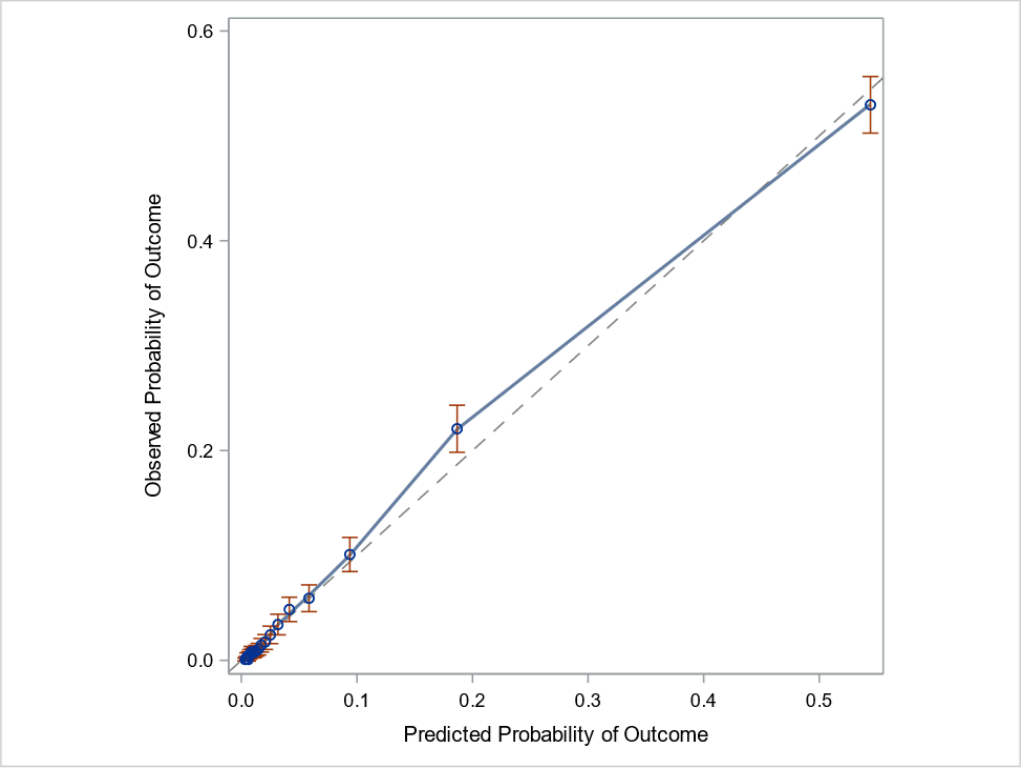


Figure 4.11: Preliminary Logistic LASSO ROC Curve (Training)

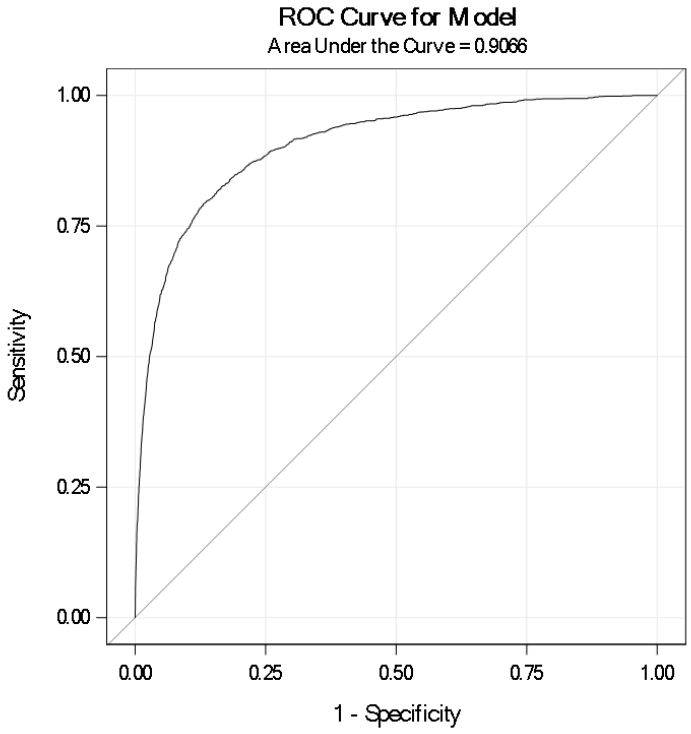


Table 4.9: Threshold-Specific Performance Across Preliminary Models, Top 5% (Training)

Model	Sensitivity	Specificity	PPV	NPV	Accuracy	Event Rate	Alert Rate
Negative Binomial LASSO	46.17%	97.42%	51.21%	96.86%	94.58%	5.55%	5.00%
Linear LASSO	46.44%	97.43%	51.52%	96.87%	94.61%	5.55%	5.00%
Logistic LASSO	47.74%	97.51%	52.96%	96.95%	94.75%	5.55%	5.00%

Figure 4.12: Logistic LASSO Model Selection Plot: Five Fold Cross Validation using Average Squared Error (Training)

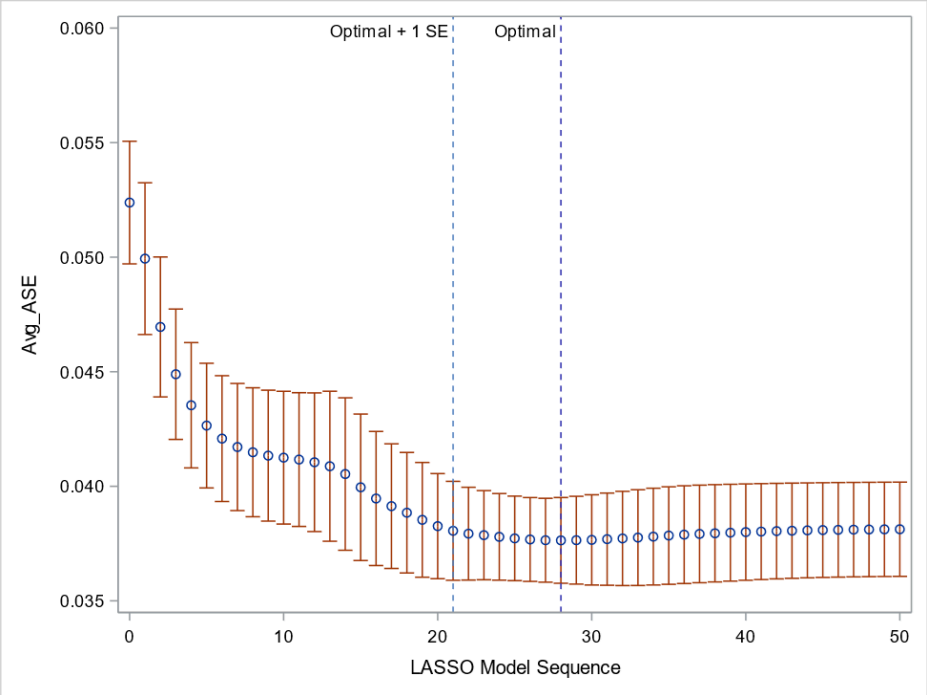


Table 4.10: Model Selection Details for Minimum ASE and Optimal ASE + 1 SE Models Using Five Fold Cross-Validation

Model	Step	Avg Features	Avg ASE	Avg BIC	Avg Lambda
Minimum ASE + 1 SE	**21**	**23**	**0.0381**	**1828**	**0.0092**
	22	25	0.0379	1856	0.0074
	23	30	0.0379	1923	0.0059
	24	35	0.0379	1995	0.0047
	25	41	0.0377	2106	0.0038
	26	45	0.0377	2178	0.0030
	27	48	0.0377	2219	0.0024
Minimum ASE	**28**	**53**	**0.0377**	**2280**	**0.0019**

Table 4.11: Comparison of Variables Chosen by SBC vs. Cross-Validation LASSO Selection for Logistic Model

SBC Selected Parameters	Cross-Validation Selected Parameters (2+ Folds)
AbuseorFosterorPsychosocial	AbuseorFosterorPsychosocial
AGE	
AlcoholDrug	AlcoholDrug 1 vs 0
Anxiety or Depression	Anxiety or Depression
	Anxiety
BHvisitFlag	BHvisitFlag
Depression	Depression
Dev Disorders Flag	Dev Disorders Flag
Drug Use	Drug Use
dx_adhd_pci	dx_adhd_pci
dx_asthma_pci	dx_asthma_pci
dx_obesity_pci	dx_obesity_pci
ED	
ED_cat	ED_cat
LGBTQ Flag	LGBTQ Flag
log_EP_LIMENG	
OPvisitCat	OPvisitCat
PCI	PCI
	PCvisitCat5
PCvisitCat	PCvisitCat
SDOHall	SDOHall
SEX (Female vs. Male)	SEX (Female vs. Male)
SMI_ADHD_dx	SMI_ADHD_dx
Smoking	Smoking
Suicidality	Suicidality

Note: Variables do not align exactly with those in the preliminary logistic model (Table 4.4) and the final logistic model (Table 4.12) because this table presents the variables selected by the LASSO step, before manual refinement.

Table 4.12: Final Logistic Model Parameters

	Crude Parameter Estimates			Adjusted Parameter Estimates		
Parameter	**Estimate**	**Stand. Error**	**Wald P-Value**	**Adj. Estimate**	**Adj. Stand. Error**	**Adj. Wald P-Value**
Abuseorfosterorpsych 1 Vs 0	0.8879	0.0284	<.0001	0.2072	0.0383	<.0001
Alcoholdrug 1 Vs 0	0.843	0.041	<.0001	0.3764	0.0628	<.0001
Anxiety 1 Vs 0	1.0809	0.0293	<.0001	0.2243	0.0428	<.0001
Smidx 1 Vs 0	0.9628	0.0346	<.0001	0.2304	0.0464	<.0001
Bhvisitflag 1 Vs 0	1.3448	0.0364	<.0001	0.1454	0.0573	0.0112
Depression 1 Vs 0	1.0492	0.0281	<.0001	0.2889	0.0423	<.0001
Devdelayflag 1 Vs 0	0.7256	0.0316	<.0001	0.3953	0.0454	<.0001
Dx_adhd_pci 1 Vs 0	1.0572	0.0336	<.0001	0.4926	0.047	<.0001
Dx_asthma_pci 1 Vs 0	0.517	0.0291	<.0001	0.2212	0.0385	<.0001
Dx_eatingdisorder_pc 1 Vs 0	1.0475	0.0834	<.0001	0.2026	0.1112	0.0684
Dx_obesity_pci 1 Vs 0	0.2186	0.0271	<.0001	-0.0432	0.0356	0.2245
Dx_sicklecell 1 Vs 0	0.5788	0.1516	0.0001	0.5593	0.1815	0.0021
Ed_cat	0.0223	945.322		0.247	0.031	<.0001
Lgbtq+ Vs Not Lgbtq+	1.0464	0.0532	<.0001	0.3345	0.0708	<.0001
Missing Data Vs Not Lgbtq+	-1.1466	0.0443	<.0001	-0.5512	0.0547	<.0001
Opvisitcat	0.00774	1738.7425		0.1076	0.012	<.0001
Pci	0.016	2224.3054		0.3136	0.0255	<.0001
Pcvisitcat	0.0147	466.4309		0.0881	0.0195	<.0001
Sdohall 1 Vs 0	0.3943	0.0298	<.0001	0.1866	0.0376	<.0001
Sex Female Vs Male	0.2745	0.0278	<.0001	0.2306	0.0364	<.0001
Curform Smoke Vs Never	1.1607	0.0667	<.0001	0.2984	0.1008	0.0031

Unknown Smoke Vs Never	-0.6981	0.0443	<.0001	-0.0642	0.0637	0.3134
Suicidality 1 Vs 0	1.0391	0.0361	<.0001	0.1477	0.0523	0.0048

Table 4.13: Univariate and Adjusted Odds Ratios for Predictors in Final Logistic Model

	Unadjusted			**Adjusted Model**		
Effect	**OR**	**OR LCL**	**OR UCL**	**aOR**	**aOR LCL**	**aOR UCL**
Abuseorfosterorpsych 1 Vs 0	5.904	5.283	6.599	1.513	1.302	1.759
Alcoholdrug 1 Vs 0	5.398	4.597	6.338	2.123	1.66	2.716
Anxiety 1 Vs 0	8.687	7.744	9.745	1.566	1.324	1.852
Smidx 1 Vs 0	6.859	5.99	7.855	1.585	1.322	1.902
Bhvisitflag 1 Vs 0	14.727	12.767	16.987	1.337	1.068	1.674
Depression 1 Vs 0	8.153	7.302	9.102	1.782	1.51	2.103
Devdelayflag 1 Vs 0	4.268	3.771	4.832	2.205	1.845	2.634
Dx_adhd_pci 1 Vs 0	8.284	7.263	9.449	2.678	2.228	3.22
Dx_asthma_pci 1 Vs 0	2.812	2.509	3.152	1.556	1.338	1.81
Dx_eatingdisorder_pc 1 Vs 0	8.125	5.86	11.267	1.5	0.97	2.319
Dx_obesity_pci 1 Vs 0	1.548	1.392	1.722	0.917	0.798	1.054
Dx_sicklecell 1 Vs 0	3.182	1.757	5.766	3.061	1.503	6.233
Ed_cat	1.986	1.901	2.075	1.28	1.205	1.36
Lgbtq+ Vs Not Lgbtq+	2.576	2.196	3.022	1.125	0.911	1.39
Missing Data Vs Not Lgbtq+	0.287	0.254	0.325	0.464	0.401	0.537
Opvisitcat	1.381	1.36	1.402	1.114	1.088	1.14
Pci	2.128	2.062	2.196	1.368	1.302	1.439
Pcvisitcat	1.373	1.334	1.413	1.092	1.051	1.135
Sdohall 1 Vs 0	2.200	1.957	2.473	1.452	1.253	1.683
Sex Female Vs Male	1.731	1.553	1.931	1.586	1.375	1.829
Curform Smoke Vs Never	5.07	4.141	6.208	1.703	1.264	2.296
Unknown Smoke Vs Never	0.79	0.707	0.883	1.185	1.033	1.36
Suicidality 1 Vs 0	7.99	6.935	9.205	1.344	1.095	1.649

Table 4.14: Final Logistic Model Fit and Performance Statistics (Training)

Statistic	**Value**
Calibration Slope (SE)	0.9586 (0.0154)
Calibration Intercept (SE)	0.0023 (0.0022)
AUROC (95% CI - Wald)	0.9070 (0.8989-0.9151)

Figure 4.13: Final Logistic Model Calibration Curve (Training)

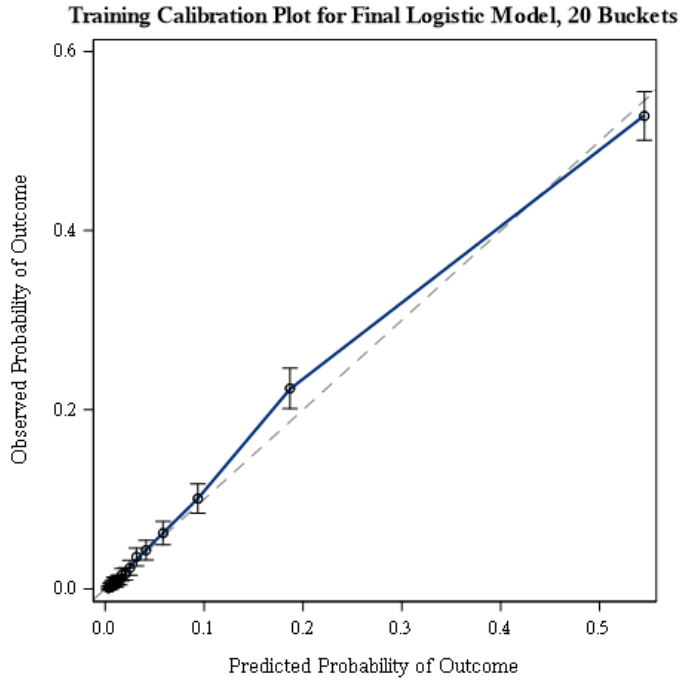


Figure 4.14: Final Logistic Model ROC Curve (Training Sample)

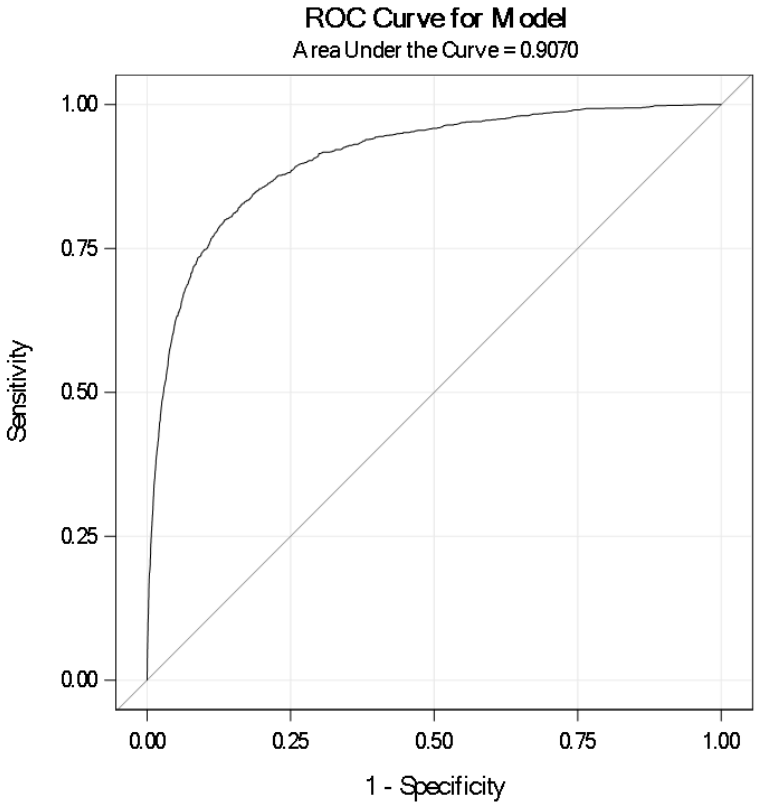


Table 4.15: Threshold-specific Performance of Final Logistic Model across Training and Validation Samples

Model	Sensitivity	Specificity	PPV	NPV	Accuracy	Event Rate	Alert Rate
Final Model: Training	47.61%	97.50%	52.81%	96.94%	94.73%	5.55%	5.00%
Final Model: Validation	46.64%	97.12%	55.62%	95.92%	93.49%	7.19%	6.03%

Table 4.16: Patient Characteristics by Predicted High Risk Status (Training)

Characteristic	Predicted High Risk			P-value
	No (N=25044)	Yes (N=1318)	Total (N=26362)	
Age, n (%)				<.0001[1]
12	4360 (17.4%)	148 (11.2%)	4508 (17.1%)	
13	4256 (17.0%)	181 (13.7%)	4437 (16.8%)	
14	4165 (16.6%)	231 (17.5%)	4396 (16.7%)	
15	4143 (16.5%)	212 (16.1%)	4355 (16.5%)	
16	4057 (16.2%)	280 (21.2%)	4337 (16.5%)	
17	4063 (16.2%)	266 (20.2%)	4329 (16.4%)	
Sex at Birth, n (%)				<.0001[1]
Female	12334 (49.2%)	874 (66.3%)	13208 (50.1%)	
Male	12710 (50.8%)	444 (33.7%)	13154 (49.9%)	
Insurance, n (%)				<.0001[1]
Commercial	5057 (20.2%)	195 (14.8%)	5252 (19.9%)	
Medicaid	18408 (73.5%)	1100 (83.5%)	19508 (74.0%)	
Uninsured	1579 (6.3%)	23 (1.7%)	1602 (6.1%)	
Race/Ethnicity, n (%)				<.0001[1]
Asian	1370 (5.5%)	42 (3.2%)	1412 (5.4%)	
Black	9144 (36.5%)	397 (30.1%)	9541 (36.2%)	
Latinx	11137 (44.5%)	752 (57.1%)	11889 (45.1%)	
Other	3012 (12.0%)	111 (8.4%)	3123 (11.8%)	
White	381 (1.5%)	16 (1.2%)	397 (1.5%)	
Preferred Language, n (%)				0.0002[1]
English	15353 (61.3%)	808 (61.3%)	16161 (61.3%)	
Other	1135 (4.5%)	29 (2.2%)	1164 (4.4%)	
Spanish	8556 (34.2%)	481 (36.5%)	9037 (34.3%)	
LGBTQ+ Status, n (%)				<.0001[1]

LGBTQ+	943 (3.8%)	278 (21.1%)	1221 (4.6%)	
Missing Data	14544 (58.1%)	192 (14.6%)	14736 (55.9%)	
Not LGBTQ+	9557 (38.2%)	848 (64.3%)	10405 (39.5%)	
1+ Social Needs, n (%)				<.0001[1]
1+ SDOH Needs	906 (3.6%)	105 (8.0%)	1011 (3.8%)	
No SDOH Needs Recorded	24138 (96.4%)	1213 (92.0%)	25351 (96.2%)	
Homelessness, n (%)				0.0500[1]
Experienced Homelessness (Last 12m)	522 (2.1%)	38 (2.9%)	560 (2.1%)	
Not Homeless	24522 (97.9%)	1280 (97.1%)	25802 (97.9%)	
Public Housing, n (%)				<.0001[1]
Lives in Public Housing	2477 (9.9%)	239 (18.1%)	2716 (10.3%)	
Not in Public Housing	22567 (90.1%)	1079 (81.9%)	23646 (89.7%)	
Census Tract Poverty Level, n (%)				<.0001[1]
High Poverty (>=30%)	12729 (50.8%)	763 (57.9%)	13492 (51.2%)	
Medium to Low Poverty (0-29%)	12315 (49.2%)	555 (42.1%)	12870 (48.8%)	
Primary Care Visits				<.0001[2]
Mean (SD)	1.8 (1.50)	3.0 (2.17)	1.9 (1.57)	
Median	1.0	3.0	1.0	
Range	1.0, 145.0	1.0, 23.0	1.0, 145.0	
Specialty Visits				<.0001[2]
Mean (SD)	0.9 (2.92)	8.8 (10.05)	1.3 (4.01)	
Median	0.0	5.0	0.0	
Range	0.0, 180.0	0.0, 56.0	0.0, 180.0	
ED Visits				<.0001[2]
Mean (SD)	0.3 (0.68)	1.5 (1.96)	0.3 (0.83)	
Median	0.0	1.0	0.0	
Range	0.0, 10.0	0.0, 22.0	0.0, 22.0	
Inpatient Visits				<.0001[2]
Mean (SD)	0.0 (0.14)	0.2 (0.53)	0.0 (0.19)	
Median	0.0	0.0	0.0	
Range	0.0, 10.0	0.0, 5.0	0.0, 10.0	
Chronic Comorbidities (PCI)				<.0001[2]
Mean (SD)	1.5 (1.31)	5.2 (1.85)	1.7 (1.57)	
Median	1.0	5.0	1.0	
Range	0.0, 8.0	1.0, 15.0	0.0, 15.0	
2+ Somatic Conditions				<.0001[1]
No	24939 (99.6%)	1250 (94.8%)	26189 (99.3%)	
Yes	105 (0.4%)	68 (5.2%)	173 (0.7%)	
Serious Mental Illness Dx				<.0001[1]
No	24067 (96.1%)	882 (66.9%)	24949 (94.6%)	
Yes	977 (3.9%)	436 (33.1%)	1413 (5.4%)	

