2023 Call for Proposals

Proposal Deadline: August 18, 2023 (3 p.m. ET)

Health Data for Action (Data Access Award)

Leveraging Health Data for Actionable Insights

BACKGROUND

The Robert Wood Johnson Foundation (RWJF) is committed to building a national Culture of Health that enables all in a diverse society to lead healthier lives now and for generations to come. To reach a Culture of Health that provides everyone in America a fair and just opportunity for health, it is necessary to identify, understand, confront, and remove the structural barriers to health and wellbeing. At RWJF, we do this by gathering evidence and producing tools to advance health and health equity.

We believe deeply in the importance of research, evaluation, and learning to build a transdisciplinary evidence base that helps inform efforts to address these complex and interrelated determinants which shape our health. We recognize that access to rich data is a cornerstone of producing timely and objective research. The Health Data for Action (HD4A) program, launched by RWJF in 2017 and managed by AcademyHealth, is designed to leverage health data for actionable insights. HD4A, and the other signature research programs—Evidence for Action, Policies for Action, and Systems for Action—underscore RWJF's commitment to engage with health and related sectors in building a rigorous evidence base to address health and wellbeing and achieve health equity.

The HD4A program further aligns with RWJF efforts to engage the healthcare sector in promoting population health and broader health policy considerations. With greater access to health data, including data from delivery systems, payers, and health information technology, researchers can better answer important questions to build a Culture of Health and inform health policy to make healthcare more affordable and equitable.

HD4A aims to reduce the barriers often faced in accessing rich data by serving as a conduit between data owners and interested researchers. Through each HD4A call for proposals (CFP),

2023 Call for Proposals

Proposal Deadline: August 18, 2023 (3 p.m. ET)

RWJF will make valuable data from unique data owners available to researchers. The grantees awarded under the previous CFPs are available on the AcademyHealth website.

THE PROGRAM

The 2023 HD4A CFP is primarily for data access, with successful applicants receiving no-cost access to data from one of the 16 available data sets. While most selected projects will not be accompanied by a financial award for personnel or other project-related costs, limited funding is available for a small number of projects from principal investigators who are first-time investigators on a grant award at the time of application. Consistent with our Equity, Diversity, and Inclusion Commitment, we encourage applications with project investigators or team members from historically underrepresented backgrounds in research disciplines who conduct health equity research. Please see the eligibility criteria for more information on the award details.

Eligible research projects can focus on a variety of topics, including national practice patterns in long-term care, public health surveillance and population health, diagnostic quality, the health effects of COVID-19, risk adjustment, health outcomes, rehospitalization rates, the opioid epidemic, maternal and infant health, home- and community-based services, social determinants of health, racial and ethnic disparities, healthcare access, healthcare costs and utilization, quality of care, prescription patterns and medication adherence, prevalence of chronic disease, trends in insurance markets, consolidation and competition, and rate comparisons, among others.

The HD4A program will support innovative research that uses the available data to answer important research questions and inform health policy. Applicants under this CFP will write a proposal for a research study using data from one of the data sets below. Projects proposing to use and/or link two data sets under the HD4A CFP will be evaluated for feasibility and awarded on a case-by-case basis. Data providers have differing stipulations on linking or merging data sets, described in the data descriptions below and in the hyperlinked supplemental information.

Electronic Health Record & Clinical Data	Claims & Encounter Data	Transparency in Coverage Data
American Health Associates	All Payer Claims Database–Colorado	Clarify Health
COVID-19 Research Database	All Payer Claims Database–Virginia	Serif Health
HealthShare Exchange (HSX)	Health Care Cost Institute (HCCI)	Turquoise Health
Nebraska Healthcare Collaborative	Medicare Advantage (MA) enrollment and encounter data	Mathematica Data Innovation Lab
New Jersey Integrated Population Health Data (iPHD)	Transformed Medicaid Statistical Information System (T-MSIS) Analytic Files (TAF), from Mathematica Data Innovation Lab	
OCHIN Community Health Equity EHR Database		
Truveta, in partnership with Mathematica Data Innovation Lab		

2023 Call for Proposals

Proposal Deadline: August 18, 2023 (3 p.m. ET)

We strongly encourage applicants to consult the following resources, which are either included within the text of this CFP or hyperlinked from each data description below.

