1. Background and Rationale:

It is estimated that 40% of children seen at tertiary children's hospitals and 25% of children nationally, have complex chronic conditions. These conditions are responsible for nearly a quarter of deaths in the U.S. among people 0 to 24 years old. Among those who survive, these conditions are often associated with significant and life-long disease burdens that impact the quality of life and productivity of both children and their families and costs families and the health care system dearly. Further, it is known that patients from some demographic, economic, or rural/urban backgrounds are more likely to die or have chronic complications. What is not known is why and what can be done about it. Congenital heart defects, for example, effect ~1% of live births and result in annual acute care costs of >$10 billion (US $ 2024). Treatment typically involves at least one—if not multiple—open heart surgeries, with half of children undergoing initial operation in their first year of life. Infant mortality for these children exceeds 10%, and survivors often describe significant healthcare burdens and challenging, lifelong, journeys. Our team and others have described 15-20% higher mortality for non-Hispanic Black and Hispanic children, even after adjusting for cardiac anatomy, clinical risk factors, family income (payer), and neighborhood sociodemographics.

This project aims to use the CO All Payer Claims Data (CO APCD), linked to the National Death Index, neighborhood characteristics from the U.S. Census Bureau, and clinical nuance for risk stratification from clinical registry data to identify modifiable drivers (mediators) of health outcomes and healthcare utilization / expenditures, to inform targeted interventions (policies, programs, and system organization) to improve outcomes and healthcare value for children and young adults.

In prior NIH-funded research, we focused on NY State Medicaid patients with congenital heart disease, as it is the most prevalent and resource intensive birth defect managed in the United States. We quantified—for the first time—longitudinal healthcare utilization and health expenditures for congenital heart patients and validated novel longitudinal outcomes for this congenital heart patients. We found that the average child on NY Medicaid spent >90 days in hospitals and doctors' offices in their first 5 postoperative years ($139K in Medicaid expenditures in first 5 years, not including initial surgery). We developed the first longitudinal congenital heart risk models the first congenital heart risk models that consider neighborhood characteristics, and the first congenital heart risk model for ICD-10 data.

We validated novel measures of access to high-quality / high-experience congenital heart providers as predictors of outcomes among NY Medicaid patients. Drawing from Penchansky and Thomas, access is often operationalized as patient-to-provider geographic distance. We expanded on this, evaluating both distance to and a range of potential descriptors of the quality / experience of the birth centers, primary care physicians, prenatal / pediatric cardiologists, congenital heart surgeons, surgical centers, and emergency rooms that patients used. Candidate measures were selected for consideration by expert consensus and include nationally endorsed measures, as available. We tested associations with risk-adjusted, acute and longitudinal outcomes, retaining measures reaching statistical significance.

We established the scientific premise for evaluation of provider characteristics as mediators of differences in congenital heart outcomes across patient demographics and neighborhood characteristics. Among NY Medicaid patients, for example, we found not only that Non-Hispanic Black and Hispanic children saw less experienced healthcare provider (defined across multiple measures) than their Non-Hispanic White, Medicaid-insured peers, but also that some of these differences explained 10-20% of observed differences in risk-adjusted outcomes—even when examining differences in providers within surgical centers.

In preliminary work, we have also described strong mediating effects of potentially modifiable neighborhood-level characteristics (e.g. exposure to environmental toxins / air particulate matter levels)

on outcomes and healthcare utilization. We are working currently to disentangle these effects from the impacts of poverty at large and neighborhood / geographic effects on access.

2. Study Aims:

We propose to use APCD to expand upon this work, examining aspects of healthcare system as mediators of differences in outcomes and healthcare resource utilization / expenditures for children both with congenital heart disease and with other chronic conditions. We will bring together APCDs from New York, Colorado, Texas, and Massachusetts, allowing us to assess the generalizability of our prior findings across four states in four distinct regions of the country, and including publicly and commercially insured patients to allow us to explore the impacts of poverty. In particular, we are interested in answering the following specific research questions:

Aim 1: We will first quantify the degree to which access to greater experience / quality multidisciplinary providers explains differences in risk-adjusted, longitudinal congenital heart patient outcomes across the life course.

> Aim 1.1. We will describe outcomes and healthcare utilization for congenital heart patients across the life course by disease severity, state, payer, patient and neighborhood demographics / economics / rurality. To assist in interpretation of our results, we will compare outcomes and healthcare utilization patients with congenital heart disease to the general, healthy, pediatric population. Heretofore, the phrase outcomes and healthcare utilization will be used to describe: our primary outcomes (longitudinal mortality and days alive and out of healthcare) and our secondary outcomes (Number / duration of inpatient admissions, number of emergency department visits, ambulatory / observation stays, outpatient primary care and subspecialty visits, and prescriptions; total payer expenditures; and disease-specific healthcare encounters).
>
> Aim 1.2. We will apply mediation analyses to quantify the degree to which measures of distance to or the experience / quality of obstetricians, birth centers, primary care providers, prenatal / pediatric subspecialists, surgeons, or surgical centers explain differences in outcomes and healthcare utilization across the lifespan, considering mediating and modifying effects of state, payer, patient and neighborhood demographics / economics / rurality / environmental toxins.