Depression or Anxiety Dx				<.0001[1]
No	24067 (96.1%)	882 (66.9%)	24949 (94.6%)	
Yes	977 (3.9%)	436 (33.1%)	1413 (5.4%)	
Suicidality				<.0001[1]
No	24300 (97.0%)	894 (67.8%)	25194 (95.6%)	
Yes	744 (3.0%)	424 (32.2%)	1168 (4.4%)	
Pregnancy or STI				<.0001[1]
No	24648 (98.4%)	1249 (94.8%)	25897 (98.2%)	
Yes	396 (1.6%)	69 (5.2%)	465 (1.8%)	
Alcohol or Substance Use				<.0001[1]
No	24310 (97.1%)	1051 (79.7%)	25361 (96.2%)	
Yes	734 (2.9%)	267 (20.3%)	1001 (3.8%)	
Tobacco Use				<.0001[1]
Current or Former	426 (1.7%)	179 (13.6%)	605 (2.3%)	
Never	12031 (48.0%)	662 (50.2%)	12693 (48.1%)	
Unknown	12587 (50.3%)	477 (36.2%)	13064 (49.6%)	
Abuse or Trauma				<.0001[1]
No	23796 (95.0%)	779 (59.1%)	24575 (93.2%)	
Yes	1248 (5.0%)	539 (40.9%)	1787 (6.8%)	
Developmental Delay-Related Disorder				<.0001[1]
No	23102 (92.2%)	863 (65.5%)	23965 (90.9%)	
Yes	1942 (7.8%)	455 (34.5%)	2397 (9.1%)	
School Challenges				<.0001[1]
No	24131 (96.4%)	1171 (88.8%)	25302 (96.0%)	
Yes	913 (3.6%)	147 (11.2%)	1060 (4.0%)	
ADHD Dx				<.0001[1]
No	24109 (96.3%)	820 (62.2%)	24929 (94.6%)	
Yes	935 (3.7%)	498 (37.8%)	1433 (5.4%)	

[1]Chi-Square p-value; [2]Kruskal-Wallis p-value;

Table 4.17: Outcome Distribution by Predicted High Risk Status (Training)

	Predicted High Risk			
	0 (N=25044)	1 (N=1318)	Total (N=26362)	P-value
Actual High Risk, n (%)				<.0001[1]
Not Top 5%	24278 (96.9%)	622 (47.2%)	24900 (94.5%)	
Top 5%	766 (3.1%)	696 (52.8%)	1462 (5.5%)	
2023 Adverse Outcome Score				<.0001[2]
Mean (SD)	1.6 (1.63)	6.3 (3.48)	1.8 (2.05)	
Median	1.0	6.0	1.0	
Range	0.0, 20.0	1.0, 25.0	0.0, 25.0	
Outcome Score Distribution, n (%)				<.0001[1]
0	5370 (21.4%)	0 (0.0%)	5370 (20.4%)	
1	10720 (42.8%)	18 (1.4%)	10738 (40.7%)	
2	4496 (18.0%)	72 (5.5%)	4568 (17.3%)	
3	2016 (8.0%)	193 (14.6%)	2209 (8.4%)	
4	1075 (4.3%)	182 (13.8%)	1257 (4.8%)	
5	601 (2.4%)	157 (11.9%)	758 (2.9%)	
6	766 (3.1%)	696 (52.8%)	1462 (5.5%)	

[1]Chi-Square p-value; [2]Kruskal-Wallis p-value;

Figure 4.15: Outcome Distribution by Predicted High Risk Status (Training)

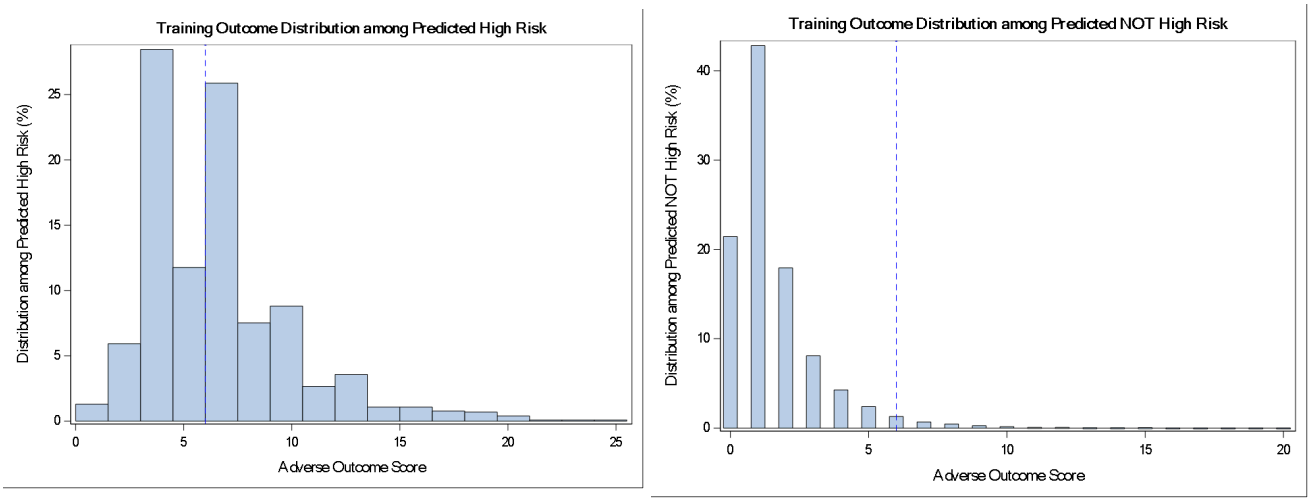


Table 4.18: Patient Characteristics by Lost to Outpatient Care Status (Training)

	Lost to Outpatient Care - 2023			
	0 (N=19372)	1 (N=6990)	Total (N=26362)	P-value
Age, n (%)				<.0001[1]
12	3424 (17.7%)	1084 (15.5%)	4508 (17.1%)	
13	3373 (17.4%)	1064 (15.2%)	4437 (16.8%)	
14	3281 (16.9%)	1115 (16.0%)	4396 (16.7%)	
15	3245 (16.8%)	1110 (15.9%)	4355 (16.5%)	
16	3168 (16.4%)	1169 (16.7%)	4337 (16.5%)	
17	2881 (14.9%)	1448 (20.7%)	4329 (16.4%)	
Sex at Birth, n (%)				0.0003[1]
Female	9837 (50.8%)	3371 (48.2%)	13208 (50.1%)	
Male	9535 (49.2%)	3619 (51.8%)	13154 (49.9%)	
Race/Ethnicity, n (%)				<.0001[1]
Asian	1047 (5.4%)	365 (5.2%)	1412 (5.4%)	
Black	6842 (35.3%)	2699 (38.6%)	9541 (36.2%)	
Latinx	8989 (46.4%)	2900 (41.5%)	11889 (45.1%)	
Other	2254 (11.6%)	869 (12.4%)	3123 (11.8%)	
White	240 (1.2%)	157 (2.2%)	397 (1.5%)	
Preferred Language, n (%)				<.0001[1]
English	11539 (59.6%)	4622 (66.1%)	16161 (61.3%)	
Other	819 (4.2%)	345 (4.9%)	1164 (4.4%)	
Spanish	7014 (36.2%)	2023 (28.9%)	9037 (34.3%)	

LGBTQ+ Status, n (%)				<.0001[1]
LGBTQ+	988 (5.1%)	233 (3.3%)	1221 (4.6%)	
Missing Data	10002 (51.6%)	4734 (67.7%)	14736 (55.9%)	
Not LGBTQ+	8382 (43.3%)	2023 (28.9%)	10405 (39.5%)	
1+ Social Needs, n (%)				0.1088[1]
1+ SDOH Needs	765 (3.9%)	246 (3.5%)	1011 (3.8%)	
No SDOH Needs Recorded	18607 (96.1%)	6744 (96.5%)	25351 (96.2%)	
Homelessness, n (%)				<.0001[1]
Experienced Homelessness (Last 12m)	369 (1.9%)	191 (2.7%)	560 (2.1%)	
Not Homeless	19003 (98.1%)	6799 (97.3%)	25802 (97.9%)	
Public Housing, n (%)				0.2299[1]
Lives in Public Housing	2022 (10.4%)	694 (9.9%)	2716 (10.3%)	
Not in Public Housing	17350 (89.6%)	6296 (90.1%)	23646 (89.7%)	
Census Tract Poverty Level, n (%)				0.0720[1]
High Poverty (>=30%)	9979 (51.5%)	3513 (50.3%)	13492 (51.2%)	
Medium to Low Poverty (0-29%)	9393 (48.5%)	3477 (49.7%)	12870 (48.8%)	
Primary Care Visits				<.0001[2]
Mean (SD)	2.0 (1.72)	1.6 (0.95)	1.9 (1.57)	
Median	2.0	1.0	1.0	
Range	1.0, 145.0	1.0, 13.0	1.0, 145.0	
Specialty Visits				<.0001[2]
Mean (SD)	1.6 (4.54)	0.5 (1.65)	1.3 (4.01)	
Median	0.0	0.0	0.0	
Range	0.0, 180.0	0.0, 39.0	0.0, 180.0	
Behavioral Health Visits				<.0001[2]
Mean (SD)	0.5 (3.16)	0.1 (1.21)	0.4 (2.79)	
Median	0.0	0.0	0.0	
Range	0.0, 56.0	0.0, 39.0	0.0, 56.0	
ED Visits				<.0001[2]
Mean (SD)	0.4 (0.85)	0.3 (0.77)	0.3 (0.83)	
Median	0.0	0.0	0.0	
Range	0.0, 22.0	0.0, 15.0	0.0, 22.0	
Inpatient Visits				0.3443[2]
Mean (SD)	0.0 (0.17)	0.0 (0.22)	0.0 (0.19)	
Median	0.0	0.0	0.0	
Range	0.0, 5.0	0.0, 10.0	0.0, 10.0	
Chronic Comorbidities (PCI)				<.0001[2]
Mean (SD)	1.8 (1.62)	1.3 (1.37)	1.7 (1.57)	
Median	1.0	1.0	1.0	
Range	0.0, 15.0	0.0, 13.0	0.0, 15.0	

[1]Chi-Square p-value; [2]Kruskal-Wallis p-value;

Table 4.19: Sensitivity Analysis for Top 5% Performance by Lost to Outpatient Care Status (Training)

Sample	Sensitivity	Specificity	PPV	NPV	Accuracy	Alert Rate	Event Rate
Full Training Sample	47.61%	97.50%	52.81%	96.94%	94.73%	5.55%	5.00%
Lost to Outpatient Care (2023)	40.74%	98.21%	21.02%	99.30%	97.54%	2.25%	1.16%
***Not* Lost to Outpatient Care (2023)**	47.94%	97.23%	57.07%	96.05%	93.72%	5.99%	7.13%

Table 4.20: Sociodemographic Bias Evaluation for Top 5% High Risk (Training)

Class	Subgroup	PPV	PPD	Bias by PPD	FNR	EOD	Bias by EOD	Event Rate	Alert Rate	Sensitivity	Specificity	Accuracy
Insurance	Commercial	49.23%	(4.45%)	No	59.49%	9.16%	No	4.51%	3.71%	40.51%	98.03%	95.43%
	Medicaid	53.69%	.		50.34%	.		6.09%	5.63%	49.66%	97.22%	94.33%
	Uninsured	39.13%	(14.55%)	Yes	75.68%	25.34%	Yes	2.31%	1.44%	24.32%	99.11%	97.38%
Language	Other	65.52%	12.67%	Yes	48.65%	-2.16%	No	3.18%	2.49%	51.35%	99.11%	97.59%
	English	52.85%	.		50.81%	.		5.37%	5.00%	49.19%	97.51%	94.91%
	Spanish	51.88%	-0.97%	No	55.30%	-4.49%	No	6.16%	5.31%	44.70%	97.28%	94.04%
Race	Asian	45.24%	(7.49%)	No	61.22%	10.11%	Yes	3.47%	2.97%	38.78%	98.31%	96.25%
	Black	55.16%	2.43%	No	52.80%	1.69%	No	4.86%	4.16%	47.20%	98.04%	95.57%
	Latinx	52.73%	.		51.11%	.		6.81%	6.32%	48.89%	96.80%	93.53%
	Other	48.65%	(4.08%)	No	56.45%	5.34%	No	3.97%	3.55%	43.55%	98.10%	95.93%
	White	43.75%	(8.98%)	No	53.33%	2.22%	No	3.78%	4.03%	46.67%	97.64%	95.72%
Sex	Female	53.15%	.		49.46%	.		6.95%	6.61%	50.54%	96.67%	93.47%
	Male	52.03%	(1.12%)	No	57.54%	8.08%	No	4.14%	3.38%	42.46%	98.31%	96.00%

Table 4.21: Validation Calibration and Discrimination

Statistic	Value
Calibration Slope (SE)	1.0296 (0.0150)
Calibration Intercept (SE)	0.0023 (0.0023)
AUROC (95% CI - Wald)	0.8930 (0.8809-0.9052)

Figure 4.16: Validation Calibration Curve

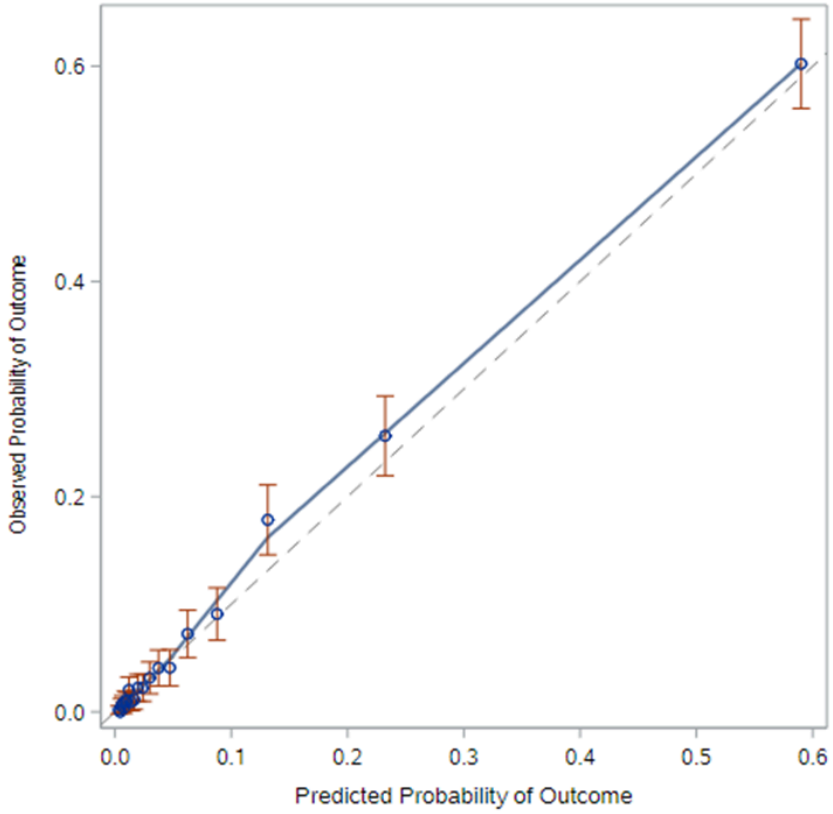
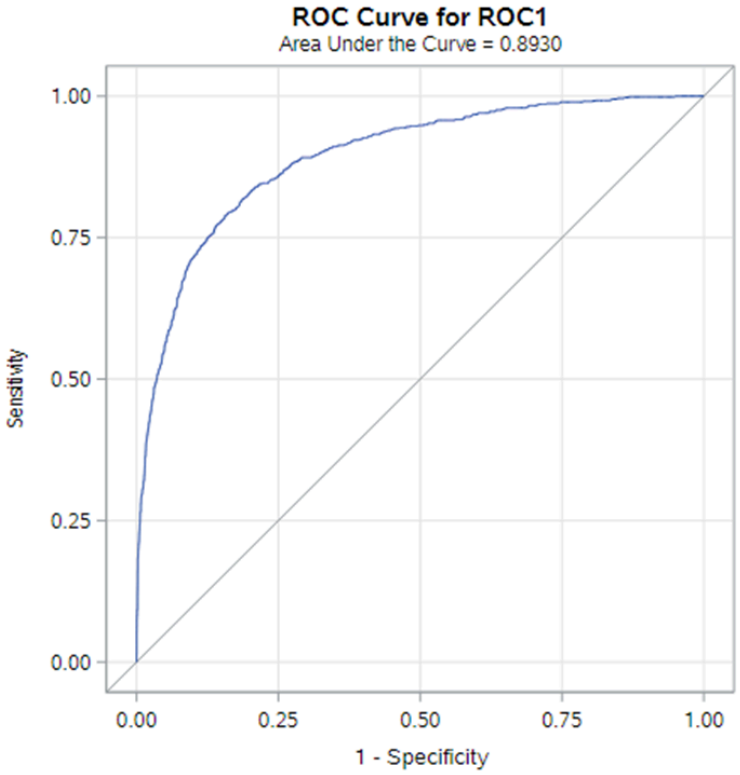


Figure 4.17: Validation ROC Curve



Discussion

Table 5.1: Logic Model for Implementation of an Adolescent Risk Score in Primary Care at NYC Health+Hospitals

Goal	Identify high risk adolescents and connect them to effective primary care resources

Resources	Activities	Outputs	Short-term Outcomes	Mid-term Outcomes	Long-term Outcomes
• Office of Population Health staff time to run model monthly • IT time to integrate model results into Epic EMR • IT time to build new Epic workflows for HR flag patients • Data from Enterprise Data Warehouse • Analyst time to monitor drift and bias • Peds time to develop and deliver training	• Adolescent risk model runs monthly to identify children at highest risk of experiencing ACE-associated adverse outcomes in the following year • High risk flag and risk score are pushed into Epic and flagged in "suggested risk tier" in patient's chart • Best practice alert fires for high risk adolescents during primary care visit, advising provider to complete full set of screenings (Subject Adolescent Council approval)	• High Risk adolescent flag in Epic • Best practice alert for screening, referrals for High Risk patients • High risk order set of screenings, referrals for High Risk patients • Workbench report of high risk patients with a visit in the next week (pre-visit planning tool) • Workbench report of High Risk patients due or overdue for a visit, with no visit scheduled	• PCP completes full set of SHADESS screenings including social needs at well child visits and provides tailored guidance to adolescent • PCP provides warm handoff to CHW (in facilities with live CHW in adolescence program) • PCP refers patients with mental health needs to Collaborative Care OR to Behavioral Health • PCP refers patients with chronic medical needs that cannot be managed in Primary Care to specialty care • PCP refers patients with social needs to internal supports (legal, SNAP enrollment) or community supports (food insecurity, housing) via Find Help platform • Patient scheduled for follow-up visit and prioritized for no-show reminders (subject to Adolescent Council approval) • Patients requiring additional follow-up attend additional primary care visits in the year after HR identification	• Adolescents and families receive CHW support to address top family challenges • Adolescents with mild to moderate depression, anxiety or ADHD receive regular coaching and appropriate medication management from Collaborative Care program • Adolescents with serious mental health needs receive counseling and appropriate medication management from Behavioral Health • Adolescents with chronic conditions that require additional management visit relevant specialty provider • Eligible families with adolescents and social needs enrolled in SNAP • Families with food insecurity visit local food pantries identified by Find Help • Families with legal needs receive appropriate guidance and legal advice from NYLAG on immigration status, eviction and other issues	• High risk adolescents experience higher quality of life • Adolescent school attendance and performance improves • Family food insecurity reduced • Family housing security increased • Family stress level reduced • Adolescent self-efficacy increased • Adolescents' with depression or anxiety are well managed and symptoms are below clinical thresholds • Adolescent with more serious mental health needs achieve improvement in symptoms

Appendices

Appendix A: Key Informant Interview Methods

Key Informant Overview No single gold standard definition exists to define risk in adolescence. The adolescent risk outcome definition portion of this work was initiated before the author began the Marywood doctoral program in August of 2023. This work defined adolescent risk by sourcing adverse outcomes associated with ACEs in the literature (S. Nelson et al., 2021) and refining them using patient and provider key informants at H+H, to produce a highly localized outcome of interest representing local practice and needs.

H+H key informant experts and patients were asked to evaluate individual and domain-level ACE-associated outcomes for their importance to risk assessment in adolescent primary care. They were asked to flag crucial missing domains. Finally, they were asked to suggest meaningful ways to measure these domains as specific variables inside an electronic medical record. The key informant feedback shaped the final analytic dataset and informed a continuous score representing the number of ACE-associated adverse outcomes experienced by each adolescent in 2023.

Sample Key informant feedback was collected before this study. As such, key informant selection criteria and processes are covered here only briefly. From March to May 2023, NYC Health+Hospitals conducted 14 key informant interviews, with 10 NYC Health+Hospitals staff members and four adolescent patients. Participants were recruited with a snowball methodology. The seed key informants were solicited from a group of co-investigators participating in a pediatric risk study focusing on replicating a HealthySteps methodology for children ages 0-4. Suggestions for additional key informants were solicited from this group and from each key informant who completed an interview. Informed consent was not solicited, as the key informants were deemed not to be human subjects by the IRB. No compensation was offered for participation. Key informants were informed that participation was entirely voluntary and that individual quotations would not be attributed to them by name. Permission to record the

session was solicited from participants. Recordings were done via Webex and stored within the H+H network.

Overview: The key informant interview followed a semi-structured format. As outlined by Nina Kammerer in her guide to preparing qualitative proposals, the interviews were purposefully designed to be adaptive, following the wisdom gleaned from the key informants (Kammerer 2013). Each key informant started the interview with an overview of risk prediction goals and possible approaches. Key informants then had a chance to ask questions and to offer overall reactions to the proposed goals and approach. Finally, key informants were asked to rate domains on importance to risk assessment in primary care for adolescents.

Procedure: Key informants were asked to review adverse outcome domains identified by Nelson et al to be significantly associated with ACEs exposure. They were instructed to identify domains that were essential, domains which should *not* be measured, and domains that might be missing or incomplete in the provided list. The initial domains (depression, anxiety, ADHD, suicidality, self-harm, early sexual debut/high risk sexual activity, substance use, alcohol use, victim of violence, higher ED utilization, lower primary care utilization, poor asthma control, teen pregnancy, disordered eating) expanded as key informants identified other key areas that they felt should be included. Key informants were also asked what they felt were the most important risk predictors in their own practice, and which screenings were routinely done in primary care that should be leveraged for risk prediction. Key informants rated each proposed measure on a pseudo Lickert scale (4= strongly agree, 3= agree, 2= disagree, 1=strongly disagree). The list of measures changed through the process in response to suggestions about additions and deletions from the key informants. The interviews closed with a request for any additional thoughts or feedback that the interview process had missed. Notes were taken via excel throughout key informant interviews, and Webex recordings were consulted to improve notes where full quotes were not captured. Interview transcript notes were analyzed twice. They were analyzed at the midpoint of the interview process, to ensure that any key themes that were

emerging were added to the interview process, and again once all interviews were complete. Interview transcript notes were analyzed using Microsoft Excel.