- Brief description of the available data: included below
- More detailed description of the available data: hyperlinked from each of the data descriptions below
- Recorded video overview of the available data set: hyperlinked from each of the descriptions below
- Data dictionary: hyperlinked from each of the data descriptions below

A snapshot of each data set is included on the next few pages:

Electronic Health Record & Clinical Data

American Health Associates (AHA)

Type of Data
Clinical laboratory, Longitudinal, Demographics
Number of Patients
500,000+
Number of Providers
1000+
Provider Setting
Skilled Nursing Facilities, Long-Term Acute Care, etc.
Years Available
2020-2022
Insurance Types Included
Multi-payer
Patient Sociodemographic Elements
Patient's facility ZIP code, birth date, sex
Patient Clinical Data Elements
Specimen type collected (blood, urine, sputum, etc), Diagnostic laboratory tests ordered, Laboratory test results, Clinical Laboratory Ordering Physician Data (NPI)
Method of Data Access
Custom extract with secure file transfer to requester
Length of Data Access
12 months, renewable at-cost (researcher responsible for cost)
Ability/Restriction on Merging Data Resources
Can be merged at individual patient level and for providers with NPI
Types of Data Arrangements Needed
Data Use Agreement

COVD-19 Research Database

Type of Data
COVID 19 Research database offers claims, EHR and SDOH–85 Billion Records
Number of Patients
316M Patients–72M de-duped
Number of Providers
Over 600k
Provider Setting
Across all settings nationally
Years Available
Three years
Insurance Types Included
Commercial, Medicaid, and Medicare
Patient Sociodemographic Elements
Over 83 fields available–including ZIP code, birth date, sex, gender, ethnicity, race, marital status
Patient Clinical Data Elements
Includes Allergies, Appointments, Demographic, Diagnosis, Encounter, Immunization, Lab Order, Lab Result, Medication, Procedure, Social History, Vitals
Method of Data Access
Access via HTG's Hosted CD-19 RD Platform
Length of Data Access
Up to a 12 months (more depending on study)
Ability/Restriction on Merging Data Resources
Reviewed by data governance council and depending if expert determination is required. Tokenization available
Types of Data Arrangements Needed
Application, DUA and potential IRB

HealthShare Exchange (HSX)

Type of Data
Longitudinal Medical Records
Number of Patients
Approximately 10M patients across Southeastern Pennsylvania, Northern Delaware, and Southern New Jersey
Number of Providers
The majority of the hospitals in the Greater Philadelphia region, but also ambulatory care settings, skilled nursing/long-term care, specialty hospitals, home care agencies, and community support organizations
Provider Setting
Inpatient, Emergency Department, Outpatient, Skilled Nursing Facilities, Home Health Care
Years Available
2017–Present
Insurance Types Included
Commercial, Medicaid, Medicare, Uninsured
Patient Sociodemographic Elements
Demographics, ZIP codes, SDOH ICD-10 Codes
Patient Clinical Data Elements
All encounters, diagnoses, vital signs, lab results, imaging reports, immunizations, medications
Method of Data Access
Linked tables via SFTP
Length of Data Access
1-2 years subject to DUA
Ability/Restriction on Merging Data Resources
If the data to be linked with HSX is patient-level or otherwise indirectly identifying we may need to do the linking at HSX prior to anonymization
Types of Data Arrangements Needed
Data Use Agreements

Nebraska Healthcare Collaborative

Type of Data
Electronic Health Records
Number of Patients
10 million
Number of Providers
6,000
Provider Setting
Nebraska, Iowa
Years Available
2018–2023
Insurance Types Included
Any reported
Patient Sociodemographic Elements
Age, sex, location, ethnicity
Patient Clinical Data Elements
USCDI
Method of Data Access
SFTP or virtual environment
Length of Data Access
12 months
Ability/Restriction on Merging Data Resources
None
Types of Data Arrangements Needed
IRB approval, Data Governance Committee approval, Data Use Agreement

New Jersey Integrated Population Health Data (iPHD)