Aim 2: Quantify how much congenital heart multidisciplinary provider team (network) and payer characteristics explain differences in risk-adjusted, longitudinal patient outcomes and healthcare utilization across the lifespan.

> Aim 2.1. We will leverage innovative network analyses to describe congenital heart provider team characteristics -- the connections between multidisciplinary providers who share congenital heart patients across time.
>
> Aim 2.2 We will apply mediation analyses to quantify whether and by how much provider network and payer characteristics explain observed differences in outcomes and healthcare utilization across the lifespan, considering mediating and modifying effects of state, payer, and neighborhood demographics / economics / rurality / environmental toxins.

Aim 3: We will leverage the lessons learned in Aims 1 and 2, to quantify the degree to which access to greater experience / quality multidisciplinary providers, and provider team / payer characteristics explains differences in risk-adjusted, outcomes and healthcare utilization for all children across the life span.

Aim 3.1 We will define longitudinal, disease and risk-stratified outcomes and healthcare utilization for children and young adults with chronic disease, adjusting for clinical characteristics, demographics, and family and neighborhood characteristics, comparing to the general population, and identifying populations with the highest healthcare burdens / the greatest variation in outcomes. Children with chronic conditions will initially be identified using the Complex Chronic Conditions v3, developed by Colorado researcher and colleague, James Feinstein. We will consider the use of other pediatric, administrative data specific algorithms, such as the Pediatric Medical Complexity Algorithm v3.

Aim 3.2. We will apply mediation analyses, as above, to quantify the degree to which measures of distance to or measures of the experience / quality of multidisciplinary providers explain observed explains differences in risk-adjusted, outcomes and healthcare utilization for non-cardiac children with chronic disease across the lifespan.

Aim 3.3. We will apply mediation analyses to quantify whether and by how much provider network and payer characteristics explain observed differences in outcomes and healthcare utilization across the lifespan, considering mediating and modifying effects of state, payer, and neighborhood demographics / economics / rurality / environmental toxins.

3. Study Design:

In Aim 1, we validate across states and payers (income) the predictive validity and mediating effects of **provider** characteristics assessed in the parent R01. **In Aim 2**, we combine network analysis, causal mediation, and advanced statistical modeling, to characterize multidisciplinary congenital heart **provider networks** (patient-sharing teams) and to understand how patients' access to provider networks impacts which downstream providers they see and, ultimately, their outcomes. Obstetricians, prenatal / pediatric cardiologists, congenital heart surgeons, and others work formally and informally as teams, providing complementary care for patients across the life course. We hypothesize that these teams are more important than any one provider alone. Further, we hypothesize that both upstream providers and health insurers (payers) influence team compositions. Understanding these relationships has potential to influence how healthcare system leaders' approach, for example, future patient referrals, teamwork, mentorship, and states' managed care contracts. **In Aim 3**, we expand beyond congenital heart disease, using the lessons learned from congenital heart disease to identify modifiable mediators of outcomes and healthcare utilization across childhood chronic disease. We hypothesize that upstream healthcare factors (e.g. insurance, prenatal care, birth center) impact downstream care access (e.g. quality or experience of congenital heart team) and outcomes, and that analyses will reveal modifiable healthcare system features as targets for care improvement.

4. Data Sources:

1) All Payer Claims Databases (APCDs): APCDs are state-administered / state-mandated data repositories that include direct identifiers and longitudinal data on all billed services associated with—but not limited to— inpatient, outpatient, emergency room, urgent care, pharmacy, rehabilitation, and home healthcare services, across institutions and across time. APCDs capture both 100% of Medicaid and 60-80% of non-Medicaid (commercial and other governmental) insurance claims. Unique identifiers follow patients even when they change insurers. We will use APCDs from NY, Massachusetts, Colorado, and Texas. While the NY and Texas ACPDs are not broadly available, we have worked with the NY and Texas APCDs and have been given permission to access these data; these APCDs are ready for our use.

2) National Death Index: As in our prior work, longitudinal mortality will be obtained from the Centers for Disease Control and Prevention's National Death Index (NDI).