Appendix B: Key Informant Select Results

Key Informant Interviews

Key Informant Characteristics The key informants included a diverse group of viewpoints in both lived experience and clinical practice (see *Key Informant Characteristics* table below). The final key informant group included five pediatric and adolescent primary care providers, two of whom also hold administrative leadership positions in addition to clinical care, and all of whom hold either adjunct or full professorship at NYU Grossman School of Medicine. The key informants also included two behavioral health providers, including a child psychiatrist and a pediatrician who runs the NYC Health+Hospitals pediatric child abuse program. The adult participants also included three pediatric and adolescent social workers, one of whom works primarily in administration with a focus on adolescent health. The four adolescent participants were recruited from the Youth Leadership Council, and were current or former patients at NYC Health+Hospitals. A limitation of the key informant interviews was the lack of inclusion of the nursing perspective, due to inability to gain permission to interview the nurses recruited for the study. A majority of the key informants were female (71.4%), broadly mirroring the demographics of those working in this area. Half of participants were mid-career (aged 30-49). Half the participants were white and 35.7% were African-American (35.7%), with one Asian-American participant, and one Two or More Races/Other participant. A limitation of the key informant perspectives represented in this study is the lack of inclusion of a Hispanic/Latinx participant, which would have been remedied had the two solicited nurses been able to participate. Findings should be investigated for their applicability across ethnicities and spoken languages during the implementation phase, after study conclusion.

Table: Key Informant Characteristics

Characteristics	**Number of Participants**

Role	Behavioral Health Provider	2
	PCP	5
	Social Worker	3
	Student	4
Age	14-17	4
	30-49	7
	50-65	3
Gender	Female	10
	Male	4
Race/Ethnicity	African-American	5
	Asian-American	1
	Other	1
	White	7
Total Participants		**14**

Findings: Risk Prediction Overview Overall, key informants thought that the outcome domains significantly associated with ACEs exposure in the literature covered the major outcome domains of adolescent risk in their clinical and lived experience. As one key informant said, “Some of these are going to be easier to capture than others, but I think they are all outcomes worthy of inclusion." Multiple key informants flagged that creation of an aggregated adolescent risk score was a challenging but worthy endeavor:

> "[Traditional] risk scores were developed in disease care systems, which are oriented around hospital costs. You are looking to predict risk based on health *outcomes*, which

> has a venn diagram with healthcare utilization, but is a distinct thing.... Making that referral and making sure that they get...that service? That's where there's a big gap in the tools that we have to help kids and adults."

Key informants felt that an integrated risk score which triggered services within the EMR would be beneficial, and were particularly excited about the possibility of adjusting primary care patient panel sizes by risk score, as well as risk score-informed pre-visit planning in the context of an expanded Community Health Worker program roll-out.

Areas of Consensus: Strongly Recommended for Inclusion Key informants agreed on the importance of many proposed socio-medical outcome domains and measures (see Key Informant Measure Feedback table below). Using a pseudo-lichert scale of 4 (strongly agree) to 1 (strongly disagree), key informants rated at a 3.5 or above visits for depression, anxiety, disordered eating, self-harm, perpetration or victimhood of violence, abuse, suicidal ideation or attempt, substance use disorder or alcohol use disorder. According to one key informant, behavioral health concerns were the most important area for this age group: “"Certainly, [I am most concerned about] mental health, like suicidality, history of suicide, depression or anxiety. Those for me are the kids I want to spend more time with, getting [them] the care they need and [making sure] that they have good safety plans [in place].” The key informants were unified in their recommendations to include visits for substance use disorder (SUD) or alcohol use disorder (AUD). They also favored including visits to a chemical dependency or methadone clinic or a prescription of buprenorphine or methadone (for opioid use), but there were no such visits or prescriptions in 2023, upon exploratory data analysis. All measures of pregnancy and sexually transmitted infection (STI) outcomes were uniformly recommended for inclusion as adverse outcomes. Finally, a mid-interview addition of separate Psych ED and Psych inpatient visits (split from overall ED/inpatient) received universally strong agreement.

Additions: Medical Complexity Multiple key informants flagged the need for an addition of autism, developmental disorders and learning disabilities, which they felt were responsible for a

significant amount of acute utilization: "Kids with developmental disabilities are often left out, and I think it's important to include them. Especially those who have significant autism disorder or cognitive challenges. They have disruptive behaviors, and if they don't have good access to care, they end up using our emergency room a lot." This variable was added via the developmental delay and pervasive developmental disorder PCI flags. Two key informants also identified the need to capture other chronic childhood illnesses like diabetes, epilepsy, IBD, Crohn's and sickle cell, to capture children facing significant medical complexity who were at elevated medical risk. Most key informant feedback was oriented simultaneously towards both predictor and outcome measures. When asked if medical complexity measures needed to be captured as outcomes, most key informants agreed that these measures might be better served as predictors.

Additions: Behavioral Health Key informants reported additional available measures of suicidality, including the final question of the PHQ-9 screener, the ASQ, and the CCSR (although the ASQ was found to be paper-based and therefore not eligible for inclusion). Multiple key informants recommended adding a prescription of PrEP, a medication to prevent transmission of HIV, as another way to capture higher risk sexual activity, although this variable was not adequately captured for the reporting period to be included. Key informants reported use of PTSD, trauma and abuse ICD-10 codes in visit diagnoses, which were added to the tracked categories. One key informant specifically recommended the CDC Social Vulnerability Index for geographic SDOH enrichment, which was adopted. Another key informant expressed doubt that zip-code level geographic data would have predictive power in New York City, due to the challenge of having public housing complexes or shelters located in affluent areas, and recommended use of census tract data if available. This suggestion was also adopted. Key informants recommended the addition of tobacco use and vaping as an outcome, and suggested that any included substance use screening focus on non-marijuana drug use in

particular. Marijuana couldn't be reliably disaggregated from other substance use in screening data. The ultimate SUD variable captured all forms of drug and substance use.

Another key addition area centered around somatic symptoms, or physical symptoms which express underlying distress. In the literature, allergies, dizziness, headache, nausea, insomnia, and weight gain or loss are significantly associated with exposure to ACEs. All of these symptoms are non-specific, meaning that they are associated with many other things as well as ACEs and may not in and of themselves indicate that a child is at elevated risk. As part of the key informant interview, key informants were asked whether any of these symptoms rose to the level of concern to merit inclusion, in their opinion. All key informants agreed that allergies should not be included. Key informants felt that BMI trajectory, and in particular rapid weight loss or gain, was cause for concern. Most but not all key informants felt that insomnia was an important risk factor. For other symptoms, key informants felt that presence of multiple symptoms were an important somatic clue, particularly among diverse populations, as is outlined in the text box below.

Key Informants on the Importance of Somatic Symptoms

- " Especially coming from cultures that don't express emotions, it comes out somatically...The ones that are sticking out to me are headache, dizziness and insomnia. "
- "I think this gets into our somatizing symptoms. For patients where there's lots going on for trauma and mental health, there may be physical symptoms like headache or nausea. If you have three symptoms (headaches, GI symptoms, dizziness - those would be my top three)... that might be worth considering."
- "I think a lot of patients who present with depression, anxiety or mental health concerns often come across as having a lot of somatic complaints like headache or belly pain. I think it's hard to pick up in the medical record [but]...one thing I do see a

lot in kids is headaches, and that does come up a lot with mental health comorbidities. GI symptoms, headache, dizziness/nausea, and insomnia. Those four [are the ones where] we typically see issues."

- "The ones I think should stay would be insomnia, headaches/migraines and GI symptoms, as well as if there's been any shift in the BMI percentile. "
- "I think headache or insomnia are vague complaints. People with vague complaints likely have anxiety or depression, so maybe that's worth including."
- "Dizziness, nausea - they definitely can be muddy but I wouldn't eliminate them."

Key informant feedback was synthesized as the presence of 2 or more symptoms among insomnia, headache and GI issues, using diagnoses. BMI trajectory (top or bottom 5%) was included as a separate outcome.

Areas of Mixed Feedback: Utilization Some key informants felt that utilization data would also act as an important proxy outcome for medical complexity, with the suggestions to include number of primary care visits, number of specialty visits, and number of different types of specialists visited. However, key informants were divided about the inclusion of these measures as outcomes, particularly the primary care measure, as they felt this might actually be measuring the receipt of needed care versus flagging children who were not getting care adequate to their risk state. As one key informant said, “With youth, one of the problems is that they are not seeking help for their problems.” Another worried that aggregate visit counts wouldn’t capture the appropriateness of care: “One of the things I struggle with a little bit, in terms of predictors and outcomes, is that you want the right kind of utilization at the right time. When you have low value care, whether it's not the right amount of something or too much of something, it's a problem.” Key informants did feel more comfortable using very high outpatient utilization (≥95th percentile) as an outcome. Ultimately, this study used the number of primary care visits as a predictor but not as an outcome, and tested ≥85th percentile specialty visits at

H+H as an outcome. Counting the number of different specialists visited was not feasible with currently available data.

Areas of Mixed Feedback: Bias and Behavioral Health Measures The issue of bias in healthcare was a key concern among key informants. Despite a very lively conversation about whether to include race/ethnicity in predictive and clinical algorithms that is currently underway in academia, key informants did *not* object to the use of race/ethnicity within the H+H adolescent risk algorithm. Key informants recognized the impact of the lived experience of racism: “In that socio-ecological model, one of the things I would put in there is racism. Because it's really big…having switched my practice to central Harlem where my patients are mostly black from [facilities] where it was mostly hispanic, the impact of racism is mind-boggling." Concerns about bias extended to bio-measures, which might be ordered differentially. For example, two key informants suggested including positive urine toxicology results, but one key informant was concerned about bias in toxicology screening orders, with screens more likely to be ordered for Black youth. That variable was ultimately not included because data were not available.

The most lively conversation centered around bias and aggression-associated visit diagnoses. Five providers and all teen key informants felt that aggression-related diagnoses should be included as predictors of outcomes in some form. However the majority of providers expressed ambivalence about including a variable that has been robustly been shown in the literature to be differentially assigned as a label to children of color, particularly Black children - see text box below.

Key Informants on Aggression-Related Disorders and Bias

- "I don't really like these terms. It's a little tricky. I'm of two minds. If anyone is coming in with a primary diagnosis of aggression or behavioral health problems, inherently there's going to be bias, especially around kids who are BIPOC and marginalized. I hesitate to include it in an algorithm because of the implicit bias. Underlying them are

depression, anxiety and underlying trauma - I see these more as symptoms and not a diagnosis."

- "I don't like the ODD diagnosis. I think that underneath, when you look at ODD, this comes out in more of skipping school, getting into fights. I think you want to be more concrete with the behavior so that it's not judgemental...Acting out is acting to get a need met...I don't know what the options are, but if we could pull out behaviors [it's better]. If we do the ODD diagnosis, it's already creating bias."
- "There is a lot of bias in the diagnosis, but it doesn't mean that there's not a problem if it's there [and this kid needs more support.] I will say, having switched my practice to central Harlem...it's mind-boggling the impact of racism. It’s unbelievable how many kids have school dysfunction and behavioral issues, and then get an identity around [it]."
- "I can understand why [it's controversial] - this is also baked into racism and how we document in the EMR, and need to be really thoughtful about it. And so much of this is also associated with autism. That's a whole other bag of worms as well. Similarly to our non-specific things associated with ACEs, is there a certain threshold where this doesn't become noise and signals an issue? And is this also correlated with high ED or IP numbers?"

Key informants identified several proxy measures for aggression, such as psychiatric ED or Inpatient visits, which could be used in place of aggression-related diagnoses. As stated by one provider, “"ED visits potentially related to [aggression], and psychiatric [issues] overlap here. You're going to have parents or caregivers sending a teen who is out of control to the psych emergency room." Another provider suggested that aggression-related disorders might often occur with other conditions that could more accurately be captured, “like autism, schizophrenia/affective disorder, bipolar, or major depression with psychosis.” Providers also

flagged common alternative diagnoses, such as parent-child conflict that might more robustly get at the underlying challenge. Ultimately, conduct disorder and psychosis were both included in a composite serious mental illness prediction flag, in addition to PCI and behavioral health utilization flags.

Missing Data Key informants lamented the lack of standardized documentation around school difficulties such as documentation of an IEP, which they felt were key predictors *and* outcomes in and of themselves. As one key informant stated, “I don't know that there's any way, [but I’d want to capture] kids who are struggling in school. That's not necessarily in the EMR, but if they're having a hard time at school, they're more likely to get into trouble in other ways, [like] maybe using drugs, hanging out in the streets, or getting involved in gang activities." Bullying was a key area of risk for adolescents that all felt was missing from the EHR. Multiple key informants also felt that social media and phone use were important but missing data that are not well captured within the EMR.

Summary The key informant process was an invaluable component of the inquiry. The key informants clarified ambiguous areas, affirmed the overall approach and direction of the inquiry, and highlighted essential additional outcomes and predictors that were added to the candidate variable list. They also identified areas where data entry was likely too heterogeneous or sparse to use. Where data availability allowed, outcome domains and measures recommended by the key informants were included in the study. Where data were not available, or exploratory data analysis showed that no data were recorded for a given domain, the variable was not included despite key informant recommendations. Additional details are available by domain in the Key Informant Measure Ratings table.

Table: Key Informant Measure Ratings

Domain	Variable	Average Key Inf. Score	N	Action
ADHD	ADHD Dx	3.1	10	Included
Alcohol Use Disorder	Alcohol Use Disorder Dx	3.9	10	Included and augmented w/screening
Alcohol Use Disorder	Antabuse Rx	2	3	Not included - no Rx recorded in time frame
Depression & Anxiety	Depression Dx	3.6	10	Included and augmented w/screening
Depression & Anxiety	Anxiety Dx	3.6	10	Included and augmented w/screening
Higher Acute Utilization	# ED Psych Visits	4	6	Collapsed with Non-Psych ED and included in 3+ ED Visit Flag + Behavioral Health Dx-Specific ED Visits
Higher Acute Utilization	# Inpatient Psych Visits	4	6	Included
Higher Acute Utilization	# ED Visits	3.4	10	Collapsed with Non-Psych ED and included

Higher Acute Utilization	# Inpatient Visits	3.1	10	Included
Higher Acute Utilization	# PC-sensitive ED Visits	2.8	10	Not included - poor quality data without external tool
Higher Risk Sexual Behavior	Pregnancy/Live Birth/Abortion Dx	3.8	10	Included in pregnancy flag
Higher Risk Sexual Behavior	Prenatal/Abortion /Miscarriage Visit	3.9	10	Included in pregnancy flag
Higher Risk Sexual Behavior	STI Dx	3.9	10	Included in STI flag
Higher Risk Sexual Behavior	Positive STI Lab Result	3.9	10	Included in STI flag
Higher Risk Sexual Behavior	PrEP Rx	3.3	6	Not included - no Rx recorded in time frame
Key Informant Addition	Disordered Eating Dx	3.7	10	Included

Key Informant Addition	# Rx (Polypharmacy)	2.5	8	Not included - not feasible in EHR data. Proxy captured in comorbidity index.
Key Informant Addition	>11 Outpatient Visits	3.2	5	Included - changed to 5+ visits to align to 85th percentile at HH
Key Informant Addition	# Specialty Visits	3.1	8	Collapsed with above
Lower PC Utilization	# No-Show/Canceled PC Visits	3.1	10	Not included - poor quality data
Lower PC utilization	No Well-Child Visit	2.5	10	Included as proxy for loss-to-follow-up in outcome year
Lower PC utilization	# Primary Care Visits	2.4	5	Included as predictor but not outcome
Poor Asthma Control	Persistent Asthma Dx	3.1	10	Included for ED visits only - key informants felt poor control was the source of risk, and other measures of poor control not available.
Substance Use Disorder	Substance Use Disorder Dx	3.9	10	Included and augmented w/screening

Substance Use Disorder	Positive Urine Toxicology Result	3.3	7	Not included - concerns about stigma and low data availability
Substance Use Disorder	Urine Toxicology order	2.5	4	Not included - concerns about stigma and low data availability
Substance Use Disorder	Buprenorphine or Methadone Rx	4	5	Not included - no Rx recorded in time frame
Substance Use Disorder	Chem. Dependency or Methadone Visit	3.8	5	Not included - no adolescent visits in prediction year
Suicidality & Self-Harm	Suicidal Ideation/Attempt Dx	4	10	Included and augmented w/screening
Suicidality & Self-Harm	Self Harm Dx	3.7	10	Collapsed with Suicidality
Trauma, Abuse & Maltreatment	Abuse Visit Dx	4	10	Included
Trauma, Abuse &	Violent Injury Dx	3.9	10	Included

Maltreatment				
Trauma, Abuse & Maltreatment	Perpetrator of Violence Dx	3.8	5	Not Included - diagnoses not available. Victim of violence diagnoses were included instead, within the abuse, trauma and neglect composite flag.

Note: adolescent key informants provided feedback on all variable domains but did not score individual variables, due to the highly clinical nature of the variables. Max N is therefore 10.

Appendix C: Data Dictionary

Domain	Variable	Var Type	Definition
Abuse, Trauma, Neglect	AbuseNeg_23	Outcome	1 = Positive SSHADESS safety issue screening or 1+ visits with a visit diagnosis of abuse, maltreatment or neglect (AHRQ CCSR) in 2023. 0 = No positive screen or visit diagnosis in 2023
Abuse, Trauma, Neglect	dx_eneuresis_23	Outcome	1 = 1+ visits with a visit diagnosis of enuresis or encopresis (AHRQ CCSR) in 2023. 0 = No such visit diagnosis in 2023
Abuse, Trauma, Neglect	dx_psychosocial_23	Outcome	1 = 1+ visits with a visit diagnosis of socioeconomic or psychosocial difficulties (CCSR, Z60.x - 65.x) in 2023. 0 = no such visit diagnosis in 2023.
Abuse, Trauma, Neglect	dx_trauma_ptsd_23	Outcome	1 = 1+ visits with a visit diagnosis of emotional trauma or post traumatic stress disorder (CCSR) in 2023. 0 = No such visits in 2023
Alcohol & Substance Use	alcohol_23	Outcome	1 = Positive alcohol use screen or 1+ visits with a visit diagnosis of alcohol use disorder (PCI) in 2023

			0 = No positive screen or visit diagnosis in 2023
Alcohol & Substance Use	druguse_all_23	Outcome	1 = Patient screened positive for marijuana use or substance use; or 1+ visits with a diagnosis of substance use (PCI) or overdose (CCSR) in 2023. 0 = No positive screen or visit diagnosis in 2023
Alcohol & Substance Use	dx_overdose_23	Outcome	1 = 1+ visits with a visit diagnosis of a substance overdose (AHRQ CCSR) in 2023. 0 = No such visit diagnosis in 2023
Alcohol & Substance Use	smoke_23	Outcome	Curform = Patient screened positive for current or former tobacco use or vaping on SSHADESS screener or received 1+ visit diagnoses for tobacco use in 2023. Never = no current or former tobacco use recorded in 2023. Missing = Smoking status not collected or marked as "unknown" in 2023.

Complexity	pciflag_23	Outcome	0 = 0-2 score on the Pediatric Comorbidity Index for 2023, aka patient in the bottom 75th percentile of adolescents at H+H 1 = 3-4 score on the Pediatric Comorbidity Index for 2023 2 = 5+ score on the Pediatric Comorbidity Index for 2023, aka in the top 94th percentile of adolescents at H+H (closest integer to 95th)
General Chronic Disease	dx_sicklecell_23	Outcome	1 = 1+ visits with a visit diagnosis of sickle cell disorder (AHRQ CCSR) in 2023. Note that PCI category was not used because it includes sickle cell trait, which this definition does not. 0 = No such visit diagnosis in 2023
Healthy Eating	bmitraj_23	Outcome	1 = Patients' change in BMI from 2022 to 2023 was ≤5th percentile or ≥95th percentile of adolescents with a BMI measured at H+H in 2022 and 2023. 0 = Patient not in top or bottom 5% of BMI trajectory change from 2022 to 2023.

Healthy Eating	dx_eatingdisorder_pci_23	Outcome	1 = 1+ visits with a visit diagnosis of eating disorders (PCI) in 2023. 0 = No such visit diagnosis in 2023
Healthy Eating	dx_weightloss_pci_ed23	Outcome	1 = 1+ visits to the Emergency Department for weight loss (PCI) in 2023. 0 = No such ED visits in 2023
Mental Health	anxiety_23	Outcome	1 = Patient screened positive on the GAD-7 or 1+ visits with a visit diagnosis of anxiety (PCI) in 2023. 0 = No positive screen or visit diagnosis in 2023
Mental Health	depression_23	Outcome	1 = Patient screened positive on the PHQ-9 or 1+ visits with a visit diagnosis of depression (PCI) in 2023. 0 = No positive screen or visit diagnosis in 2023
Mental Health	dx_anxiety_pci_ed23	Outcome	1 = 1+ visits to the Emergency Department for anxiety (PCI) in 2023. 0 = No such ED visits in 2023
Mental Health	dx_depression_pci_ed23	Outcome	1 = 1+ visits to the Emergency Department for depression (PCI) in 2023. 0 = No such ED visits in 2023

Mental Health	SMIdx_23	Outcome	1 = 1+ visits with a visit diagnosis of: disordered conduct, obsessive compulsive disorder, psychosis, or specified personality disorders (PCI); or 1+ visits with a diagnosis of having witnessed or perpetrated violence (AHRQ CCSR) in 2023. 0 = No such visit diagnosis in 2023
Mental Health	somatic_23	Outcome	1 = Patient had 1+ visits in 2023 for two of the following three somatic symptom diagnosis categories: sleep problems, headache or migraine, nausea or vomiting (PCI). 0 = Patient did not record 2+ somatic symptom diagnoses in 2023.
Mental Health	suicidality_23	Outcome	1=Patient screened positive for suicidality on the CCSR or question 9 of the PHQ-9, or patient had 1+ visits with a suicidality diagnosis (PCI) in 2023. 0 = No screening or diagnosis of suicidality recorded in 2023.
School Challenges	DevDelayFlag_23	Outcome	1 = 1+ visits with a visit diagnosis of: chromosomal abnormalities, developmental delay or pervasive

			developmental disorders (PCI) in 2023. 0 = No such visit diagnosis in 2023
School Challenges	dx_adhd_pci_23	Outcome	1 = 1+ visits with a visit diagnosis of ADHD (PCI) in 2023. 0 = No such visit diagnosis in 2023
School Challenges	SchoolProb_23	Outcome	1 = Challenges in school or poor school performance SSHADESS screening in 2023 or 1+ visits with a visit diagnosis of educational difficulties (AHRQ CCSR) in 2023. 0= No screening or visit diagnosis recorded in 2023.
Sexual Health	PREG_DEL_ABO_MIS_FLAG_23	Outcome	1 = Pregnancy, delivery, abortion or miscarriage documented in 2023. 0 = No pregnancy, delivery, abortion or miscarriage documented in 2023.
Sexual Health	STI_FLAG_23	Outcome	1 = Patient tested positive for syphilis, chlamydia, gonorrhea; received an STI-specific medication prescription, or was diagnosed with 1+ STIs in 2023 (internal H+H grouper, curated by Sexual Health providers). 0 = No lab, medication or diagnostic evidence of STI in 2023.