Type of Data
State of New Jersey linked administrative data from five sources: all-payer hospital inpatient and emergency department billing records, mortality and birth records, COVID-19 lab-verified case surveillance records, and emergency medical services records
Number of Patients
Over 20 million individuals
Number of Providers
Over 20,000 providers
Provider Setting
State of New Jersey hospitals and EMS services
Years Available
NJ birth data from 2000 to 2021, NJ mortality data from 2000 to 2021, NJ UB hospital discharge data from 2010 to 2021, NJ CDRSS COVID-19 surveillance data from 2020-2021, and NJ EMS data from 2017 to 2022. An additional year of birth, mortality, hospital discharge, and EMS data may be available at the time of award
Insurance Types Included
All (including uninsured)
Patient Sociodemographic Elements
Age, sex, race/ethnicity, ZIP code, marital status, language (varies by data source)
Patient Clinical Data Elements
Diagnostic, procedure, and cause of death codes
Method of Data Access
Secure portal or HIPAA compliant Limited Data Set for analysis on grantee servers if deemed secure
Length of Data Access
Duration of grant award, potentially extendable
Ability/Restriction on Merging Data Resources
Only de-identified linked data will be provided, researcher can merge on ZIP code-level characteristics, external data sources may be linked by the iPHD in some circumstances
Types of Data Arrangements Needed
Data use agreement and data management plan required

OCHIN Community Health Equity EHR Database

Type of Data
Outpatient electronic health records enhanced with mortality data, Medicaid claims, and geo-linked community-level social determinants of health indicators
Number of Patients
>6 million
Number of Providers
>27,000
Provider Setting
National network of ambulatory 'safety-net' community-based healthcare organizations providing primary care, behavioral health, dental, and other specialties
Years Available
2012–2023
Insurance Types Included
Medicaid, Medicare, commercial, uninsured
Patient Sociodemographic Elements
Sex, age, race, ethnicity, primary language, federal poverty level, sexual orientation, gender identity, unhoused status, migrant/seasonal worker status, veteran status, residential information (e.g., urban/rural, zip code, census tract)
Patient Clinical Data Elements
Encounters, diagnoses, procedures, vitals, lab results, medications, immunizations, patient-reported outcomes including social needs
Method of Data Access
Curated Limited Data Set meeting specific study's needs. Applicants may access a web-based Cohort Discovery tool pre-award for feasibility assessment
Length of Data Access
Duration of award
Ability/Restriction on Merging Data Resources
May be explored with OCHIN; linkages requiring direct identifiers must be done at OCHIN
Types of Data Arrangements Needed
Data release requires IRB review (approval or exemption) and Data Use Agreement; data will be transferred via SFTP and must be stored in an access-controlled encrypted environment

Truveta Studio in Partnership with Mathematica Data Innovation Lab

Type of Data
Full patient medical records linked across health systems from structured fields and augmented with SDOH, mortality, pharmacy, and claims data; grantees will receive a summary aggregated for specific populations at a geographic level

Number of Patients
~86M

Number of Providers
20,000+

Provider Setting
Hospital inpatient, outpatient, clinic, laboratory

Years Available
Minimum of 5 years of longitudinal clinical history, refreshed nightly

Insurance Types Included
All

Patient Sociodemographic Elements
Age, sex, race/ethnicity, county, SDOH

Patient Clinical Data Elements
Full records, claims, mortality, pharmacy, social drivers of health

Method of Data Access
Indirect access: Mathematica conducts analysis on Truveta Studio and delivers aggregate data to Grantee

Length of Data Access
Ability to derive cohort and specify analyses according to an approved protocol or analytic plan during the grant period.

Ability/Restriction on Merging Data Resources
Individual-level data are already linked in the platform. Can be merged to county- or cohort-level data