3) Patient Residential Neighborhood Characteristics: Measures of neighborhood-level demographics / economics / rurality / environmental toxins will be obtained via the US Census American Community Survey 5-year estimates. As in our prior work, key measures will be identified via principal components analysis. Composite measures will again be considered (e.g. Yost, Childhood Opportunity Index). Rurality will again be assessed as commuting time (patients' home addresses to both the nearest and used (accessed) providers), as well as via Rural-Urban Commuting Area / Combined Statistical Areas or other publicly available measure as appropriate for the geography. We already have these data on our local server.

6) Provider Characteristics: Provider years of experience, board certifications, and addresses come via the American Medical Association Physician Masterfile and Departments of Health's rosters. Other measures will be derived via claims.

7) Clinical Registry: For Aims 1 and 2, cases will be identifies / clinical detail enhanced via linkage to The Society of Thoracic Surgeons-Congenital Heart Surgery Database (STS-CHSD). The STS-CHSD is the world's largest congenital heart surgery registry. It captures >90% of US congenital heart operations, with great clinical granularity. Data include direct identifiers, key demographics, 110 anatomically specific cardiac diagnoses, 187 procedures, 22 comorbidities / preoperative risk-factors, bypass, cross-clamp, and circulatory arrest times, 11 major postoperative morbidities, and operative mortality. Missingness for key demographics in the last decade is <1%. Colorado Children's Hospital and HCA HealthONE Rocky Mountain Children's own their own locally-held STS-CHSD and are partners in this application. These institutions have agreed to share their locally held clinically registry data with us. These data will be stored on our local server. For Aim 3, the cohort will initially be defined the CCC and PMCA algorithms. To the extent to which we find key chronic diseases to lack specificity / sensitivity in these algorithms, Colorado Children's Hospital and HCA HealthONE Rocky Mountain Children's will work with us to identify the most appropriate, locally-held clinical registry necessary to enhance scientific rigor.

5. Analytic Approach

Patient characteristics, mortality, healthcare utilization, and healthcare expenditures will be summarized by year of initial surgery, disease severity, patient age, key patient and neighborhood demographics, state, and payer. To provide context to assist in interpretation, clinical outcomes, healthcare utilization, and health expenditures will be compared for patients with cardiac and other chronic diseases and for matched strata of otherwise healthy children from the comparison cohort. Univariable and standard multivariable analyses will test associations between predictor variables and outcomes. We will consider fractional logit models for percent days alive and out of healthcare; Cox proportional hazards models for mortality; Poisson models that allow for over dispersion for visit counts, number of inpatient days, and prescriptions; and log linear models for expenditures. Models will adjust for clinical characteristics, as well as key patient and neighborhood demographics, operative year and an offset term for time insurance enrolled. We will consider center fixed vs. random effects and test for interactions or effect modification by state, payer, rurality, surgical era, and neighborhood characteristics, stratifying as appropriate and quantifying effect attenuation. Fractional logit models use a generalized linear model framework with a logit link function to model outcomes as proportions on a 0-1 interval (including 0 and 1); this allows for retention of patients who die before discharge (zero day alive and out of healthcare), while ensuring predicted probabilities remain in this interval. We will conduct sensitivity analyses to understand effects of death and insurance disenrollment / change, comparing patients who: a) survived vs. died, b) enrolled

continuously in the same health plans vs. changed plans; c) enrolled continuously in a Medicaid vs. in a non-Medicaid plan vs. switched (and timing of switch); and d) enrolled continuously in any plan vs. churned (defined as <11 months per year enrolled) vs. moved or lost insurance without re-enrolling—standardized for months enrolled. We will also consider Tobit models that directly address truncation of healthcare utilization among patients who die, move, or lose insurance, or other models as appropriate.

6. Data Security and Privacy:

Each states' Department of Health will transmit APCD data to us in accordance with their own state laws and policies. Once received, data will be stored and accessed on a HIPAA compliant, password protected, and encrypted platform, Minerva. Data will be maintained and accessed exclusively on Minerva and will not be permitted to be stored on local devices or shared outside of the designated research team. The Minerva server segregates these sensitive information, from other data to prevent unauthorized access. Cloud-stored files are backed up automatically upon data modification or at regular intervals established by the system administrator. A detailed log of team members who have access to the instance of Minerva on which these data are stored is maintained by the PI and the Information Security Team. System privileges and usages are audited automatically at regular intervals by the information security and internal audit teams to ensure ongoing compliance with established security protocols. The instance of the server on which these data will be stored is configured such that raw data cannot be downloaded. Only aggregate data will be reported outside the study team. To reduce the possibility of indirect re-identification, we will also not report any small cell sizes.

All data will be retained for at least 10 years following study completion, in accordance with NIH guidelines. After this time, data will be destroyed in accordance with each state's preferences and legal requirements, as outlined in their APCD data use agreements. A modification of this IRB will be submitted with copies of these agreements once they are fully executed.

No protected health information will be disclosed to any other person or entity not IRB approved, except where required by law or for the authorized oversight of this research project.