Social Determinants	dx_lifestyle_23	Outcome	1 = 1+ visit with visit diagnosis for problems relating to lifestyle, habits and healthy behaviors (Z72.x family, AHRQ CCSR) in 2023. 0 = no such visit diagnosis in 2023.
Social Determinants	SDOHneed_23	Outcome	1 = Patient screened positive for food insecurity in SSHADESS or positive for any domain on the social determinants screener in 2023. 0 = No positive social needs recorded in 2023.
Uncontrolled Chronic Disease	dx_asthma_pci_ed23	Outcome	1 = 1+ visits to the Emergency Department for asthma (PCI) in 2023. 0 = No such ED visits in 2023
Uncontrolled Chronic Disease	dx_cardiovascular_pci_ed23	Outcome	1 = 1+ visits to the Emergency Department for cardiovascular issues (PCI) in 2023. 0 = No such ED visits in 2023
Uncontrolled Chronic Disease	dx_diabetes_pci_ed23	Outcome	1 = 1+ visits to the Emergency Department for diabetes (PCI) in 2023. 0 = No such ED visits in 2023
Uncontrolled Chronic Disease	dx_epilepsy_pci_ed23	Outcome	1 = 1+ visits to the Emergency Department for epilepsy (PCI) in 2023. 0 = No such ED visits in 2023

Uncontrolled Chronic Disease	dx_gi_pci_ed23	Outcome	1 = 1+ visits to the Emergency Department for chronic gastrointestinal disease (PCI) in 2023. 0 = No such ED visits in 2023
Unmet BH Needs	dx_conduct_pci_ed23	Outcome	1 = 1+ visits to the Emergency Department for conduct disorder (PCI) in 2023. 0 = No such ED visits in 2023
Unmet BH Needs	dx_ocd_pci_ed23	Outcome	1 = 1+ visits to the Emergency Department for obsessive compulsive disorder (PCI) in 2023. 0 = No such ED visits in 2023
Unmet BH Needs	dx_psychotic_pci_ed23	Outcome	1 = 1+ visits to the Emergency Department for psychosis (PCI) in 2023. 0 = No such ED visits in 2023
Unmet BH Needs	dx_specpersonality_pci_ed23	Outcome	1 = 1+ visits to the Emergency Department for specified personality disorders (PCI) in 2023. 0 = No such ED visits in 2023
Unmet Developmental Needs	dx_adhd_pci_ed23	Outcome	1 = 1+ visits to the Emergency Department for ADHD (PCI) in 2023. 0 = No such ED visits in 2023

Unmet Developmental Needs	dx_chromosomal_pci_ed23	Outcome	1 = 1+ visits to the Emergency Department for chromosomal abnormalities (PCI) in 2023. 0 = No such ED visits in 2023
Unmet Developmental Needs	dx_congmal_pci_ed23	Outcome	1 = 1+ visits to the Emergency Department for congenital malformations (PCI) in 2023. 0 = No such ED visits in 2023
Unmet Developmental Needs	dx_devdelay_pci_ed23	Outcome	1 = 1+ visits to the Emergency Department for developmental delay (PCI) in 2023. 0 = No such ED visits in 2023
Unmet Developmental Needs	dx_perdevdis_pci_ed23	Outcome	1 = 1+ visits to the Emergency Department for pervasive developmental disorders (PCI) in 2023. 0 = No such ED visits in 2023
Utilization	ed_flag_23	Outcome	1 = Patient had 3+ visits to the ED in 2023 (Source: PCI recommended cutoff, verified by key informants) 0 = Patient had 0-2 visits to ED in 2023.
Utilization	IP_NONPSYCH_Flag_23	Outcome	1 = Patient had 2+ visits to inpatient setting, not including inpatient psychiatric units, in 2023.

			0 = Patient had 0-1 visits to non-psychiatric inpatient setting in 2023.
Utilization	IP_PSYCH_Flag_23	Outcome	1 = Patient had a visit to a psychiatric inpatient unit in 2023. 0 = no visits to inpatient psychiatric units in 2023.
Utilization	OPvisitFlag_23	Outcome	1 = Patient had 5+ outpatient visits in 2023 exclusive of primary care, and was therefore in the 95th percentile of adolescent outpatient utilization at H+H. 0 = 0-4 outpatient visits in 2023.
Utilization	WCV_FLAG_23	Outcome	1 = Patient did not attend a Well Child Visit in 2023, either because it was missed or it was not scheduled. 0 = Patient attended a Well Child Visit in 2023.
Abuse, Trauma, Neglect	dx_abuse_neglect	Predictor	1= Patient had 1+ visits with a visit diagnosis of abuse, maltreatment or neglect (AHRQ CCSR) in 2022. 0 = No such visits in 2022.
Abuse, Trauma, Neglect	dx_eneuresis	Predictor	1 = 1+ visits with a visit diagnosis of enuresis or encopresis (AHRQ

			CCSR) in 2022. 0 = No such visit diagnosis in 2022
Abuse, Trauma, Neglect	dx_psychosocial	Predictor	1 = 1+ visits with a visit diagnosis of socioeconomic or psychosocial difficulties (CCSR, Z60.x - 65.x) in 2022. 0 = no such visit diagnosis in 2022.
Abuse, Trauma, Neglect	dx_trauma_ptsd	Predictor	1 = 1+ visits with a visit diagnosis of emotional trauma or post traumatic stress disorder (CCSR) in 2022. 0 = No such visit diagnosis in 2022
Abuse/Trauma	AbuseorFoster	Predictor	1 = Positive SSHADESS safety issue screening or Foster screening or 1+ visits with a visit diagnosis of abuse, maltreatment or neglect (AHRQ CCSR) in 2022. 0 = No positive screen or visit diagnosis in 2022
Abuse/Trauma	AbuseorFosterorPsychosocial	Predictor	1 = Positive SSHADESS safety issue screening or Foster screening or 1+ visits with a visit diagnosis of abuse, maltreatment, neglect or psychosocial issues (AHRQ CCSR) in 2022. 0 = No positive screen or visit diagnosis in 2022

Alcohol & Substance Use	AlcoholDrug	Predictor	1 = Positive alcohol use screen OR substance use screen or 1+ visits with a visit diagnosis of alcohol use disorder or substance use (PCI) or overdose (AHRQ CCSR) in 2022 0 = No positive screen or visit diagnosis in 2022
Alcohol & Substance Use	druguse_all	Predictor	1 = Positive substance use screen or 1+ visits with a visit diagnosis of substance use (PCI) or overdose (AHRQ CCSR) in 2022 0 = No positive screen or visit diagnosis in 2022
Alcohol & Substance Use	smoke	Predictor	Curform = Patient screened positive for current or former tobacco use or vaping on SSHADESS screener or received 1+ visit diagnoses for tobacco use in 2023. Never = no current or former tobacco use recorded in 2022. Missing = Smoking status not collected or marked as "unknown" in 2022.
Alcohol & Substance Use	alcohol	Predictor	1 = Positive alcohol use screen or 1+ visits with a visit diagnosis of alcohol

			use disorder (PCI) in 2022 0 = No positive screen or visit diagnosis in 2022
Complexity	PCI	Predictor	Continuous pediatric comorbidity index score for 2022, censored at 10+.
Demographics	AGE	Predictor	Age as of end of 2022
Demographics	English	Predictor	1 = English recorded as patient's preferred language. 0 = English not recorded as patient's preferred language.
Demographics	Insurance	Predictor	if (PRIMARY_PAYER_GRP='Commercial' or PRIMARY_PAYER_GRP='Other') then Insurance='Commercial'; if (PRIMARY_PAYER_GRP='Medicaid' or PRIMARY_PAYER_GRP='MedicaidMC' or PRIMARY_PAYER_GRP='MedicareMC') then insurance='Medicaid'; if PRIMARY_PAYER_GRP='Self-pay' then insurance='Uninsured' ;
Demographics	Language_lang3	Predictor	English = patient's preferred language listed as English. Spanish = patient's preferred

			language listed as Spanish. Other = patient's preferred language not listed as Spanish or English
Demographics	LGBTQ_FLAG	Predictor	Not LGBTQ+ = patient's sexual orientation listed as straight/heterosexual and patients' gender identity listed as cis female or cis male on SOGI screening Missing = Both sexual orientation and gender identity SOGI screenings missing or unknown/declined to answer LGBTQ+ = SOGI screening completed for gender identity or sexual orientation and NOT equal to unknown/declined, straight/heterosexual, or cisgender
Demographics	race	Predictor	Hispanic/Latinx = patient's ethnicity listed as Hispanic/Latinx. Black or African American = patient's race listed as 'Black' and ethnicity not listed as Hispanic/Latinx. Asian = patient's race listed as 'Asian/Native Hawaiian/Pacific Islander' and ethnicity not listed as

			Hispanic/Latinx. White = patient's race listed as 'White' and ethnicity not listed as Hispanic/Latinx. Other = patient's race and ethnicity not listed as any of the above categories.
Demographics	SEX	Predictor	Sex at birth, as recorded in registration. Captures only Male, Female in this sample.
General Chronic Disease	dx_asthma_pci	Predictor	1 = 1+ visits with a visit diagnosis of asthma (PCI) in 2022. 0 = No such visit diagnosis in 2022
General Chronic Disease	dx_cardiovascular_pci	Predictor	1 = 1+ visits with a visit diagnosis of cardiovascular disease (PCI) in 2022. 0 = No such visit diagnosis in 2022
General Chronic Disease	dx_diabetes_pci	Predictor	1 = 1+ visits with a visit diagnosis of diabetes (PCI) in 2022. 0 = No such visit diagnosis in 2022
General Chronic Disease	dx_epilepsy_pci	Predictor	1 = 1+ visits with a visit diagnosis of epilepsy (PCI) in 2022. 0 = No such visit diagnosis in 2022
General Chronic Disease	dx_gi_pci	Predictor	1 = 1+ visits with a visit diagnosis of chronic gastrointestinal disease (PCI)

			in 2022. 0 = No such visit diagnosis in 2022
General Chronic Disease	dx_sicklecell	Predictor	1 = 1+ visits with a visit diagnosis of sickle cell disorder (AHRQ CCSR) in 2022. Note that PCI category was not used because it includes sickle cell trait, which this definition does not. 0 = No such visit diagnosis in 2022
Healthy Eating	dx_eatingdisorder_pci	Predictor	1 = 1+ visits with an eating disorder visit diagnosis (PCI) in 2022. 0 = No such visit diagnosis in 2022
Healthy Eating	dx_obesity_pci	Predictor	1 = 1+ visits with a visit diagnosis of obesity (PCI) in 2022. 0 = No such visit diagnosis in 2022
Healthy Eating	dx_weightloss_pci	Predictor	1 = 1+ visits with a visit diagnosis of weight loss (PCI) in 2022. 0 = No such ED visits in 2022.
Mental Health	anxiety	Predictor	1 = Patient screened positive on the GAD-7 or 1+ visits with a visit diagnosis of anxiety (PCI) in 2022. 0 = No positive screen or visit diagnosis in 2022
Mental Health	anxietydepress	Predictor	1 = Patient screened positive on the PHQ-9 or GAD-7 or 1+ visits with a visit diagnosis of depression or anxiety

			(PCI) in 2022. 0 = No positive screen or visit diagnosis in 2022
Mental Health	SMI_ADHD_dx	Predictor	1 = 1+ visits with a visit diagnosis of: ADHD, disordered conduct, obsessive compulsive disorder, psychosis, or specified personality disorders (PCI); or 1+ visits with a diagnosis of having witnessed or perpetrated violence (AHRQ CCSR) in 2022. 0 = No such visit diagnosis in 2022
Mental Health	SMIdx	Predictor	1 = 1+ visits with a visit diagnosis of: disordered conduct, obsessive compulsive disorder, psychosis, or specified personality disorders (PCI); or 1+ visits with a diagnosis of having witnessed or perpetrated violence (AHRQ CCSR) in 2022. 0 = No such visit diagnosis in 2022
Mental Health	dx_conduct_pci	Predictor	1 = 1+ visits with a visit diagnosis of conduct disorder (PCI) in 2022. 0 = No such visit diagnosis in 2022
Mental Health	dx_ocd_pci	Predictor	1 = 1+ visits with a visit diagnosis of obsessive compulsive disorder (PCI)

			in 2022. 0 = No such visit diagnosis in 2022
Mental Health	dx_psychotic_pci	Predictor	1 = 1+ visits with a visit diagnosis of psychosis (PCI) in 2022. 0 = No such visit diagnosis in 2022
Mental Health	dx_violence	Predictor	1 = 1+ visits with perpetration of violence-related diagnosis (CCSR, r45.x, 46.x) in 2022. 0 = no such visit.
Mental Health	somatic	Predictor	1 = Patient had 1+ visits in 2022 for two of the following three somatic symptom diagnosis categories: sleep problems, headache or migraine, nausea or vomiting (PCI). 0 = Patient did not record 2+ somatic symptom diagnoses in 2022.
Mental Health	suicidality	Predictor	1=Patient screened positive for suicidality on the CCSR or question 9 of the PHQ-9, or patient had 1+ visits with a suicidality diagnosis (PCI) in 2022. 0 = No screening or diagnosis of suicidality recorded in 2022.
Mental Health	Depression	Predictor	1 = Patient screened positive on the PHQ-9 or 1+ visits with a visit

			diagnosis of depression (PCI) in 2022. 0 = No positive screen or visit diagnosis in 2022
PCI Other	dx_anemia_pci	Predictor	1 = 1+ visits with a visit diagnosis of anemia-related disorders, including sickle cell trait and sickle cell disease (PCI) in 2022. 0 = No such visit diagnosis in 2022
PCI Other	dx_injuries_pci	Predictor	1 = 1+ visits with a visit diagnosis of injury (PCI) in 2022. 0 = No such visit diagnosis in 2022
PCI Other	dx_joint_pci	Predictor	1 = 1+ visits with a visit diagnosis of joint disorders (PCI) in 2022. 0 = No such visit diagnosis in 2022
PCI Other	dx_menstrual_pci	Predictor	1 = 1+ visits with a visit diagnosis related to menstrual-related problems (PCI) in 2022. 0 = No such visit diagnosis in 2022
PCI Other	dx_migraine_pci	Predictor	1 = 1+ visits with a visit diagnosis related to headache or migraine (PCI) in 2022. 0 = No such visit diagnosis in 2022
PCI Other	dx_nauseavomit_pci	Predictor	1 = 1+ visits with a visit diagnosis related to nausea or vomiting (PCI) in

			2022. 0 = No such visit diagnosis in 2022
PCI Other	dx_pain_pci	Predictor	1 = 1+ visits with a visit diagnosis of a pain-related disorder (PCI) in 2022. 0 = No such visit diagnosis in 2022
PCI Other	dx_sleep_pci	Predictor	1 = 1+ visits with a visit diagnosis of insomnia or a sleep disorder (PCI) in 2022. 0 = No such visit diagnosis in 2022
School Challenges	DevDelayFlag	Predictor	1 = 1+ visits with a visit diagnosis of: chromosomal abnormalities, developmental delay or pervasive developmental disorders (PCI) in 2022. 0 = No such visit diagnosis in 2022
School Challenges	dx_adhd_pci	Predictor	1 = 1+ visits with a visit diagnosis of ADHD (PCI) in 2022. 0 = No such visit diagnosis in 2022
School Challenges	dx_chromosomal_pci	Predictor	1 = 1+ visits with a visit diagnosis of chromosomal abnormalities (PCI) in 2022. 0 = No such visit diagnosis in 2022
School Challenges	dx_congmal_pci	Predictor	1 = 1+ visits with a visit diagnosis of congenital malformation (PCI) in

			2022. 0 = No such visit diagnosis in 2022
School Challenges	dx_devdelay_pci	Predictor	1 = 1+ visits with a visit diagnosis of developmental delay (PCI) in 2022. 0 = No such visit diagnosis in 2022
School Challenges	dx_education	Predictor	1 = 1+ visits with a visit diagnosis of problems related to education and literacy (Z55.x family, AHRQ CCSR) in 2022. 0 = no such visit diagnosis in 2022.
School Challenges	dx_perdevdis_pci	Predictor	1 = 1+ visits with a visit diagnosis of a pervasive developmental disability (PCI) in 2022. 0 = No such visit diagnosis in 2022
School Challenges	SchoolProb	Predictor	1 = Challenges in school or poor school performance SSHADESS screening in 2022 or 1+ visits with a visit diagnosis of educational difficulties (AHRQ CCSR) in 2022. 0= No screening or visit diagnosis recorded in 2022.
Sexual Health	Risky_sex	Predictor	1 = Patient screened positive for STI (see STI definition) or received care for delivery, abortion or miscarriage in 2022. Note that pregnancy is not

			included in the definition, because of the potential for data leakage to 2023. 0 = no evidence of positive STI or delivery, abortion or miscarriage in 2022.
Sexual Health	STI_FLAG	Predictor	1 = Patient tested positive for syphilis, chlamydia, gonorrhea; received an STI-specific medication prescription, or was diagnosed with 1+ STIs in 2022 (internal H+H grouper, curated by Sexual Health providers). 0 = No lab, medication or diagnostic evidence of STI in 2022.
Social Determinants	dx_lifestyle	Predictor	1 = 1+ visit with visit diagnosis for problems relating to lifestyle, habits and healthy behaviors (Z72.x family, AHRQ CCSR) in 2022. 0 = no such visit diagnosis in 2022.
Social Determinants	HOMELESS_FLG	Predictor	1 = Patient appeared on H+H registry of patients experiencing homelessness within 12 months of a 2022 visit. 0 = Patient did not appear on H+H registry of patients experiencing homelessness in this time frame.

Social Determinants	Housing	Predictor	1 = Patient home address matched to NYCHA Housing OR Patient appeared on H+H registry of patients experiencing homelessness within 12 months of a 2022 visit. 0 = Patient did not appear on H+H registry of patients experiencing homelessness or match to NYCHA housing in this time frame.
Social Determinants	NYCHA_FLG	Predictor	1 = Patient's home address matched with a known NYCHA public housing property in 2022. 0 = Patient's home address did not match with NYCHA.
Social Determinants	resilience	Predictor	1 = Patient reported exercise or positive sources of social support in SSHADESS screening. 0 = SSHADESS screening complete but social support or exercise not reported Missing = SSHADESS screenings not completed.
Social Determinants	SDOHall	Predictor	1 = positive for variables sdohneed=1 or HOMELESS_FLG=1 or NYCHA_FLG=1 or dx_lifestyle= 1.

			0 = not positive for any of those social variables.
Social Determinants	SDOHneed	Predictor	1 = Patient screened positive for food insecurity in SSHADESS or positive for any domain on the social determinants screener in 2022. 0 = No positive social needs recorded in 2022.
Social Determinants	SVI_EP_CROWD	Predictor	Log transformed continuous Social Vulnerability Index measure of census tract-level residential crowding (EP_CROWD)
Social Determinants	SVI_EP_HBURD	Predictor	Log transformed continuous Social Vulnerability Index measure of census tract-level Housing Cost Burden (EP_HBURD)
Social Determinants	SVI_EP_LIMENG	Predictor	Log transformed continuous Social Vulnerability Index measure of census tract-level English Language Proficiency (EP_LIMENG)
Social Determinants	SVI_EP_NOHSDP	Predictor	Log transformed continuous Social Vulnerability Index measure of census tract-level No High School Diploma (EP_NOHSDP)

Social Determinants	SVI_EP_NOINT	Predictor	Log transformed continuous Social Vulnerability Index measure of census tract-level No broadband internet (EP_NOINT)
Social Determinants	SVI_EP_POV150	Predictor	Log transformed continuous Social Vulnerability Index measure of census tract-level Below 150% Poverty (EP_POV150)
Social Determinants	SVI_EP_SNGPNT	Predictor	Log transformed continuous Social Vulnerability Index measure of census tract-level Single-Parent Households (EP_SNGPNT).
Social Determinants	SVI_EP_UNEMP	Predictor	Log transformed continuous Social Vulnerability Index measure of census tract-level Unemployed (EP_UNEMP)
Social Determinants	SVI_EP_UNINSUR	Predictor	Log transformed continuous Social Vulnerability Index measure of census tract-level No Health Insurance (EP_UNINSUR)
Social Determinants	SVI_RPL_THEME1	Predictor	Log transformed continuous Social Vulnerability Index measure of census tract-level score for socioeconomic needs (RPL_THEME1)
Social Determinants	SVI_RPL_THEME2	Predictor	Log transformed continuous Social Vulnerability Index measure of census

			tract-level score for household characteristics (RPL_THEME2)
Social Determinants	SVI_RPL_THEMES	Predictor	Log transformed continuous Social Vulnerability Index measure of zip code level Overall Ranking within NYS (RPL_THEMES)
Utilization	BHvisitFlag	Predictor	1= Had an outpatient behavioral health visit in 2022 0= did not have an outpatient BH visit in 2022
Utilization	ED	Predictor	Number of ED visits in 2022 (continuous)
Utilization	ED_cat	Predictor	Continuous number of ED visits in 2022, censured at 5+ (representing 99% of the sample).
Utilization	ED_NONPSYCH_cat	Predictor	Continuous number of ED visits to a non-psychiatric unit in 2022, censured at 5+ (representing 99% of the sample).
Utilization	ED_NONPSYCH_Flag	Predictor	1 = Patient had 3+ visits to a non-psychiatric unit of the ED in 2022 (Source: PCI recommended cutoff, verified by key informants) 0 = Patient had 0-2 visits to ED in 2022.

Utilization	ED_PSYCH_Flag	Predictor	1 = 1+ visits to an emergency department psychiatric unit in 2022. 0 = no such visits in 2022.
Utilization	IP	Predictor	Continuous number of inpatient visits in 2022.
Utilization	IP_flag	Predictor	1 = 1+ inpatient visits in 2022. 0 = no inpatient visits in 2022.
Utilization	OPvisitCat	Predictor	Continuous number of specialty outpatient visits, censored at 10+.
Utilization	OPvisitFlag	Predictor	1 = Patients with 5+ specialty outpatient visits not including primary care, in 2022 (representing 95th percentile of H+H adolescents). 0 = Patient had 0 -4 specialty visits in 2022.
Utilization	PCvisitCat	Predictor	Continuous 2022 primary care visits, censored at 20+.
Utilization	PCvisitCat5	Predictor	Continuous 2022 primary care visits, censored at 5+.
Utilization	WCV_FLAG	Predictor	1 = Patient had a Well Child Visit in 2022. 0 = Patient did not have a Well Child Visit in 2022.