Types of Data Arrangements Needed
Data use agreement with Mathematica

Claims & Encounter Data

All Payer Claims Database–Colorado

Type of Data
Claims data
Number of Patients
8.6M unique patients over 10 years
Number of Providers
Approximately 64,000 primary care, specialists, MD, DO, RN, NP, PA
Provider Setting
All settings across Colorado
Years Available
2012–2022
Insurance Types Included
Commercial, Medicaid, Medicare Advantage, Medicare FFS (application dependent)
Patient Sociodemographic Elements
Age, gender, insurance coverage (including dual-eligible individuals), 3-digit ZIP. May be able to include ACS-based Social Vulnerability Index
Patient Clinical Data Elements
Procedures, diagnoses, prescription fills
Method of Data Access
Secure file transfer to requester
Length of Data Access
One year, then required data destruction. Extensions available at no cost or reduced cost (case dependent, researcher responsible for cost)
Ability/Restriction on Merging Data Resources
De-identified data set can be merged on NPI, Medicare provider ID, 3-digit ZIP. CIVHC may be able to merge on identifiable elements, de-identify output and deliver to recipient–dependent on complexity of merge and the other source of data. Dates of service or 5-digit ZIP may be available depending on the application. If approved, these fields can also be used by the requester for merger
Types of Data Arrangements Needed
Data Use Agreement with recipient and recipient organization. Institutional data management or privacy and security guide, and data storage plan are also required. CIVHC must review all research or papers prior to publication

All Payer Claims Database–Virginia

Type of Data
All Payer Claims Database (APCD)–Medical claims, Pharmacy claims, insurance enrollment information
Number of Patients
5M plus covered individuals per year
Number of Providers
Spans all providers of care for Virginia residents, 50,000+
Provider Setting
All settings including Hospital IP, ED, Urgent Care, ASC, Freestanding Physician Office
Years Available
2017–2022
Insurance Types Included
Commercial, Medicare, Medicaid
Patient Sociodemographic Elements
County, Zip, Age, Sex, Insurance status
Patient Clinical Data Elements
Diagnoses, procedure, prescriptions
Method of Data Access
Flat file transfer, data extraction/analytics tool depending on scope of data approval
Length of Data Access
Can exceed one year as long as proper justification is provided
Ability/Restriction on Merging Data Resources
Linkage based on provider, geographic elements, all factors except patient ID. Masked Patient ID is linked longitudinally within APCD but cannot be unhashed to link with outside data sources by individual patient
Types of Data Arrangements Needed
Data subscriber (use) agreement, approval by Virginia APCD Data Release Committee

Health Care Cost Institute (HCCI)

Type of Data
Medical and pharmacy claims data
Number of Patients
Over 50M covered lives per year
Number of Providers
Over 2M providers per year
Provider Setting
U.S.-wide care across all settings
Years Available
2012–2021
Insurance Types Included
Employer-Sponsored Health Insurance (ESI)
Patient Sociodemographic Elements
Age, sex, ZIP code
Patient Clinical Data Elements
Claims data
Method of Data Access
Secure data Enclave hosted by NORC at the University of Chicago
Length of Data Access
12 months, renewable at-cost (researcher responsible for cost)
Ability/Restriction on Merging Data Resources
Can merge by geography, National Drug Code, and limited merging by encrypted National Provider Identifier at-cost (researcher responsible for cost)
Types of Data Arrangements Needed
Data use agreement with PI's affiliated organization

Transformed Medicaid Statistical Information System (T-MSIS) Analytic Files (TAF), from Mathematica Data Innovation Lab

Type of Data
Aggregated Medicaid claims & CHIP data based on researcher specification
Number of Patients
500k deliveries/year and over 900k Home-and Community-Based Services (HCBS) users/year
Number of Providers
~ Over 500k across all service types
Provider Setting
Inpatient, professional, long-term care
Years Available
2019–2021; possibly 2022 based on release
Insurance Types Included
Medicaid only
Patient Sociodemographic Elements
Age, sex, race/ethnicity, county, eligibility group
Patient Clinical Data Elements
Claims data
Method of Data Access
Direct delivery to grantee
Length of Data Access
No limit
Ability/Restriction on Merging Data Resources
Can be merged to area-level characteristics
Types of Data Arrangements Needed
Data use agreement with Mathematica