Appendix D: TRIPOD+AI Checklist

TRIPOD+AI

Version: 11-January-2024

Section/Topic	Item	Development / evaluation[1]	Checklist item	Reported on page
TITLE				
Title	1	D;E	Identify the study as developing or evaluating the performance of a multivariable prediction model, the target population, and the outcome to be predicted	4
ABSTRACT				
Abstract	2	D;E	See TRIPOD+AI for Abstracts checklist	N/A
INTRODUCTION				
Background	3a	D;E	Explain the healthcare context (including whether diagnostic or prognostic) and rationale for developing or evaluating the prediction model, including references to existing models	3-4
	3b	D;E	Describe the target population and the intended purpose of the prediction model in the context of the care pathway, including its intended users (e.g., healthcare professionals, patients, public)	7-10
	3c	D;E	Describe any known health inequalities between sociodemographic groups	1-3
Objectives	4	D;E	Specify the study objectives, including whether the study describes the development or validation of a prediction model (or both)	8-10
METHODS				
Data	5a	D;E	Describe the sources of data separately for the development and evaluation datasets (e.g., randomised trial, cohort, routine care or registry data), the rationale for using these data, and representativeness of the data	53-54;
	5b	D;E	Specify the dates of the collected participant data, including start and end of participant accrual; and, if applicable, end of follow-up	53,57-59
Participants	6a	D;E	Specify key elements of the study setting (e.g., primary care, secondary care, general population) including the number and location of centres	53
	6b	D;E	Describe the eligibility criteria for study participants	53-54
	6c	D;E	Give details of any treatments received, and how they were handled during model development or evaluation, if relevant	N/A
Data preparation	7	D;E	Describe any data pre-processing and quality checking, including whether this was similar across relevant sociodemographic groups	64-67
Outcome	8a	D;E	Clearly define the outcome that is being predicted and the time horizon, including how and when assessed, the rationale for choosing this outcome, and whether the method of outcome assessment is consistent across sociodemographic groups	55-60
	8b	D;E	If outcome assessment requires subjective interpretation, describe the qualifications and demographic characteristics of the outcome assessors	N/A
	8c	D;E	Report any actions to blind assessment of the outcome to be predicted	N/A
Predictors	9a	D	Describe the choice of initial predictors (e.g., literature, previous models, all available predictors) and any pre-selection of predictors before model building	60-64
	9b	D;E	Clearly define all predictors, including how and when they were measured (and any actions to blind assessment of predictors for the outcome and other predictors)	60-64
	9c	D;E	If predictor measurement requires subjective interpretation, describe the qualifications and demographic characteristics of the predictor assessors	N/A
Sample size	10	D;E	Explain how the study size was arrived at (separately for development and evaluation), and justify that the study size was sufficient to answer the research question. Include details of any sample size calculation	53-54
Missing data	11	D;E	Describe how missing data were handled. Provide reasons for omitting any data	65-67
Analytical methods	12a	D	Describe how the data were used (e.g., for development and evaluation of model performance) in the analysis, including whether the data were partitioned, considering any sample size requirements	54-55
	12b	D	Depending on the type of model, describe how predictors were handled in the analyses (functional form, rescaling, transformation, or any standardisation)	65, 67
	12c	D	Specify the type of model, rationale[2], all model-building steps, including any hyperparameter tuning, and method for internal validation	68-77
	12d	D;E	Describe if and how any heterogeneity in estimates of model parameter values and model performance was handled and quantified across clusters (e.g., hospitals, countries). See TRIPOD-Cluster for additional considerations[3]	75
	12e	D;E	Specify all measures and plots used (and their rationale) to evaluate model performance (e.g., discrimination, calibration, clinical utility) and, if relevant, to compare multiple models	68-77
	12f	E	Describe any model updating (e.g., recalibration) arising from the model evaluation, either overall or for particular sociodemographic groups or settings	75-77
	12g	E	For model evaluation, describe how the model predictions were calculated (e.g., formula, code, object, application programming interface)	74, 78
Class imbalance	13	D;E	If class imbalance methods were used, state why and how this was done, and any subsequent methods to recalibrate the model or the model predictions	Not used
Fairness	14	D;E	Describe any approaches that were used to address model fairness and their rationale	77, 112-116
Model output	15	D	Specify the output of the prediction model (e.g., probabilities, classification). Provide details and rationale for any classification and how the thresholds were identified	69, 74-75

TRIPOD+AI

Version: 11-January-2024

Training versus evaluation	16	D;E	Identify any differences between the development and evaluation data in healthcare setting, eligibility criteria, outcome, and predictors	80-82
Ethical approval	17	D;E	Name the institutional research board or ethics committee that approved the study and describe the participant-informed consent or the ethics committee waiver of informed consent	54
OPEN SCIENCE				
Funding	18a	D;E	Give the source of funding and the role of the funders for the present study	
Conflicts of interest	18b	D;E	Declare any conflicts of interest and financial disclosures for all authors	iv
Protocol	18c	D;E	Indicate where the study protocol can be accessed or state that a protocol was not prepared	Not published
Registration	18d	D;E	Provide registration information for the study, including register name and registration number, or state that the study was not registered	Not registered
Data sharing	18e	D;E	Provide details of the availability of the study data	64
Code sharing	18f	D;E	Provide details of the availability of the analytical code[4]	64
PATIENT & PUBLIC INVOLVEMENT				
Patient & Public Involvement	19	D;E	Provide details of any patient and public involvement during the design, conduct, reporting, interpretation, or dissemination of the study or state no involvement.	178-189
RESULTS				
Participants	20a	D;E	Describe the flow of participants through the study, including the number of participants with and without the outcome and, if applicable, a summary of the follow-up time. A diagram may be helpful.	80-83
	20b	D;E	Report the characteristics overall and, where applicable, for each data source or setting, including the key dates, key predictors (including demographics), treatments received, sample size, number of outcome events, follow-up time, and amount of missing data. A table may be helpful. Report any differences across key demographic groups.	80-83, 65-67
	20c	E	For model evaluation, show a comparison with the development data of the distribution of important predictors (demographics, predictors, and outcome).	80-83
Model development	21	D;E	Specify the number of participants and outcome events in each analysis (e.g., for model development, hyperparameter tuning, model evaluation)	80-83
Model specification	22	D	Provide details of the full prediction model (e.g., formula, code, object, application programming interface) to allow predictions in new individuals and to enable third-party evaluation and implementation, including any restrictions to access or re-use (e.g., freely available, proprietary)[5]	89-91, 163-164
Model performance	23a	D;E	Report model performance estimates with confidence intervals, including for any key subgroups (e.g., sociodemographic). Consider plots to aid presentation.	164-166
	23b	D;E	If examined, report results of any heterogeneity in model performance across clusters. See TRIPOD Cluster for additional details[1].	Not Examined
Model updating	24	E	Report the results from any model updating, including the updated model and subsequent performance	83-92
DISCUSSION				
Interpretation	25	D;E	Give an overall interpretation of the main results, including issues of fairness in the context of the objectives and previous studies	102-116
Limitations	26	D;E	Discuss any limitations of the study (such as a non-representative sample, sample size, overfitting, missing data) and their effects on any biases, statistical uncertainty, and generalizability	124-134
Usability of the model in the context of current care	27a	D	Describe how poor quality or unavailable input data (e.g., predictor values) should be assessed and handled when implementing the prediction model	116
	27b	D	Specify whether users will be required to interact in the handling of the input data or use of the model, and what level of expertise is required of users	117-122
	27c	D;E	Discuss any next steps for future research, with a specific view to applicability and generalizability of the model	134-138

From: Collins GS, Moons KGM, Dhiman P, et al. *BMJ* 2024;385:e078378. doi:10.1136/bmj-2023-078378

Appendix E: IRB Approval

MARYWOOD UNIVERSITY
INSTITUTIONAL REVIEW BOARD
Immaculata Hall, 2300 Adams Avenue, Scranton, PA 18509

DATE:	November 25, 2024
TO:	Remle Newton-Dame, MPH
FROM:	Marywood University Institutional Review Board
STUDY TITLE:	[2256928-1] *Predicting Primary Care Risk Among Adolescents in a Large Urban Safety Net System*
MUIRB #:	
SUBMISSION TYPE:	New Project
ACTION:	APPROVED
APPROVAL DATE:	November 25, 2024
CHECK IN DUE DATE:	November 25, 2025
REVIEW TYPE:	Expedited Review
EXPEDITED REVIEW TYPE:	45 CFR 46.110 (b)(1)()() (Delete If Not Applicable)

Dear Mr./Mrs./Ms./Dr.:

PLEASE READ THIS LETTER CAREFULLY IN ITS ENTIRETY.
IT CONTAINS IMPORTANT INFORMATION ABOUT YOUR RESEARCH PROPOSAL AND YOUR RESPONSIBILITIES AS AN INVESTIGATOR. THE IRB IS REQUIRED BY FEDERAL LAW TO REPORT ALL SERIOUS OR CONTINUING NONCOMPLIANCE WITH THESE REQUIREMENTS TO FEDERAL AGENCIES.

Thank you for your submission of New Project materials for this research study. Marywood University's Institutional Review Board has **APPROVED** your submission. This approval is based on an appropriate risk/benefit ratio and a study design wherein the risks have been minimized. All research must be conducted in accordance with this approved submission.

Please remember that informed consent is a process beginning with a complete description of the study and assurance of subject understanding, followed by a signed informed consent and/or assent form, unless a waiver of documentation of informed consent was granted. Informed consent must continue

throughout the study via a dialogue between the investigator and each participating research subject. Federal regulations require each subject to receive a written copy of the consent document, unless waived.

We have applied the IRB's approval stamp to the following documents, which we have uploaded with this letter in IRBNet. The stamp must appear on versions shared with subjects whenever possible. If it is not feasible to use the stamped versions online (e.g., some email systems or survey platforms), please ensure that the language in the transmitted versions is identical to the stamped versions.

1. Informed Consent Form/Parental Permission Form/Assent Form or Script
2. Advertisement

Please also note that:

- **CLOSURE REPORTING:** Upon completion of the research, you must file a closure report form via IRBNet.
- **CHECK IN REPORTING OR CONTINUING REVIEW:** If activities will continue beyond your approval's one-year anniversary of _11/25/25_____, file a check in (most expedited) or a continuing review form (most full) by or before that date. For continuing review, submit with adequate review time, as lapses are not allowed.
- **RECORDS RETENTION:** You must retain records for a minimum of three years after the official closure date in IRBNet.
- **DEVIATION, UNANTICIPATED PROBLEM OR SERIOUS ADVERSE EVENT REPORTING:** If any of these events occur, you must file the appropriate form immediately via IRBNet.
- **REVISION REQUESTS:** If you decide to make procedural or document changes to your approved project, you must file a revision request form for review and approval prior to implementation, except when necessary to eliminate apparent, immediate hazards to the subjects. In hazardous situations, you must file the form immediately afterward.

The appropriate forms for any of the reports mentioned above may be found on the IRB's website or IRBNet's Forms Library. The library appears after you begin a follow-up package within your existing project and then click the *Designer* button on the left menu, followed by the blue "Need forms" link on the main screen (opens library under Step 1).

If you have any questions, please contact the Research Office at 570-348-6211, x.2418 or irbhelp@marywood.edu.

Please include your study title and IRBNet ID number in all correspondence with this office.

Thank you and good luck with your research!

Regards,
Institutional Review Board

Appendix F: Institutional Support Letter

NYC HEALTH+ HOSPITALS

Office of Population Health
Central Office
50 Water Street, 6th Floor
New York, NY 10004

October 28, 2024

Re: Predicting Primary Care Risk Among Adolescents in a Large Urban Safety Net System

Dear Institutional Review Board,

This letter confirms that as an authorized representative of NYC Health + Hospitals' Office of Population Health, I am aware of Remle Newton-Dame's research project and protocol. I will allow the investigator to access data to develop an evidence-based risk tiering algorithm for adolescent patients at NYC Health+Hospitals (H+H). I understand that this project will involve a secondary data analysis of data from NYC Health+Hospitals pediatric primary care sites, from 2022-2023. All data for this study will remain on our internal servers, and no patients will be contacted as a result of this study. I understand that activities for this dissertation may only commence after final approval from Marywood University's IRB.

I fully support this research study occurring in our office. This study addresses both a clinical need at NYC Health+Hospitals and a gap in the literature. Please do not hesitate to contact me if there are any questions regarding our cooperation with this project at Nichola.davis@nychhc.org .

Sincerely,

Nichola Davis, M.D., M.S.
Vice President and Chief Population Health Officer
Office of Ambulatory Care and Population Health
New York City Health + Hospitals
Clinical Professor
Department of Population Health, NYU School of Medicine
Email: nichola.davis@nychhc.org
Office: (646) 694-7123
Mobile: (347) 622-0469
bit.ly/PopHealthNYCHH

References

2022 National Survey on LGBTQ Youth Mental Health. (n.d.). The Trevor Project. Retrieved August 16, 2024, from https://www.thetrevorproject.org/

ACE Aware. (2024). *ACEs Aware May 2024 data report: More than 1.6 million Medi-Cal members screened for ACEs* [Data Brief]. California Department of Healthcare Services. https://www.acesaware.org/blog/aces-aware-may-2024-data-report-more-than-1-6-million-medi-cal-members-screened-for-aces/

Adus, S., Macklin, J., & Pinto, A. (2023). Exploring patient perspectives on how they can and should be engaged in the development of artificial intelligence (AI) applications in health care. *BMC Health Services Research*, *23*, 1163. https://doi.org/10.1186/s12913-023-10098-2

Agency for Healthcare Research and Quality. (2024). *How community engagement can support whole-person primary care* (AHRQ Topic Briefs). https://integrationacademy.ahrq.gov/products/topic-briefs/oud-low-threshold-treatment

Agency for Toxic Substances and Disease Registry. (2022, October 28). *CDC SVI documentation 2020 | Place and Health*. https://www.atsdr.cdc.gov/placeandhealth/svi/documentation/SVI_documentation_2020.html

Akoglu, H. (2018). User's guide to correlation coefficients. *Turkish Journal of Emergency Medicine*, *18*(3), 91–93. https://doi.org/10.1016/j.tjem.2018.08.001

Alderman, E. M., Breuner, C. C., Committee on Adolescence, Grubb, L. K., Powers, M. E., Upadhya, K., & Wallace, S. B. (2019). Unique needs of the adolescent. *Pediatrics*, *144*(6), e20193150. https://doi.org/10.1542/peds.2019-3150

Alderwick, H., Hood-Ronick, C. M., & Gottlieb, L. M. (2019). Medicaid investments to address social needs in Oregon and California. *Health Affairs*, *38*(5), 774–781. https://doi.org/10.1377/hlthaff.2018.05171

Aldo Crossa, M., Baquero, M., Etheredge, A. J., Seidl, L., Nieves, C., Dannefer, R., Solomon, E., Prasad, D., Jasek, J., Dongchung, T. Y., Marder, T., Deng, W. Q., Seligson, A. L., & Dumas, S. E. (2021). *Food insecurity and access in New York City during the COVID-19 pandemic, 2020-2021* (Epi Data Brief 128). New York City Department of Health and Mental Hygiene. https://www1.nyc.gov/assets/doh/downloads/pdf/epi/databrief128.pdf

Allen, M., & Donkin, A. (2015). *The impact of adverse experiences in the home on children and young people* (No. 9). Institute for Health Equity. https://www.instituteofhealthequity.org/resources-reports/the-impact-of-adverse-experiences-in-the-home-on-children-and-young-people

Allred, K. (2024). Saving lives: Suicide prevention in LGBTQ youth. *Journal of Psychosocial Nursing & Mental Health Services*, *62*(4), 6–8. https://doi.org/10.3928/02793695-20240307-01

Almeida, I. L. de L., Rego, J. F., Teixeira, A. C. G., & Moreira, M. R. (2021). Social isolation and its impact on child and adolescent development: A systematic review. *Revista Paulista de Pediatria*, *40*, e2020385. https://doi.org/10.1590/1984-0462/2022/40/2020385

Alowais, S. A., Alghamdi, S. S., Alsuhebany, N., Alqahtani, T., Alshaya, A. I., Almohareb, S. N., Aldairem, A., Alrashed, M., Bin Saleh, K., Badreldin, H. A., Al Yami, M. S., Al Harbi, S., & Albekairy, A. M. (2023). Revolutionizing healthcare: The role of artificial intelligence in clinical practice. *BMC Medical Education*, *23*(1), 689. https://doi.org/10.1186/s12909-023-04698-z

American Academy of Pediatrics. (n.d.). *HIPAA and FERPA basics*. Retrieved August 16, 2024, from https://www.aap.org/en/patient-care/school-health/hipaa-and-ferpa-basics/

American Journal of Preventive Medicine. (n.d.). *Explore Adverse Childhood Experiences in New York | AHR*. Retrieved August 12, 2024, from

https://www.americashealthrankings.org/explore/measures/ACEs_8/NY
American Lung Association. (n.d.). *Tobacco use among children and teens*. Retrieved December 17, 2023, from https://www.lung.org/quit-smoking/smoking-facts/tobacco-use-among-children
Andermann, A. (2016). Taking action on the social determinants of health in clinical practice: A framework for health professionals. *CMAJ*, *188*(17–18), E474–E483. https://doi.org/10.1503/cmaj.160177
Anderson, K. N., Swedo, E. A., Trinh, E., Ray, C. M., Krause, K. H., Verlenden, J. V., Clayton, H. B., Villaveces, A., Massetti, G. M., & Holditch Niolon, P. (2022). Adverse childhood experiences during the COVID-19 pandemic and associations with poor mental health and suicidal behaviors among high school students—Adolescent behaviors and experiences survey, United States, January-June 2021. *MMWR. Morbidity and Mortality Weekly Report*, *71*(41), 1301–1305. https://doi.org/10.15585/mmwr.mm7141a2
Andridge, R. R., & Little, R. J. A. (2010). A review of hot deck imputation for survey non-response. *International Statistical Review = Revue Internationale de Statistique*, *78*(1), 40–64. https://doi.org/10.1111/j.1751-5823.2010.00103.x
Apathy, N. (2025, June 8). *Design & data considerations for retrospective AI evaluations*. Academy Health Annual Research Meeting.
Aribot, M., Arriaga, A., Auyeung, A., Bannerman, C., Bushman-Copp, L., Cao, T., Cargill, C., Castillo, D., Cha, E., & Collyer, S. (2022). *The state of poverty and disadvantage in New York City*. Robin Hood and Center on Poverty & Social Policy.
Atreya, M. R., & Wong, H. R. (2019). Precision medicine in pediatric sepsis. *Current Opinion in Pediatrics*, *31*(3), 322–327. https://doi.org/10.1097/MOP.0000000000000753
Azur, M. J., Stuart, E. A., Frangakis, C., & Leaf, P. J. (2011). Multiple imputation by chained equations: What is it and how does it work? *International Journal of Methods in Psychiatric Research*, *20*(1), 40–49. https://doi.org/10.1002/mpr.329
Babenko, O., Kovalchuk, I., & Metz, G. A. S. (2015). Stress-induced perinatal and transgenerational epigenetic programming of brain development and mental health. *Neuroscience & Biobehavioral Reviews*, *48*, 70–91. https://doi.org/10.1016/j.neubiorev.2014.11.013
Bailit, M., Phillips, K., & Long, A. (2010). *Paying for the medical home: Payment models to support patient-centered medical home transformation in the safety net* (Safety Net Medical Home Initiative). Bailit Health Purchasing and Qualis Health. https://www.safetynetmedicalhome.org/sites/default/files/Policy-Brief-1.pdf
Baldwin, J. R., Caspi, A., Meehan, A. J., Ambler, A., Arseneault, L., Fisher, H. L., Harrington, H., Matthews, T., Odgers, C. L., Poulton, R., Ramrakha, S., Moffitt, T. E., & Danese, A. (2021). Population vs individual prediction of poor health from results of adverse childhood experiences screening. *JAMA Pediatrics*, *175*(4), 385–393. https://doi.org/10.1001/jamapediatrics.2020.5602
Balistreri, K. S., & Alvira-Hammond, M. (2016). Adverse childhood experiences, family functioning and adolescent health and emotional well-being. *Public Health*, *132*, 72–78.
Balocchini, E., Chiamenti, G., & Lamborghini, A. (2013). Adolescents: Which risks for their life and health? *Journal of Preventive Medicine and Hygiene*, *54*(4), 191–194.
Barnes, A. J., Anthony, B. J., Karatekin, C., Lingras, K. A., Mercado, R., & Thompson, L. A. (2020). Identifying adverse childhood experiences in pediatrics to prevent chronic health conditions. *Pediatric Research*, *87*(2), 362–370. https://doi.org/10.1038/s41390-019-0613-3
Baron-Lee, J., Bonner, B., Knapp, C., Bright, M., & Hinojosa, M. (2015). Factors associated with having a medical home for children at-risk of experiencing negative events: Results from a national study. *Maternal and Child Health Journal*, *19*(10), 2233–2242. https://doi.org/10.1007/s10995-015-1742-x

Bechtel, L., Gonzalez, Y., & Thompson, K. J. (n.d.). *Hot deck imputation: How many times can or should a donor be used?*

Bedoya, A. D., Economou-Zavlanos, N. J., Goldstein, B. A., Young, A., Jelovsek, J. E., O'Brien, C., Parrish, A. B., Elengold, S., Lytle, K., Balu, S., Huang, E., Poon, E. G., & Pencina, M. J. (2022). A framework for the oversight and local deployment of safe and high-quality prediction models. *Journal of the American Medical Informatics Association : JAMIA*, *29*(9), 1631–1636. https://doi.org/10.1093/jamia/ocac078

Bellis, M. A., Hughes, K., Ford, K., Hardcastle, K. A., Sharp, C. A., Wood, S., Homolova, L., & Davies, A. (2018). Adverse childhood experiences and sources of childhood resilience: A retrospective study of their combined relationships with child health and educational attendance. *BMC Public Health*, *18*(1), 792. https://doi.org/10.1186/s12889-018-5699-8

Bellis, M. A., Hughes, K., Ford, K., Ramos Rodriguez, G., Sethi, D., & Passmore, J. (2019). Life course health consequences and associated annual costs of adverse childhood experiences across Europe and North America: A systematic review and meta-analysis. *The Lancet. Public Health*, *4*(10), e517–e528. https://doi.org/10.1016/S2468-2667(19)30145-8

Benner, A. (2017). The toll of racial/ethnic discrimination on adolescents' adjustment. *Child Development Perspectives*, *11*(4), 251–256. https://doi.org/10.1111/cdep.12241

Beretta, L., & Santaniello, A. (2016). Nearest neighbor imputation algorithms: A critical evaluation. *BMC Medical Informatics and Decision Making*, *16*(3), 74. https://doi.org/10.1186/s12911-016-0318-z

Berg, K. L., Shiu, C.-S., Feinstein, R. T., Msall, M. E., & Acharya, K. (2018). Adverse childhood experiences are associated with unmet healthcare needs among children with autism spectrum disorder. *The Journal of Pediatrics*, *202*, 258–264.e1. https://doi.org/10.1016/j.jpeds.2018.07.021

Bethell, C. D., Carle, A., Hudziak, J., Gombojav, N., Powers, K., Wade, R., & Braveman, P. (2017). Methods to assess adverse childhood experiences of children and families: Toward approaches to promote child well-being in policy and practice. *Academic Pediatrics*, *17*(7S), S51–S69. https://doi.org/10.1016/j.acap.2017.04.161

Bethell, C., Blackwell, C. K., Gombojav, N., Davis, M. B., Bruner, C., & Garner, A. S. (2022). Toward measurement for a whole child health policy: Validity and national and state prevalence of the integrated child risk index. *Academic Pediatrics*, *22*(6), 952–964. https://doi.org/10.1016/j.acap.2021.12.001

Bethell, C., Jones, J., Gombojav, N., Linkenbach, J., & Sege, R. (2019). Positive childhood experiences and adult mental and relational health in a statewide sample: Associations across adverse childhood experiences levels. *JAMA Pediatrics*, *173*(11), e193007. https://doi.org/10.1001/jamapediatrics.2019.3007

Bharadwaj, P., Nicola, L., Breau-Brunel, M., Sensini, F., Tanova-Yotova, N., Atanasov, P., Lobig, F., & Blankenburg, M. (2024). Unlocking the value: Quantifying the return on investment of hospital artificial intelligence. *Journal of the American College of Radiology*, *21*(10), 1677–1685. https://doi.org/10.1016/j.jacr.2024.02.034

Bhattarai, A., Dimitropoulos, G., Marriott, B., Paget, J., Bulloch, A. G. M., Tough, S. C., & Patten, S. B. (2021). Can the adverse childhood experiences (ACEs) checklist be utilized to predict emergency department visits among children and adolescents? *BMC Medical Research Methodology*, *21*(1), 195. https://doi.org/10.1186/s12874-021-01392-w

Blakemore, S.-J., & Mills, K. L. (2014). Is adolescence a sensitive period for sociocultural processing? *Annual Review of Psychology*, *65*, 187–207. https://doi.org/10.1146/annurev-psych-010213-115202

Blum, R. W., Li, M., & Naranjo-Rivera, G. (2019). Measuring adverse child experiences among young adolescents globally: Relationships with depressive symptoms and violence perpetration. *The Journal of Adolescent Health: Official Publication of the Society for*

Adolescent Medicine, *65*(1), 86–93. https://doi.org/10.1016/j.jadohealth.2019.01.020

Boch, S. J., Warren, B. J., & Ford, J. L. (2019). Attention, externalizing, and internalizing problems of youth exposed to parental incarceration. *Issues in Mental Health Nursing*, *40*(6), 466–475. https://doi.org/10.1080/01612840.2019.1565872

Bomysoad, R. N., & Francis, L. A. (2020). Adverse childhood experiences and mental health conditions among adolescents. *The Journal of Adolescent Health: Official Publication of the Society for Adolescent Medicine*, *67*(6), 868–870. https://doi.org/10.1016/j.jadohealth.2020.04.013

Boullier, M., & Blair, M. (2018). Adverse childhood experiences. *Paediatrics and Child Health*, *28*(3), 132–137. https://doi.org/10.1016/j.paed.2017.12.008

Bradbury, K., Janicke, D. M., Riley, A. W., & Finney, J. W. (1999). Predictors of unintentional injuries to school-age children seen in pediatric primary care. *Journal of Pediatric Psychology*, *24*(5), 423–433. https://doi.org/10.1093/jpepsy/24.5.423

Brautsch, L. A. S., Lund, L., Andersen, M. M., Jennum, P. J., Folker, A. P., & Andersen, S. (2023). Digital media use and sleep in late adolescence and young adulthood: A systematic review. *Sleep Medicine Reviews*, *68*, 101742. https://doi.org/10.1016/j.smrv.2022.101742

Breehl, L., & Caban, O. (2024). Physiology, puberty. In *StatPearls*. StatPearls Publishing. http://www.ncbi.nlm.nih.gov/books/NBK534827/

Bright Futures/American Academy of Pediatrics. (2023, December). *Recommendations for preventive pediatric health care*. https://downloads.aap.org/AAP/PDF/periodicity_schedule.pdf

Brindle, P. M., McConnachie, A., Upton, M. N., Hart, C. L., Smith, G. D., & Watt, G. C. (2005). The accuracy of the Framingham risk-score in different socioeconomic groups: A prospective study. *The British Journal of General Practice*, *55*(520), 838–845.

Brockie, T. N., Dana-Sacco, G., Wallen, G. R., Wilcox, H. C., & Campbell, J. C. (2015). The relationship of adverse childhood experiences to PTSD, depression, poly-drug use and suicide attempt in reservation-based Native American adolescents and young adults. *American Journal of Community Psychology*, *55*(3–4), 411–421. https://doi.org/10.1007/s10464-015-9721-3

Brown, M. J., Masho, S. W., Perera, R. A., Mezuk, B., & Cohen, S. A. (2015). Sex and sexual orientation disparities in adverse childhood experiences and early age at sexual debut in the United States: Results from a nationally representative sample. *Child Abuse & Neglect*, *46*, 89–102. https://doi.org/10.1016/j.chiabu.2015.02.019

Bucci, M., Marques, S. S., Oh, D., & Harris, N. B. (2016). Toxic stress in children and adolescents. *Advances in Pediatrics*, *63*(1), 403–428. https://doi.org/10.1016/j.yapd.2016.04.002

Burke Harris, N. (2018). *Deepest well: Healing the long-term effects of childhood trauma and adversity*. Mariner Books.

Burke, N. J., Hellman, J. L., Scott, B. G., Weems, C. F., & Carrion, V. G. (2011). The impact of adverse childhood experiences on an urban pediatric population. *Child Abuse & Neglect*, *35*(6), 408–413. https://doi.org/10.1016/j.chiabu.2011.02.006

Burkhart, K., Asogwa, K., Muzaffar, N., & Gabriel, M. (2020). Pediatric integrated care models: A systematic review. *Clinical Pediatrics*, *59*(2), 148–153. https://doi.org/10.1177/0009922819890004

Byrne, B. M. (2013). *Structural equation modeling with AMOS: Basic concepts, applications, and programming* (2nd ed). Routledge. https://doi.org/10.4324/9780203805534

Calvo-Friedman, A., Clapp, J., Kumar, S. L., Adams, A., Gogia, K., & Davis, N. J. (2023). Scaling primary care social needs screening and referrals in New York City's safety-net health system. *NEJM Catalyst*, *4*(7), CAT.22.0371. https://doi.org/10.1056/CAT.22.0371

Cambridge Dictionary. (2023, December 6). *Diagnosis*.

Carrasquillo, O. (2013). Health care utilization. In M. D. Gellman & J. R. Turner (Eds.), *Encyclopedia of behavioral medicine* (pp. 909–910). Springer. https://doi.org/10.1007/978-1-4419-1005-9_885

Castensøe-Seidenfaden, P., Jensen, A. K., Smedegaard, H., Hommel, E., Husted, G. R., Pedersen-Bjergaard, U., & Teilmann, G. (2017). Clinical, behavioural and social indicators for poor glycaemic control around the time of transfer to adult care: A longitudinal study of 126 young people with diabetes. *Diabetic Medicine: A Journal of the British Diabetic Association*, *34*(5), 667–675. https://doi.org/10.1111/dme.13318

Centers for Disease Control and Prevention (CDC). (2021, April 2). *Adverse childhood experiences (ACEs)*. https://www.cdc.gov/violenceprevention/aces/index.html

Centers for Disease Control and Prevention (CDC). (2023, June 29). *Behavioral Risk Factor Surveillance System ACE Data*. https://www.cdc.gov/violenceprevention/aces/ace-brfss.html.

Centers for Disease Control and Prevention (CDC). (2024, April 30). *FastStats: Adolescent health*. https://www.cdc.gov/nchs/fastats/adolescent-health.htm

Centers for Disease Control and Prevention (CDC). (2021, September 30). *Social determinants of health*. https://www.cdc.gov/socialdeterminants/index.htm

Centers for Disease Control and Prevention (CDC). (2024, November 12). *Tables from STI surveillance, 2023 | STI statistics*. https://www.cdc.gov/sti-statistics/data-vis/index.html

Child and Adolescent Mental Health Coalition. (2023, April 14). *Child and adolescent mental and behavioral health principles*. https://static1.squarespace.com/static/621e47eba8741d3518d1a3cd/t/6439b56ab09c41722901299d/1681503594954/2023+CAMH+Principles.pdf

Children's Defense Fund. (n.d.). *Child population*. Retrieved October 24, 2022, from https://data.cccnewyork.org/data/map/98/child-population#10/13/3/16/62/a/a

Chin, M. H., Afsar-Manesh, N., Bierman, A. S., Chang, C., Colón-Rodríguez, C. J., Dullabh, P., Duran, D. G., Fair, M., Hernandez-Boussard, T., Hightower, M., Jain, A., Jordan, W. B., Konya, S., Moore, R. H., Moore, T. T., Rodriguez, R., Shaheen, G., Snyder, L. P., Srinivasan, M., & Ohno-Machado, L. (2023). Guiding principles to address the impact of algorithm bias on racial and ethnic disparities in health and health care. *JAMA Network Open*, *6*(12), e2345050. https://doi.org/10.1001/jamanetworkopen.2023.45050

Clapp, J., Calvo-Friedman, A., Cameron, S., Kramer, N., Kumar, S. L., Foote, E., Lupi, J., Osuntuyi, O., & Chokshi, D. A. (2020). The COVID-19 shadow pandemic: Meeting social needs for a city in lockdown. *Health Affairs (Project Hope)*, *39*(9), 1592–1596. https://doi.org/10.1377/hlthaff.2020.00928

Coalition for Health AI. (2024). *CHAI Responsible AI guide (Blueprint for trustworthy AI)*. https://assets.ctfassets.net/7s4afyr9pmov/6e7PrdrsNTQ5FjZ4uyRjTW/c4070131c523d4e1db26105aa51f087d/CHAI_Responsible-AI-Guide.pdf

Collins, G. S., Moons, K. G. M., Dhiman, P., Riley, R. D., Beam, A. L., Calster, B. V., Ghassemi, M., Liu, X., Reitsma, J. B., van Smeden, M., Boulesteix, A.-L., Camaradou, J. C., Celi, L. A., Denaxas, S., Denniston, A. K., Glocker, B., Golub, R. M., Harvey, H., Heinze, G., & Logullo, P. (2024). TRIPOD+AI statement: Updated guidance for reporting clinical prediction models that use regression or machine learning methods. *BMJ*, *385*, e078378. https://doi.org/10.1136/bmj-2023-078378

Collins, G. S., Reitsma, J. B., Altman, D. G., & Moons, K. G. M. (2015). Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): The TRIPOD statement. *Annals of Internal Medicine*, *162*(1), 55–63. https://doi.org/10.7326/M14-0697

Columbia Suicide Severity Rating Scale (C-SSRS). (2021, September 17). https://www.samhsa.gov/resource/dbhis/columbia-suicide-severity-rating-scale-c-ssrs

Committee on Adolescence, American Academy of Pediatrics. (2008). Achieving quality health

services for adolescents. *Pediatrics*, *121*(6), 1263–1270. https://doi.org/10.1542/peds.2008-0694

Committee on Educating Health Professionals to Address the Social Determinants of Health, Board on Global Health, Institute of Medicine, & National Academies of Sciences, Engineering, and Medicine. (2016). *A framework for educating health professionals to address the social determinants of health*. National Academies Press (US). http://www.ncbi.nlm.nih.gov/books/NBK395983/

Community Health Needs Assessment for Charitable Hospital organizations—Section 501(r)(3) | Internal Revenue Service. (n.d.). Retrieved May 29, 2022, from https://www.irs.gov/charities-non-profits/community-health-needs-assessment-for-charitable-hospital-organizations-section-501r3

Conger, R. D., Conger, K. J., & Martin, M. J. (2010). Socioeconomic status, family processes, and individual development. *Journal of Marriage and the Family*, *72*(3), 685–704. https://doi.org/10.1111/j.1741-3737.2010.00725.x

Cromhout, P. F., Thygesen, L. C., Moons, P., Nashef, S., Damgaard, S., & Berg, S. K. (2022). Social and emotional factors as predictors of poor outcomes following cardiac surgery. *Interactive Cardiovascular and Thoracic Surgery*, *34*(2), 193–200. https://doi.org/10.1093/icvts/ivab261

Danese, A., & McEwen, B. S. (2012). Adverse childhood experiences, allostasis, allostatic load, and age-related disease. *Physiology & Behavior*, *106*(1), 29–39. https://doi.org/10.1016/j.physbeh.2011.08.019

Daniel-Robinson, L., & Moore, J. E. (2019). *Innovation and opportunities to address social determinants of health in Medicaid managed care*. Institute for Medicaid Innovation. https://medicaidinnovation.org/wp-content/uploads/2022/11/2019-IMI-Social_Determinants_of_Health_in_Medicaid-Report-1.pdf

Declaration of Alma-Ata International Conference on Primary Health Care, Alma-Ata, USSR, 6–12 September 1978. (2004). *Development*, *47*(2), 159–161. https://doi.org/10.1057/palgrave.development.1100047

Delaplain, P. T., Guner, Y. S., Feaster, W., Wallace, E., Gibbs, D., Gholizadeh, M., Reyna, T., & Ehwerhemuepha, L. (2020). Prediction of 7-day readmission risk for pediatric trauma patients. *The Journal of Surgical Research*, *253*, 254–261. https://doi.org/10.1016/j.jss.2020.03.068

Dibben, G. O., Martin, A., Shore, C. B., Johnstone, A., McMellon, C., Palmer, V., Pugmire, J., Riddell, J., Skivington, K., Wells, V., McDaid, L., & Simpson, S. A. (2023). Adolescents' interactive electronic device use, sleep and mental health: A systematic review of prospective studies. *Journal of Sleep Research*, *32*(5), e13899. https://doi.org/10.1111/jsr.13899

Dmochowski, R. R. (2021). Foundational statistical principles in medical research: Sensitivity, specificity, positive predictive value, and negative predictive value. *Medicina*, *57*(5), 503. https://doi.org/10.3390/medicina57050503

Donaldson, M. S., Yordy, K. D., Lohr, K. N., & Vanselow, N. A. (1996). Defining primary care. In *Primary care: America's health in a new era*. National Academies Press (US). https://www.ncbi.nlm.nih.gov/books/NBK232631/

Draxler, J. M., & Ruppar, T. M. (2022). Treatment adherence in adolescents with histories of adverse childhood experiences: A systematic review. *Journal of Psychosocial Nursing and Mental Health Services*, *60*(6), 11–18. https://doi.org/10.3928/02793695-20211118-04

Earnshaw, V. A. (2020). Stigma and substance use disorders: A clinical, research, and advocacy agenda. *The American Psychologist*, *75*(9), 1300–1311. https://doi.org/10.1037/amp0000744

Economou-Zavlanos, N. J., Bessias, S., Cary, M. P., Bedoya, A. D., Goldstein, B. A., Jelovsek,

J. E., O'Brien, C. L., Walden, N., Elmore, M., Parrish, A. B., Elengold, S., Lytle, K. S., Balu, S., Lipkin, M. E., Shariff, A. I., Gao, M., Leverenz, D., Henao, R., Ming, D. Y., & Poon, E. G. (2023). Translating ethical and quality principles for the effective, safe and fair development, deployment and use of artificial intelligence technologies in healthcare. *Journal of the American Medical Informatics Association : JAMIA*, *31*(3), 705–713. https://doi.org/10.1093/jamia/ocad221

Efthimiou, O., Hoogland, J., Debray, T. P. A., Seo, M., Furukawa, T. A., Egger, M., & White, I. R. (2023). Measuring the performance of prediction models to personalize treatment choice. *Statistics in Medicine*, *42*(8), 1188–1206. https://doi.org/10.1002/sim.9665

Ehrenstein, V., Hellfritzsch, M., Kahlert, J., Langan, S. M., Urushihara, H., Marinac-Dabic, D., Lund, J. L., Sørensen, H. T., & Benchimol, E. I. (2024). Validation of algorithms in studies based on routinely collected health data: General principles. *American Journal of Epidemiology*, *193*(11), 1612–1624. https://doi.org/10.1093/aje/kwae071

Elmore, A. L., & Crouch, E. (2020). The association of adverse childhood experiences with anxiety and depression for children and youth, 8 to 17 years of age. *Academic Pediatrics*, *20*(5), 600–608. https://doi.org/10.1016/j.acap.2020.02.012

Elmore, A. L., Crouch, E., & Kabir Chowdhury, M. A. (2020). The interaction of adverse childhood experiences and resiliency on the outcome of depression among children and youth, 8-17 year olds. *Child Abuse & Neglect*, *107*, 104616. https://doi.org/10.1016/j.chiabu.2020.104616

Environmental Justice NYC Initiative. (2024). *Environmental Justice New York City report*. Mayor's Office of Climate and Environmental Justice. https://climate.cityofnewyork.us/ejnyc-report/

Eum, Y., Yoo, E., & Bowen, E. (2019). Socioeconomic determinants of pediatric asthma emergency department visits under regional economic development in western New York. *Social Science & Medicine*, *222*, 133–144. https://doi.org/10.1016/j.socscimed.2019.01.001

Exworthy, M. (2008). The enduring legacy of Alma Ata: 30 years on. *London Journal of Primary Care*, *1*(2), 81–84.

Fairchild, A. L., Rosner, D., Colgrove, J., Bayer, R., & Fried, L. P. (2010). The exodus of public health: What history can tell us about the future. *American Journal of Public Health*, *100*(1), 54–63. https://doi.org/10.2105/AJPH.2009.163956

Felitti, V. J., Anda, R. F., Nordenberg, D., Williamson, D. F., Spitz, A. M., Edwards, V., Koss, M. P., & Marks, J. S. (1998). Relationship of childhood abuse and household dysfunction to many of the leading causes of death in adults: The Adverse Childhood Experiences (ACE) Study. *American Journal of Preventive Medicine*, *14*(4), 245–258. https://doi.org/10.1016/S0749-3797(98)00017-8

Finkelhor, D. (2018). Screening for adverse childhood experiences (ACEs): Cautions and suggestions. *Child Abuse & Neglect*, *85*, 174–179. https://doi.org/10.1016/j.chiabu.2017.07.016

Finkelhor, D., Shattuck, A., Turner, H., & Hamby, S. (2015). A revised inventory of adverse childhood experiences. *Child Abuse & Neglect*, *48*, 13–21. https://doi.org/10.1016/j.chiabu.2015.07.011

Fisher, C. B., Wallace, S. A., & Fenton, R. E. (2000). Discrimination distress during adolescence. *Journal of Youth and Adolescence*, *29*(6), 679–695. https://doi.org/10.1023/A:1026455906512

Fitch, J. (2024, February 28). The role of the pediatrician in school-based mental health services. *Contemporary Pediatrics*. https://www.contemporarypediatrics.com/view/the-role-of-the-pediatrician-in-school-based-mental-health-services

Folk, J. B., Kemp, K., Yurasek, A., Barr-Walker, J., & Tolou-Shams, M. (2021). Adverse

childhood experiences among justice-involved youth: Data-driven recommendations for action using the sequential intercept model. *The American Psychologist*, *76*(2), 268–283. https://doi.org/10.1037/amp0000769

Folk, J. B., Ramos, L. M. C., Bath, E. P., Rosen, B., Marshall, B. D. L., Kemp, K., Brown, L., Conrad, S., & Tolou-Shams, M. (2021). The prospective impact of adverse childhood experiences on justice-involved youth's psychiatric symptoms and substance use. *Journal of Consulting and Clinical Psychology*, *89*(6), 483–498. https://doi.org/10.1037/ccp0000655

Fontanarosa, J. B., & Dai, Y. (2011). Using LASSO regression to detect predictive aggregate effects in genetic studies. *BMC Proceedings*, *5*(9), S69. https://doi.org/10.1186/1753-6561-5-S9-S69

Frieden, T. R. (2010). A framework for public health action: The health impact pyramid. *American Journal of Public Health*, *100*(4), 590–595. https://doi.org/10.2105/AJPH.2009.185652

Fry, C. H., Heppleston, E., Fluck, D., & Han, T. S. (2020). Derivation of age-adjusted LACE index thresholds in the prediction of mortality and frequent hospital readmissions in adults. *Internal and Emergency Medicine*, *15*(7), 1319–1325. https://doi.org/10.1007/s11739-020-02448-3

Garner, A., Yogman, M., & Committee on Psychosocial Aspects of Child and Family Health, Section on Developmental and Behavioral Pediatrics, Council on Early Childhood. (2021). Preventing childhood toxic stress: Partnering with families and communities to promote relational health. *Pediatrics*, *148*(2), e2021052582. https://doi.org/10.1542/peds.2021-052582

Gerald, L. B., Anderson, A., Johnson, G. D., Hoff, C., & Trimm, R. F. (1994). Social class, social support and obesity risk in children. *Child: Care, Health and Development*, *20*(3), 145–163. https://doi.org/10.1111/j.1365-2214.1994.tb00377.x

Get Screened—MyHealthfinder | health.gov. (n.d.). Retrieved December 10, 2023, from https://health.gov/myhealthfinder/doctor-visits/screening-tests/get-screened

Gibbons, F. X., Roberts, M. E., Gerrard, M., Li, Z., Beach, S. R. H., Simons, R. L., Weng, C.-Y., & Philibert, R. A. (2012). The impact of stress on the life history strategies of African American adolescents: Cognitions, genetic moderation, and the role of discrimination. *Developmental Psychology*, *48*(3), 722–739. https://doi.org/10.1037/a0026599

Ginsburg, K. R. (n.d.). Chapter 32: The SSHADESS screening: A strength-based psychosocial assessment. In K.R. Ginsburg & S.B. Kinsman (Eds), *Reaching Teens: Strength-Based Communication Strategies to Build Resilience and Support Healthy Adolescent Development*. American Academy of Pediatrics. https://www.aap.org/contentassets/0e45de0366d54ec38fbfcb72382a0c6c/rt2e_ch32_saham.pdf.

Giovanelli, A., Mondi, C. F., Reynolds, A. J., & Ou, S.-R. (2020). Adverse childhood experiences: Mechanisms of risk and resilience in a longitudinal urban cohort. *Development and Psychopathology*, *32*(4), 1418–1439. https://doi.org/10.1017/S095457941900138X

Goldstein, B. A., Navar, A. M., Pencina, M. J., & Ioannidis, J. P. A. (2017). Opportunities and challenges in developing risk prediction models with electronic health records data: A systematic review. *Journal of the American Medical Informatics Association: JAMIA*, *24*(1), 198–208. https://doi.org/10.1093/jamia/ocw042

Gottlieb, L. M., Wing, H., & Adler, N. E. (2017). A systematic review of interventions on patients' social and economic needs. *American Journal of Preventive Medicine*, *53*(5), 719–729. https://doi.org/10.1016/j.amepre.2017.05.011

Graf, G. H.-J., Chihuri, S., Blow, M., & Li, G. (2021). Adverse childhood experiences and justice system contact: A systematic review. *Pediatrics*, *147*(1), e2020021030.

https://doi.org/10.1542/peds.2020-021030
Gunn, H. J., Rezvan, P. H., Fernández, M. I., & Comulada, W. S. (2023). How to apply variable selection machine learning algorithms with multiply imputed data: A missing discussion. *Psychological Methods*, *28*(2), 452–471. https://doi.org/10.1037/met0000478
Hair, J. F., Black, W. C., Babin, B. J., & Anderson, R. E. (2013). *Multivariate data analysis*. Pearson Education Limited.
Hall, A., Perez, A., West, X., Brown, M., Kim, E., Salih, Z., & Aronoff, S. (2021). The association of adverse childhood experiences and resilience with health outcomes in adolescents: An observational study. *Global Pediatric Health*, *8*, 2333794X20982433. https://doi.org/10.1177/2333794X20982433
Hare, A., Honig, K., Cronholm, P., Shabazz-McKlaine, S., & Morgan, A. (2023). Patient perspectives on technology-based approaches to social needs screening. *The American Journal of Managed Care*, *29*(SP9), SP725-SP727. https://www.ajmc.com/view/patient-perspectives-on-technology-based-approaches-to-social-needs-screening
Hastie, T. J., Tibshirani, R., & Friedman, J. H. (2009). *The elements of statistical learning: Data mining, inference, and prediction* (2nd ed). Springer.
Health Resources and Services Administration. (2020). *Adverse childhood experiences: NSCH data brief* (National Survey on Children's Health).
HealthIT.gov (n.d.) What are the differences between electronic medical records, electronic health records, and personal health records? Retrieved December 17, 2023, from https://www.healthit.gov/faq/what-are-differences-between-electronic-medical-records-electronic-health-records-and-personal
HealthySteps. (2022). *Tiers and core components*. https://www.healthysteps.org/what-we-do/our-model/tiers-and-core-components/
Healthy People 2020. (n.d.). *Social determinants of health.* Retrieved May 29, 2022, from https://www.healthypeople.gov/2020/topics-objectives/topic/social-determinants-health/interventions-resources
Healthy People 2030. (n.d.). *Social determinants of health.* Retrieved December 17, 2023, from https://health.gov/healthypeople/priority-areas/social-determinants-health
Healthy Teen Network. (n.d.). *Youth 360°: How & where youth live, learn, & play matters* [Presentation]. Retrieved November 5, 2022, from https://www.healthyteennetwork.org/resources/youth360-presentation/
Herzog, J. I., & Schmahl, C. (2018). Adverse childhood experiences and the consequences on neurobiological, psychosocial, and somatic conditions across the lifespan. *Frontiers in Psychiatry*, *9*. https://doi.org/10.3389/fpsyt.2018.00420
Hillis, S. D., Anda, R. F., Felitti, V. J., & Marchbanks, P. A. (2001). Adverse childhood experiences and sexual risk behaviors in women: A retrospective cohort study. *Family Planning Perspectives*, *33*(5), 206–211.
Hirshon, J. M., Risko, N., Calvello, E. J., Stewart de Ramirez, S., Narayan, M., Theodosis, C., & O'Neill, J. (2013). Health systems and services: The role of acute care. *Bulletin of the World Health Organization*, *91*(5), 386–388. https://doi.org/10.2471/BLT.12.112664
Horwitz, L. I., & Krelle, H. A. (2023). Using rapid randomized trials to improve health care systems. *Annual Review of Public Health*, *44*, 445–457. https://doi.org/10.1146/annurev-publhealth-071521-025758
Huang, J., Galal, G., Etemadi, M., & Vaidyanathan, M. (2022). Evaluation and mitigation of racial bias in clinical machine learning models: Scoping review. *JMIR Medical Informatics*, *10*(5), e36388. https://doi.org/10.2196/36388
Hughes, K., Bellis, M. A., Hardcastle, K. A., Sethi, D., Butchart, A., Mikton, C., Jones, L., & Dunne, M. P. (2017). The effect of multiple adverse childhood experiences on health: A systematic review and meta-analysis. *The Lancet. Public Health*, *2*(8), e356–e366.

https://doi.org/10.1016/S2468-2667(17)30118-4
Hunt, A. (2021). Development of a screening tool for social determinants of health at a Federally Qualified Health Center. *Pediatrics*, *147*(3_MeetingAbstract), 40. https://doi.org/10.1542/peds.147.3MA1.40
Ignoffo, S., Gu, S., Ellyin, A., & Benjamins, M. R. (2024). A review of community health worker integration in health departments. *Journal of Community Health*, *49*(2), 366–376. https://doi.org/10.1007/s10900-023-01286-6
Indiana Department of Health. (2021, December). *Adverse childhood experiences (ACEs) and adult smoking*. https://www.in.gov/health/tpc/files/ACEs-and-Current-Smoking_2021.pdf
Institute of Medicine (US), Board on Population Health and Public Health Practice, & Committee on the Recommended Social and Behavioral Domains and Measures for Electronic Health Records. (2015). *Capturing social and behavioral domains and measures in electronic health records: Phase 2*. National Academies Press.
Institute of Medicine (US) Committee on Integrating Behavioral and Social Sciences Research to Improve the Public's Health, Smedley, B. D., & Syme, S. L. (2000). Preadolescent and adolescent influences on health. In *Promoting health: Intervention strategies from social and behavioral research*. National Academies Press (US). https://www.ncbi.nlm.nih.gov/books/NBK222829/
Institute of Medicine (US) and National Research Council (US) Committee on the Science of Adolescence. (2011). Adolescents and the risks that affect them. In *The science of adolescent risk-taking: Workshop report*. National Academies Press (US). https://www.ncbi.nlm.nih.gov/books/NBK53412/
Jackson, S. L. (2018). Hypertension among youths—United States, 2001–2016. *MMWR. Morbidity and Mortality Weekly Report*, *67*. https://doi.org/10.15585/mmwr.mm6727a2
James, C., Jimenez, M. E., Wade, R., & Nepomnyaschy, L. (2021). Adverse childhood experiences and teen behavior outcomes: The role of disability. *Academic Pediatrics*, *21*(8), 1395–1403. https://doi.org/10.1016/j.acap.2021.05.006
Jaworska, N., & MacQueen, G. (2015). Adolescence as a unique developmental period. *Journal of Psychiatry & Neuroscience : JPN*, *40*(5), 291–293. https://doi.org/10.1503/jpn.150268
Jbaily, A., Zhou, X., Liu, J., Lee, T.-H., Kamareddine, L., Verguet, S., & Dominici, F. (2022). Air pollution exposure disparities across US population and income groups. *Nature*, *601*(7892), 228–233. https://doi.org/10.1038/s41586-021-04190-y
Ji, D., Goudy, S. L., Raval, M. V., & Raol, N. (2019). Pediatric surgical risk assessment tools: A systematic review. *The Journal of Surgical Research*, *234*, 277–282. https://doi.org/10.1016/j.jss.2018.09.051
Johnson, M. K., & Johnson, K. (n.d.). *8.4 encoding missingness | Feature engineering and selection: A practical approach for predictive models*. Retrieved February 28, 2025, from https://bookdown.org/max/FES/encoding-missingness.html
Johnson, T. L., Brewer, D., Estacio, R., Vlasimsky, T., Durfee, M. J., Thompson, K. R., Everhart, R. M., Rinehart, D. J., & Batal, H. (2015). Augmenting predictive modeling tools with clinical insights for care coordination program design and implementation. *EGEMS (Washington, DC)*, *3*(1), 1181. https://doi.org/10.13063/2327-9214.1181
Kalmakis, K. A., & Chandler, G. E. (2015). Health consequences of adverse childhood experiences: A systematic review. *Journal of the American Association of Nurse Practitioners*, *27*(8), 457. https://doi.org/10.1002/2327-6924.12215
Kang, H. (2013). The prevention and handling of the missing data. *Korean Journal of Anesthesiology*, *64*(5), 402–406. https://doi.org/10.4097/kjae.2013.64.5.402
Kent, D. M. (2020). Predictably unequal: Understanding and addressing concerns that algorithmic clinical prediction may increase health disparities. *NPJ Digital Medicine*, *3*, 99. https://doi.org/10.1038/s41746-020-0304-9
Kerker, B. D., Zhang, J., Nadeem, E., Stein, R. E. K., Hurlburt, M. S., Heneghan, A., Landsverk,

J., & McCue Horwitz, S. (2015). Adverse childhood experiences and mental health, chronic medical conditions, and development in young children. *Academic Pediatrics*, *15*(5), 510–517. https://doi.org/10.1016/j.acap.2015.05.005

Khalaf, A. M., Alubied, A. A., Khalaf, A. M., & Rifaey, A. A. (n.d.). The impact of social media on the mental health of adolescents and young adults: A systematic review. *Cureus*, *15*(8), e42990. https://doi.org/10.7759/cureus.42990

Kim, J. H. J., Tsai, W., Kodish, T., Trung, L. T., Lau, A. S., & Weiss, B. (2019). Cultural variation in temporal associations among somatic complaints, anxiety, and depressive symptoms in adolescence. *Journal of Psychosomatic Research*, *124*, 109763. https://doi.org/10.1016/j.jpsychores.2019.109763

King, C. A., Grupp-Phelan, J., Brent, D., Dean, J. M., Webb, M., Bridge, J. A., Spirito, A., Chernick, L. S., Mahabee-Gittens, E. M., Mistry, R. D., Rea, M., Keller, A., Rogers, A., Shenoi, R., Cwik, M., Busby, D. R., Casper, T. C., & Pediatric Emergency Care Applied Research Network (PECARN). (2019). Predicting 3-month risk for adolescent suicide attempts among pediatric emergency department patients. *Journal of Child Psychology and Psychiatry and Allied Disciplines*, *60*(10), 1055–1064. https://doi.org/10.1111/jcpp.13087

Klein, D. A., Paradise, S. L., & Landis, C. A. (2020). Screening and counseling adolescents and young adults: A framework for comprehensive care. *American Family Physician*, *101*(3), 147–158.

Knight, J. R., Shrier, L. A., Bravender, T. D., Farrell, M., Vander Bilt, J., & Shaffer, H. J. (1999). A new brief screen for adolescent substance abuse. *Archives of Pediatrics & Adolescent Medicine*, *153*(6), 591–596. https://doi.org/10.1001/archpedi.153.6.591

Koita, K., Long, D., Hessler, D., Benson, M., Daley, K., Bucci, M., Thakur, N., & Harris, N. B. (2018). Development and implementation of a pediatric adverse childhood experiences (ACEs) and other determinants of health questionnaire in the pediatric medical home: A pilot study. *PLOS ONE*, *13*(12), e0208088. https://doi.org/10.1371/journal.pone.0208088

Kroening, A. L. H., & Dawson-Hahn, E. (2019). Health considerations for immigrant and refugee children. *Advances in Pediatrics*, *66*, 87–110. https://doi.org/10.1016/j.yapd.2019.04.003

Krstajic, D., Buturovic, L. J., Leahy, D. E., & Thomas, S. (2014). Cross-validation pitfalls when selecting and assessing regression and classification models. *Journal of Cheminformatics*, *6*, 10. https://doi.org/10.1186/1758-2946-6-10

Kuppermann, N., Holmes, J. F., Dayan, P. S., Hoyle, J. D., Atabaki, S. M., Holubkov, R., Nadel, F. M., Monroe, D., Stanley, R. M., Borgialli, D. A., Badawy, M. K., Schunk, J. J., Quayle, K. S., Mahajan, P., Lichenstein, R., Lillis, K. A., Tunik, M. G., Jacobs, E. S., Callahan, J. M., & Pediatric Emergency Care Applied Research Network (PECARN). (2009). Identification of children at very low risk of clinically-important brain injuries after head trauma: A prospective cohort study. *Lancet (London, England)*, *374*(9696), 1160–1170. https://doi.org/10.1016/S0140-6736(09)61558-0

L, S., & K, K. (2023). *National Hospital Ambulatory Medical Care Survey: 2019 National Summary Tables*. National Center for Health Statistics (U.S.). https://doi.org/10.15620/cdc:123251

Lahey, J. (2021). *The addiction inoculation: Raising healthy kids in a culture of dependence* (Unabridged). HarperCollins.

Leary, J. C., Price, L. L., Scott, C. E. R., Kent, D., Wong, J. B., & Freund, K. M. (2019). Developing prediction models for 30-day unplanned readmission among children with medical complexity. *Hospital Pediatrics*, *9*(3), 201–208. https://doi.org/10.1542/hpeds.2018-0174

Lee, X., Griffin, A., Ragavan, M. I., & Patel, M. (2024). Adolescent mental health and academic performance: Determining evidence-based associations and informing approaches to support in educational settings. *Pediatric Research*, *95*(6), 1395–1397.

https://doi.org/10.1038/s41390-024-03098-3
Lefchak, B., Cushwa, A., Kersten, H., Courts, K., & McPeak, K. (2022). Characterization of social risk factors among newborns seen at an urban pediatric primary care predictive of appointment nonattendance during the first 6 months of life. *Health Equity*, *6*(1), 40–48. https://doi.org/10.1089/heq.2021.0053
Lent, D. G., & Hickson, A. (2023). *Displaced and disconnected: The experience of asylum seekers and migrants in New York City in 2023*. Make the Road NYC. https://maketheroadny.org/wp-content/uploads/2023/06/Displaced-and-Disconnected_MRNY-Report.pdf
Li, C., Parpia, C., Sriharan, A., & Keefe, D. T. (2022). Electronic medical record-related burnout in healthcare providers: A scoping review of outcomes and interventions. *BMJ Open*, *12*(8), e060865. https://doi.org/10.1136/bmjopen-2022-060865
Li, Z., Gogia, S., Tatem, K. S., Cooke, C., Singer, J., Chokshi, D. A., & Newton-Dame, R. (2023). Developing a model to predict high health care utilization among patients in a New York City safety net system. *Medical Care*, *61*(2), 102–108. https://doi.org/10.1097/MLR.0000000000001807
Lin, D., Li, X., Fang, X., & Lin, X. (2011). Childhood sexual abuse and sexual risks among young rural-to-urban migrant women in Beijing, China. *AIDS Care*, *23*(Suppl 1), 113–119. https://doi.org/10.1080/09540121.2010.534434
Liu, F. (2025, June 10). *Data-driven ethical AI for clinical risk predictive modeling*. Academy Health Annual Research Meeting, Minneapolis, MN.
Shilkofski, N. (2021). Addressing trauma and building resilience in children and families: Standardized patient cases for pediatric residents. *MedEdPORTAL: The Journal of Teaching and Learning Resources*, *17*, 11193. https://doi.org/10.15766/mep_2374-8265.11193
Loades, M. E., Chatburn, E., Higson-Sweeney, N., Reynolds, S., Shafran, R., Brigden, A., Linney, C., McManus, M. N., Borwick, C., & Crawley, E. (2020). Rapid systematic review: The impact of social isolation and loneliness on the mental health of children and adolescents in the context of COVID-19. *Journal of the American Academy of Child & Adolescent Psychiatry*, *59*(11), 1218–1239.e3. https://doi.org/10.1016/j.jaac.2020.05.009
Long, D., Hessler, D., Koita, K., Bucci, M., Benson, M., Gilgoff, R., Thakur, N., & Burke Harris, N. (2022). Screening for adverse childhood experiences in pediatrics: A randomized trial of aggregate-level versus item-level response screening formats. *PloS One*, *17*(12), e0273491. https://doi.org/10.1371/journal.pone.0273491
Loveday, S., Hall, T., Constable, L., Paton, K., Sanci, L., Goldfeld, S., & Hiscock, H. (2022). Screening for adverse childhood experiences in children: A systematic review. *Pediatrics*, *149*(2), e2021051884. https://doi.org/10.1542/peds.2021-051884
Luo, W., Phung, D., Tran, T., Gupta, S., Rana, S., Karmakar, C., Shilton, A., Yearwood, J., Dimitrova, N., Ho, T. B., Venkatesh, S., & Berk, M. (2016). Guidelines for developing and reporting machine learning predictive models in biomedical research: A multidisciplinary view. *Journal of Medical Internet Research*, *18*(12), e323. https://doi.org/10.2196/jmir.5870
Lupton-Smith, C., Stuart, E. A., McGinty, E. E., Dalcin, A. T., Jerome, G. J., Wang, N.-Y., & Daumit, G. L. (2022). Determining predictors of weight loss in a behavioral intervention: A case study in the use of Lasso Regression. *Frontiers in Psychiatry*, *12*. https://www.frontiersin.org/articles/10.3389/fpsyt.2021.707707
Mackin, S., Major, V. J., Chunara, R., & Newton-Dame, R. (2025). Identifying and mitigating algorithmic bias in the safety net. *Npj Digital Medicine*, *8*(1), 1–11. https://doi.org/10.1038/s41746-025-01732-w
Marmot, M. (2005). Social determinants of health inequalities. *Lancet (London, England)*, *365*(9464), 1099–1104. https://doi.org/10.1016/S0140-6736(05)71146-6

Mazurenko, O., Harle, C. A., Blackburn, J., Menachemi, N., Hirsh, A., Grannis, S., Boustani, M., Musey, P. I., Schleyer, T. K., Sanner, L. M., & Vest, J. R. (2025). Effectiveness of a clinical decision support system with prediction modeling to identify patients with health-related social needs in the emergency department: Study protocol. *PLOS One*, *20*(5), e0323094. https://doi.org/10.1371/journal.pone.0323094

Merriam-Webster. (2023, December 5). *Definition of demographic*.

Merrick, M. T., Ford, D. C., Ports, K. A., Guinn, A. S., Chen, J., Klevens, J., Metzler, M., Jones, C. M., Simon, T. R., Daniel, V. M., Ottley, P., & Mercy, J. A. (2019). Vital signs: Estimated proportion of adult health problems attributable to adverse childhood experiences and implications for prevention - 25 states, 2015-2017. *MMWR. Morbidity and Mortality Weekly Report*, *68*(44), 999–1005. https://doi.org/10.15585/mmwr.mm6844e1

Meyers, J. (2022, December 1). *Demographic table and subgroup summary macro %TABLEN*. SAS Communities Library. https://communities.sas.com/t5/SAS-Communities-Library/Demographic-Table-and-Subgroup-Summary-Macro-TABLEN/ta-p/634030

Mironova, O., & Bach, V. (2018). *Tenants at the edge: Rising insecurity among renters in New York City*. Community Service Society.

Monaghan, T. F., Rahman, S. N., Agudelo, C. W., Wein, A. J., Lazar, J. M., Everaert, K., & Dmochowski, R. R. (2021). Foundational statistical principles in medical research: Sensitivity, specificity, positive predictive value, and negative predictive value. *Medicina*, *57*(5), 503. https://doi.org/10.3390/medicina57050503

Moore, K. A. (2006). *Defining the term "at risk"* [Dataset]. https://doi.org/10.1037/e574062006-001

Morgan, C. A., Chang, Y.-H., Choy, O., Tsai, M.-C., & Hsieh, S. (2021). Adverse childhood experiences are associated with reduced psychological resilience in youth: A systematic review and meta-analysis. *Children (Basel, Switzerland)*, *9*(1), 27. https://doi.org/10.3390/children9010027

Moussaoui, D., & Grover, S. R. (2022). The association between childhood adversity and risk of dysmenorrhea, pelvic pain, and dyspareunia in adolescents and young adults: A systematic review. *Journal of Pediatric and Adolescent Gynecology*, *35*(5), 567–574. https://doi.org/10.1016/j.jpag.2022.04.010

Muisyo, T., Bernardo, E. O., Camazine, M., Colvin, R., Thomas, K. A., Borgman, M. A., & Spinella, P. C. (2019). Mortality prediction in pediatric trauma. *Journal of Pediatric Surgery*, *54*(8), 1613–1616. https://doi.org/10.1016/j.jpedsurg.2018.08.045

Mukhopadhyay, P. K. (n.d.). *Survey data imputation with PROC SURVEYIMPUTE*.

Musoro, J. Z., Zwinderman, A. H., Puhan, M. A., ter Riet, G., & Geskus, R. B. (2014). Validation of prediction models based on lasso regression with multiply imputed data. *BMC Medical Research Methodology*, *14*(1), 116. https://doi.org/10.1186/1471-2288-14-116

Nacht, C. L., Kelly, M. M., Edmonson, M. B., Sklansky, D. J., Shadman, K. A., Kind, A. J. H., Zhao, Q., Barreda, C. B., & Coller, R. J. (2022). Association between neighborhood disadvantage and pediatric readmissions. *Maternal and Child Health Journal*, *26*(1), 31–41. https://doi.org/10.1007/s10995-021-03310-4

Nahm, F. S. (2022). Receiver operating characteristic curve: Overview and practical use for clinicians. *Korean Journal of Anesthesiology*, *75*(1), 25–36. https://doi.org/10.4097/kja.21209

National Academy for State Health Policy. (2019, December 5). *New York CHIP fact sheet*. NASHP. https://nashp.org/new-york-chip-fact-sheet/

National Academy for State Health Policy. (2024, January 11). *State community health worker policies*. NASHP. https://nashp.org/state-tracker/state-community-health-worker-policies/

National Center for Health Statistics (U.S.). (2023). *National Hospital Ambulatory Medical Care Survey: 2019 National Summary Tables*. https://doi.org/10.15620/cdc:123251

National Collaborating Centre for Determinants of Health. (n.d.). *Rio Political Declaration on Social Determinants of Health*. Retrieved March 7, 2022, from https://nccdh.ca/resources/entry/rio-political-declaration-on-social-determinants-of-health

National HIV Curriculum. (n.d.). *Generalized Anxiety Disorder 7-item (GAD-7)—Mental Health Screening*. Retrieved February 20, 2024, from https://www.hiv.uw.edu/page/mental-health-screening/gad-7

National Institute of Environmental Health Sciences. (n.d.). *Air pollution and your health*. Retrieved August 14, 2024, from https://www.niehs.nih.gov/health/topics/agents/air-pollution

National Library of Medicine. (n.d.). *The PHQ-9—PMC*. Retrieved February 20, 2024, from https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1495268/

National Library of Medicine. (n.d.). *How to understand your lab results: MedlinePlus Medical Test*. Retrieved December 10, 2023, from https://medlineplus.gov/lab-tests/how-to-understand-your-lab-results/

National Research Council. (Ed.). (2011). *The science of adolescent risk-taking: Workshop report*. National Academies Press.

Nelson, C. A., Bhutta, Z. A., Burke Harris, N., Danese, A., & Samara, M. (2020). Adversity in childhood is linked to mental and physical health throughout life. *The BMJ*, *371*, m3048. https://doi.org/10.1136/bmj.m3048

Nelson, S., Beveridge, J. K., Mychasiuk, R., & Noel, M. (2021). Adverse childhood experiences (ACEs) and internalizing mental health, pain, and quality of life in youth with chronic pain: A longitudinal examination. *The Journal of Pain*, *22*(10), 1210–1220. https://doi.org/10.1016/j.jpain.2021.03.143

New York City Department of Health and Mental Hygiene. (n.d.). *Cockroaches data for NYC*. Environment & Health Data Portal. Retrieved August 14, 2024, from https://a816-dohbesp.nyc.gov/IndicatorPublic/data-explorer/cockroaches/

New York City Department of Health and Mental Hygiene. (n.d.). *Community Health Survey: Health Insurance Type*. Epiquery. Retrieved June 19, 2025, from https://a816-health.nyc.gov/hdi/epiquery/visualizations?PageType=ps&PopulationSource=YRBS

New York City Department of Health and Mental Hygiene. (n.d.). *Economic conditions data for NYC*. Environment & Health Data Portal. Retrieved August 14, 2024, from https://a816-dohbesp.nyc.gov/IndicatorPublic/data-explorer/economic-conditions/

New York City Department of Health and Mental Hygiene. (n.d.). *Health impacts of air pollution: Data for NYC*. Environment & Health Data Portal. Retrieved August 14, 2024, from https://a816-dohbesp.nyc.gov/IndicatorPublic/data-explorer/health-impacts-of-air-pollution/

New York City Department of Health and Mental Hygiene. (n.d.). *Housing maintenance data for NYC*. Environment & Health Data Portal. Retrieved August 14, 2024, from https://a816-dohbesp.nyc.gov/IndicatorPublic/data-explorer/housing-maintenance/

New York City Department of Health and Mental Hygiene. (n.d.). *Interactive heat vulnerability index*. Environment & Health Data Portal. Retrieved August 14, 2024, from https://a816-dohbesp.nyc.gov/IndicatorPublic/data-features/hvi/

New York City Department of Health and Mental Hygiene. (n.d.). *Mice and rats data for NYC*. Environment & Health Data Portal. Retrieved August 14, 2024, from https://a816-dohbesp.nyc.gov/IndicatorPublic/data-explorer/mice-and-rats/

New York City Department of Health and Mental Hygiene. (n.d.). *NYC youth risk behavior survey, 2009-2019*. Retrieved August 7, 2024, from https://a816-health.nyc.gov/hdi/epiquery/visualizations?PageType=ps&PopulationSource=YRBS

New York City Department of Health and Mental Hygiene. (n.d.). *Violence data for NYC*.

Environment & Health Data Portal. Retrieved August 14, 2024, from https://a816-dohbesp.nyc.gov/IndicatorPublic/data-explorer/violence/

New York State Department of Health. (n.d.). *Memo: Reimbursement of HRSN service providers for services delivered via the SCN program*. Retrieved June 10, 2025, from https://www.health.ny.gov/health_care/medicaid/redesign/sdh/scn/docs/cbo_funding_guide.pdf

New York State Department of Health. (2021). *1115 waiver demonstration: Conceptual framework. A federal-state partnership to address health disparities exacerbated by the COVID-19 pandemic*. https://www.health.ny.gov/health_care/medicaid/redesign/2021/docs/2021-08_1115_waiver_concept_paper.pdf

Newton-Dame, R., Jacobson, L., Wallach, A. B., Silverman, E., Dreyer, B., & Long, T. (2025). Caring for patients seeking asylum: Early data from the safety net system in New York City. *Journal of Public Health Management and Practice: JPHMP*, *31*(3), 477–485. https://doi.org/10.1097/PHH.0000000000002106

NYC Health+Hospitals. (2022, January 7). *Sen. Kirsten Gillibrand meets with Public Health Corps workers at Bellevue*. https://www.nychealthandhospitals.org/pressrelease/sen-kirsten-gillibrand-meets-with-public-health-corps-workers-at-bellevue/

Obermeyer, Z., & Topol, E. J. (2021). Artificial intelligence, bias, and patients' perspectives. *Lancet (London, England)*, *397*(10289), 2038. https://doi.org/10.1016/S0140-6736(21)01152-1

Obermeyer, Z., Nissan, R., Stern, M., Eaneff, S., Bembeneck, E. J., & Mullainathan, S. (2021). *Algorithmic bias playbook*. Chicago Booth, Center for Applied Artificial Intelligence. https://www.chicagobooth.edu/-/media/project/chicago-booth/centers/caai/docs/algorithmic-bias-playbook-june-2021

Obermeyer, Z., Powers, B., Vogeli, C., & Mullainathan, S. (2019). Dissecting racial bias in an algorithm used to manage the health of populations. *Science*, *366*(6464), 447–453. https://doi.org/10.1126/science.aax2342

Office of Community and Family Affairs. (n.d.). *About the CRAFFT – CRAFFT*. Retrieved February 20, 2024, from https://crafft.org/about-the-crafft/

Oh, D. L., Jerman, P., Silvério Marques, S., Koita, K., Purewal Boparai, S. K., Burke Harris, N., & Bucci, M. (2018). Systematic review of pediatric health outcomes associated with childhood adversity. *BMC Pediatrics*, *18*(1), 83. https://doi.org/10.1186/s12887-018-1037-7

Rijnbeek, P., & Reps, J. (2021). *The book of observational health data sciences and informatics*. https://ohdsi.github.io/TheBookOfOhdsi/

Ortiz, R., & Sibinga, E. M. (2017). The role of mindfulness in reducing the adverse effects of childhood stress and trauma. *Children (Basel, Switzerland)*, *4*(3), Article 3. https://doi.org/10.3390/children4030016

Overhage, J. M., & Johnson, K. B. (2020). Pediatrician electronic health record time use for outpatient encounters. *Pediatrics*, *146*(6), e20194017. https://doi.org/10.1542/peds.2019-4017

Ozer, E. M., & Irwin Jr., C. E. (2011). Chapter 66. Psychological development. In C. D. Rudolph, A. M. Rudolph, G. E. Lister, L. R. First, & A. A. Gershon (Eds.), *Rudolph's pediatrics* (22nd ed.). The McGraw-Hill Companies. https://accesspediatrics.mhmedical.com/content.aspx?aid=6736874

P. Iqbal, M., Walpola, R., Harris‐Roxas, B., Li, J., Mears, S., Hall, J., & Harrison, R. (2022). Improving primary health care quality for refugees and asylum seekers: A systematic review of interventional approaches. *Health Expectations : An International Journal of Public Participation in Health Care and Health Policy*, *25*(5), 2065–2094.

https://doi.org/10.1111/hex.13365
Paneth, N. (2004). Assessing the contributions of John Snow to epidemiology: 150 years after removal of the Broad Street pump handle. *Epidemiology*, *15*(5), 514–516. https://doi.org/10.1097/01.ede.0000135915.94799.00
Parade, S. H., Huffhines, L., Daniels, T. E., Stroud, L. R., Nugent, N. R., & Tyrka, A. R. (2021). A systematic review of childhood maltreatment and DNA methylation: Candidate gene and epigenome-wide approaches. *Translational Psychiatry*, *11*, 134. https://doi.org/10.1038/s41398-021-01207-y
Paulus, J. K., & Kent, D. M. (2020). Predictably unequal: Understanding and addressing concerns that algorithmic clinical prediction may increase health disparities. *NPJ Digital Medicine*, *3*, 99. https://doi.org/10.1038/s41746-020-0304-9
Pavilon, J., & Virgin, V. (2022). *Social determinants of immigrants' health in New York City: A study of six neighborhoods in Brooklyn and Queens*. Candid. https://search.issuelab.org/resource/social-determinants-of-immigrants-health-in-new-york-city-a-study-of-six-neighborhoods-in-brooklyn-and-queens.html
Perry, R., Braileanu, G., Palmer, T., & Stevens, P. (2019). The economic burden of pediatric asthma in the United States: Literature review of current evidence. *PharmacoEconomics*, *37*(2), 155–167. https://doi.org/10.1007/s40273-018-0726-2
Petruccelli, K., Davis, J., & Berman, T. (2019). Adverse childhood experiences and associated health outcomes: A systematic review and meta-analysis. *Child Abuse & Neglect*, *97*, 104127. https://doi.org/10.1016/j.chiabu.2019.104127
Pfaltz, M. C., Halligan, S. L., Haim-Nachum, S., Sopp, M. R., Åhs, F., Bachem, R., Bartoli, E., Belete, H., Belete, T., Berzengi, A., Dukes, D., Essadek, A., Iqbal, N., Jobson, L., Langevin, R., Levy-Gigi, E., Lüönd, A. M., Martin-Soelch, C., Michael, T., & Seedat, S. (2022). Social functioning in individuals affected by childhood maltreatment: Establishing a research agenda to inform interventions. *Psychotherapy and Psychosomatics*, *91*(4), 238–251. https://doi.org/10.1159/000523667
Pham, T. M., Pandis, N., & White, I. R. (2024). Missing data: Issues, concepts, methods. *Seminars in Orthodontics*, *30*(1), 37–44. https://doi.org/10.1053/j.sodo.2024.01.007
Pinto, N. R., Applebaum, M. A., Volchenboum, S. L., Matthay, K. K., London, W. B., Ambros, P. F., Nakagawara, A., Berthold, F., Schleiermacher, G., Park, J. R., Valteau-Couanet, D., Pearson, A. D. J., & Cohn, S. L. (2015). Advances in risk classification and treatment strategies for neuroblastoma. *Journal of Clinical Oncology: Official Journal of the American Society of Clinical Oncology*, *33*(27), 3008–3017. https://doi.org/10.1200/JCO.2014.59.4648
Piotrowski, C. C., Talavera, G. A., & Mayer, J. A. (2009). Healthy Steps: A systematic review of a preventive practice-based model of pediatric care. *Journal of Developmental & Behavioral Pediatrics*, *30*(1), 91–103. https://doi.org/10.1097/DBP.0b013e3181976a95
Poon, E. G., Lemak, C. H., Rojas, J. C., Guptill, J., & Classen, D. (2025). Adoption of artificial intelligence in healthcare: Survey of health system priorities, successes, and challenges. *Journal of the American Medical Informatics Association*, *ocaf065*. https://doi.org/10.1093/jamia/ocaf065
Rajaguru, V., Han, W., Kim, T. H., Shin, J., & Lee, S. G. (2022). LACE Index to Predict the High Risk of 30-Day Readmission: A Systematic Review and Meta-Analysis. *Journal of Personalized Medicine*, *12*(4), 545. https://doi.org/10.3390/jpm12040545
Richter, L., Komárek, A., Desmond, C., Celentano, D., Morin, S., Sweat, M., Chariyalertsak, S., Chingono, A., Gray, G., Mbwambo, J., & Coates, T. (2014). Reported physical and sexual abuse in childhood and adult HIV risk behaviour in three African countries: Findings from project accept (HPTN-043). *AIDS and Behavior*, *18*(2), 381–389. https://doi.org/10.1007/s10461-013-0439-7
Rifkin, S. B. (2018). Alma Ata after 40 years: Primary health care and health for all—from

consensus to complexity. *BMJ Global Health*, *3*(Suppl 3), e001188. https://doi.org/10.1136/bmjgh-2018-001188

Rijnbeek, P., & Reps, J. (2021). *The book of observational health data sciences and informatics*. OHDSI. https://ohdsi.github.io/TheBookOfOhdsi/

Riley, M., Patterson, V., Lane, J. C., Won, K. M., & Ranalli, L. (2018). The adolescent champion model: Primary care becomes adolescent-centered via targeted quality improvement. *The Journal of Pediatrics*, *193*, 229–236.e1. https://doi.org/10.1016/j.jpeds.2017.09.084

Rinehart, D. J., Oronce, C., Durfee, M. J., Ranby, K. W., Batal, H. A., Hanratty, R., Vogel, J., & Johnson, T. L. (2018). Identifying subgroups of adult superutilizers in an urban safety-net system using latent class analysis: Implications for clinical practice. *Medical Care*, *56*(1), e1–e9. https://doi.org/10.1097/MLR.0000000000000628

Robertshaw, L., Dhesi, S., & Jones, L. L. (2017). Challenges and facilitators for health professionals providing primary healthcare for refugees and asylum seekers in high-income countries: A systematic review and thematic synthesis of qualitative research. *BMJ Open*, *7*(8), e015981. https://doi.org/10.1136/bmjopen-2017-015981

Rogosch, F. A., Dackis, M. N., & Cicchetti, D. (2011). Child maltreatment and allostatic load: Consequences for physical and mental health in children from low-income families. *Development and Psychopathology*, *23*(4), 1107–1124. https://doi.org/10.1017/S0954579411000587

Rothman, K. J., & Greenland, S. (2005). Hill's criteria for causality. In P. Armitage & T. Colton (Eds.), *Encyclopedia of biostatistics* (p. b2a03072). John Wiley & Sons, Ltd. https://doi.org/10.1002/0470011815.b2a03072

Russell, S. T., & Fish, J. N. (2016). Mental health in lesbian, gay, bisexual, and transgender (LGBT) youth. *Annual Review of Clinical Psychology*, *12*, 465–487. https://doi.org/10.1146/annurev-clinpsy-021815-093153

Saewyc, E. M. (2011). Research on adolescent sexual orientation: Development, health disparities, stigma and resilience. *Journal of Research on Adolescence: The Official Journal of the Society for Research on Adolescence*, *21*(1), 256–272. https://doi.org/10.1111/j.1532-7795.2010.00727.x

Sahiner, B., Chen, W., Samala, R. K., & Petrick, N. (2023). Data drift in medical machine learning: Implications and potential remedies. *The British Journal of Radiology*, *96*(1150), 20220878. https://doi.org/10.1259/bjr.20220878

Salerno, J. P. (2023). Impact of racist microaggressions and LGBTQ-related minority stressors: Effects on psychological distress among LGBTQ+ young people of color. *Preventing Chronic Disease*, *20*. https://doi.org/10.5888/pcd20.220371

Sandhu, S., Xu, J., Eisenson, H., & Prvu Bettger, J. (2021). Workforce models to screen for and address patients' unmet social needs in the clinic setting: A scoping review. *Journal of Primary Care & Community Health*, *12*, 21501327211021021. https://doi.org/10.1177/21501327211021021

Sarkar, P. (2015, September 24). *STAT 383C: Statistical modeling I, lecture 9*. https://www.cs.cmu.edu/~psarkar/sds383c_16/lecture9_scribe.pdf

Scholes-Balog, K. E., Hemphill, S. A., Heerde, J. A., Toumbourou, J. W., & Patton, G. C. (2020). Childhood social environmental and behavioural predictors of early adolescent onset cannabis use. *Drug and Alcohol Review*, *39*(4), 384–393. https://doi.org/10.1111/dar.13077

Scott, K., Beckham, S. W., Gross, M., Pariyo, G., Rao, K. D., Cometto, G., & Perry, H. B. (2018). What do we know about community-based health worker programs? A systematic review of existing reviews on community health workers. *Human Resources for Health*, *16*(1), 39. https://doi.org/10.1186/s12960-018-0304-x

Shamseer, L., Moher, D., Clarke, M., Ghersi, D., Liberati, A., Petticrew, M., Shekelle, P., Stewart, L. A., & PRISMA-P Group. (2015). Preferred reporting items for systematic review and

meta-analysis protocols (PRISMA-P) 2015: Elaboration and explanation. *BMJ*, *349*, g7647. https://doi.org/10.1136/bmj.g7647

Shekelle, P., Paige, N., Apaydin, E., & Goldhaber-Fiebert, J. (2019). *What is the optimal panel size in primary care? A systematic review* (Evidence Synthesis Program). Veterans Health Administration. https://www.hsrd.research.va.gov/publications/esp/reports.cfm

Sherwin, R. P. (1983). What is an adverse health effect? *Environmental Health Perspectives*, *52*, 177–182.

Shonkoff, J. P., Boyce, W. T., Levitt, P., Martinez, F. D., & McEwen, B. (2021). Leveraging the biology of adversity and resilience to transform pediatric practice. *Pediatrics*, *147*(2), e20193845. https://doi.org/10.1542/peds.2019-3845

Shonkoff, J. P., Slopen, N., & Williams, D. R. (2021). Early childhood adversity, toxic stress, and the impacts of racism on the foundations of health. *Annual Review of Public Health*, *42*, 115–134. https://doi.org/10.1146/annurev-publhealth-090419-101940

Siegel, D. J. (2013). *Brainstorm: The power and purpose of the teenage brain*. Jeremy P. Tarcher/Penguin.

Sills, M. R., Hall, M., Cutler, G. J., Colvin, J. D., Gottlieb, L. M., Macy, M. L., Bettenhausen, J. L., Morse, R. B., Fieldston, E. S., Raphael, J. L., Auger, K. A., & Shah, S. S. (2017). Adding social determinant data changes children's hospitals' readmissions performance. *The Journal of Pediatrics*, *186*, 150–157.e1. https://doi.org/10.1016/j.jpeds.2017.03.056

Slopen, N., Koenen, K. C., & Kubzansky, L. D. (2012). Childhood adversity and immune and inflammatory biomarkers associated with cardiovascular risk in youth: A systematic review. *Brain, Behavior, and Immunity*, *26*(2), 239–250. https://doi.org/10.1016/j.bbi.2011.11.003

Sosnowski, D. W., Musci, R. J., & Johnson, S. B. (2021). Positive functioning moderates the association between classes of adverse childhood experiences and adolescent depressive symptoms. *Adversity and Resilience Science*, *2*(3), 157–168. https://doi.org/10.1007/s42844-021-00033-x

Spencer, A. E., Baul, T. D., Sikov, J., Adams, W. G., Tripodis, Y., Buonocore, O., Jellinek, M., Michael Murphy, J., & Garg, A. (2020). The relationship between social risks and the mental health of school-age children in primary care. *Academic Pediatrics*, *20*(2), 208–215. https://doi.org/10.1016/j.acap.2019.11.006

Stevens, R. J., & Poppe, K. K. (2020). Validation of clinical prediction models: What does the "calibration slope" really measure? *Journal of Clinical Epidemiology*, *118*, 93–99. https://doi.org/10.1016/j.jclinepi.2019.09.016

Steyerberg, E. W., & Vergouwe, Y. (2014). Towards better clinical prediction models: Seven steps for development and an ABCD for validation. *European Heart Journal*, *35*(29), 1925–1931. https://doi.org/10.1093/eurheartj/ehu207

Stirling, J., Gavril, A., Brennan, B., Sege, R. D., Dubowitz, H., & American Academy of Pediatrics, Committee on Child Abuse and Neglect. (2024). The pediatrician's role in preventing child maltreatment: Clinical report. *Pediatrics*, *154*(2), e2024067608. https://doi.org/10.1542/peds.2024-067608

Struja, T., Baechli, C., Koch, D., Haubitz, S., Eckart, A., Kutz, A., Kaeslin, M., Mueller, B., & Schuetz, P. (2020). What are they worth? Six 30-day readmission risk scores for medical inpatients externally validated in a Swiss cohort. *Journal of General Internal Medicine*, *35*(7), 2017–2024. https://doi.org/10.1007/s11606-020-05638-z

Student Homelessness in New York City. (2021). Advocates for Children of New York. https://www.advocatesforchildren.org/sites/default/files/library/nyc_student_homelessness_20-21.pdf?pt=1

Su, C., Aseltine, R., Doshi, R., Chen, K., Rogers, S. C., & Wang, F. (2020). Machine learning for suicide risk prediction in children and adolescents with electronic health records. *Translational Psychiatry*, *10*, 413. https://doi.org/10.1038/s41398-020-01100-0

Sun, J. W., Bourgeois, F. T., Haneuse, S., Hernández-Díaz, S., Landon, J. E., Bateman, B. T., & Huybrechts, K. F. (2021). Development and validation of a pediatric comorbidity index. *American Journal of Epidemiology*, *190*(5), 918–927. https://doi.org/10.1093/aje/kwaa244

Sundararajan, V., Henderson, T., Perry, C., Muggivan, A., Quan, H., & Ghali, W. A. (2004). New ICD-10 version of the Charlson comorbidity index predicted in-hospital mortality. *Journal of Clinical Epidemiology*, *57*(12), 1288–1294. https://doi.org/10.1016/j.jclinepi.2004.03.012

Tolkkinen, A., Madanat-Harjuoja, L., Taskinen, M., Rantanen, M., Malila, N., & Pitkäniemi, J. (2018). Impact of parental socioeconomic factors on childhood cancer mortality: A population-based registry study. *Acta Oncologica (Stockholm, Sweden)*, *57*(11), 1547–1555. https://doi.org/10.1080/0284186X.2018.1478125

Traub, F., & Boynton-Jarrett, R. (2017). Modifiable resilience factors to childhood adversity for clinical pediatric practice. *Pediatrics*, *139*(5), e20162569. https://doi.org/10.1542/peds.2016-2569

Truchil, A., Dravid, N., Singer, S., Martinez, Z., Kuruna, T., & Waulters, S. (2018). Lessons from the Camden Coalition of Healthcare Providers' First Medicaid Shared Savings Performance Evaluation. *Population Health Management*, *21*(4), 278–284. https://doi.org/10.1089/pop.2017.0164

UCLA Advanced Research Computing. (n.d.). *Negative binomial regression*. Statistical Methods and Data Analysis. Retrieved March 2, 2025, from https://stats.oarc.ucla.edu/sas/output/negative-binomial-regression/

US EPA, Office of Air and Radiation. (2016, March 21). *What is ozone?* [Data and Tools]. https://www.epa.gov/ozone-pollution-and-your-patients-health/what-ozone

Vest, J. (2025, June 9). *Prospective trial of prediction modeling to identify patients with health-related social needs in the emergency department*. Academy Health Annual Research Meeting, Minneapolis, MN.

Wade, R., Becker, B. D., Bevans, K. B., Ford, D. C., & Forrest, C. B. (2017). Development and evaluation of a short adverse childhood experiences measure. *American Journal of Preventive Medicine*, *52*(2), 163–172. https://doi.org/10.1016/j.amepre.2016.09.033

Walters, S., Konty, K., Day, S., Agerton, T., Olson, C., & Wilson, L. (2021). *Disparities among children with asthma in New York City* (Epi Data Brief 126). New York City Department of Health and Mental Hygiene. https://www1.nyc.gov/assets/doh/downloads/pdf/epi/databrief126.pdf

Wang, M. C., Bangaru, S., & Zhou, K. (2023). Care for vulnerable populations with chronic liver disease: A safety-net perspective. *Healthcare*, *11*(20), 2725. https://doi.org/10.3390/healthcare11202725

Weitzman Institute and AcademyHealth, Damian, A. J., Bruner, C., Samudio, M., Schweer-Collins, M., McCrae, J., Hails, K. A., Kuklinski, M., Wissow, L. S., Earls, M. F. E., Smith, J., Berkel, C., Shearman, N., Javier, J., Hemady, K. T., Dumitriu, D., Barker, T., Bethell, C., & Polfuss, M. (2025). The state of healthy parenting in primary care interventions in advancing health equity. *NAM Perspectives*, *3*. https://doi.org/10.31478/202503a

Wolf, K., & Schmitz, J. (2024). Scoping review: Longitudinal effects of the COVID-19 pandemic on child and adolescent mental health. *European Child & Adolescent Psychiatry*, *33*(5), 1257–1312. https://doi.org/10.1007/s00787-023-02206-8

World Health Organization. (n.d.). *The adolescent health indicators recommended by the Global Action for Measurement of Adolescent Health*. Retrieved August 6, 2024, from https://www.who.int/publications/i/item/9789240092198

World Health Organization. (2023, April 28). *Adolescent and young adult health*. https://www.who.int/news-room/fact-sheets/detail/adolescents-health-risks-and-solutions

Xu, J., Xiao, Y., Wang, W. H., Ning, Y., Shenkman, E. A., Bian, J., & Wang, F. (2022). Algorithmic fairness in computational medicine. *EBioMedicine*, *84*, 104250. https://doi.org/10.1016/j.ebiom.2022.104250

Xu, L., Zimmermann, M., Forkey, H., Griffin, J., Wilds, C., Morgan, W. S., Byatt, N., & McNeal, C. J. (2022). How to mitigate risk of premature cardiovascular disease among children and adolescents with mental health conditions. *Current Atherosclerosis Reports*, *24*(4), 253–264. https://doi.org/10.1007/s11883-022-00998-9

Yong, E. (2021, October 23). How public health took part in its own downfall. *The Atlantic*. https://www.theatlantic.com/health/archive/2021/10/how-public-health-took-part-its-own-downfall/620457/

Ziring, J., Gogia, S., Newton-Dame, R., Singer, J., & Chokshi, D. A. (2018). An all-payer risk model for super-utilization in a large safety net system. *Journal of General Internal Medicine*, *33*(5), 596–598. https://doi.org/10.1007/s11606-018-4346-5