Medicare Advantage (MA) Enrollment and Encounter Data

Type of Data
Medicare Advantage (MA) Enrollment and Encounter Data
Number of Patients
MA Enrollment data: 30M for MA enrollment; Encounter data: 24M
Number of Providers
1M
Provider Setting
Every hospital and facility with the exception of facilities that focus solely on specialty populations (e.g., Veterans Affairs) or facilities that do not accept the type of insurance (e.g., Assisted Living)
Years Available
Enrollment data: 2010 until current; Encounter data: 2015 until 2020 (with 2–3-year lag)
Insurance Types Included
Medicare Advantage
Patient Sociodemographic Elements
Geographic data (state, county, and CBSA), race, age, sex, dual status, and beneficiary-level MA plan/contract enrollment on a month-by-month basis
Patient Clinical Data Elements
Diagnostic (ICD-10 diagnosis codes), procedures performed (HCPCS, ICD-10 procedure, and MS-DRG codes), provider identification (organization TIN and NPI, practitioner NPI) for any relevant individual/organization on the claim
Method of Data Access
Aggregate data will be delivered in a machine-readable format (e.g., CSV) via email or file share
Length of Data Access
Up to 12 months
Ability/Restriction on Merging Data Resources
External data sources can be merged with any of the non-patient identifiers available within each table. These could include geographic information, physician and facility information, or any other number of identifiers

Transparency-in- Coverage Data

Clarify Health

Type of Data
Price Transparency Data
Number of Patients
N/A
Number of Providers
All providers within the USA referenced within the Payer released Machine Readable Files (MRFs) are represented. Currently, have over 350,000 unique provider NPIs in the data set
Provider Setting
Institutional and Professional rates are provided for both the inpatient and outpatient settings
Years Available
Payer-sourced rates data since the first CMS mandated release of July 2022 to the present day
Insurance Types Included
Products referenced in the MRFs are available, for example EPO, HMO, PPO, POS, Indemnity
Patient Sociodemographic Elements
N/A
Patient Clinical Data Elements
N/A
Method of Data Access
Data will either be delivered through a flat-file, or accessible through the Clarify Rates software described below
Length of Data Access
Access to the data, in the case of software subscription, will be available for the duration of the contract
Ability/Restriction on Merging Data Resources
No restrictions
Types of Data Arrangements Needed
Master Service License and Software Agreement and Order Form

Serif Health

Type of Data
Transparency In Coverage in-network negotiated rates data
Number of Patients
N/A–no patient data included
Number of Providers
~2.2M nationwide represented across payer released machine readable files (MRFs)
Provider Setting
U.S.-wide care across all settings covering both institutional and professional rate settings
Years Available
2022–2023, updated monthly
Insurance Types Included
Commercial–Group and Individual/Exchange Plans
Patient Sociodemographic Elements
N/A, no patient data included
Patient Clinical Data Elements
N/A, no patient data included
Method of Data Access
Secure file transfer to researcher
Length of Data Access
12 months, renewable at-cost (researcher responsible for cost of renewal)
Ability/Restriction on Merging Data Resources
No restrictions
Types of Data Arrangements Needed
Data use agreement with Serif Health

Turquoise Health

Type of Data
Payer-reported in-network negotiated rates data ("Transparency" data)
Number of Patients
N/A
Number of Providers
Distinct Provider Group IDS: 73,007; Distinct Provider NPIs: 86,416
Provider Setting
U.S.-wide hospitals, ambulatory health care facilities, hospital units, laboratories (full sub-segments in longer description)
Years Available
2023 data set, updated monthly
Insurance Types Included
Commercial and Individual/exchange
Patient Sociodemographic Elements
None
Patient Clinical Data Elements
None
Method of Data Access
Trino cloud instance where researcher can pull using SQL queries
Length of Data Access
12 months, extension up to 24 months or more available on case-by-case basis
Ability/Restriction on Merging Data Resources
Data can be merged on NPI or CPT/DRG. Merging requires approval by Turquoise on project-by-project basis
Types of Data Arrangements Needed
Data use agreement with Turquoise

Mathematica Data Innovation Lab

Type of Data
Transparency in Coverage (price information for group health plans and issuers of group or individual health insurance)
Number of Patients
N/A
Number of Providers
~ 1M
Provider Setting
U.S.-wide
Years Available
2022, 2023 may be available depending on timing of the request
Insurance Types Included
Commercial individual and group insurance plans
Patient Sociodemographic Elements
N/A
Patient Clinical Data Elements
N/A
Method of Data Access
Analytic file
Length of Data Access
12 months
Ability/Restriction on Merging Data Resources
No restrictions, possible add-ons include Medicare rates, hospital information, and provider details.
Types of Data Arrangements Needed
Data use agreement with Mathematica

The following data descriptions have been developed by the data providers. RWJF is not responsible for any inaccuracies in the descriptions.

Electronic Health Record & Clinical Data

American Health Associates

American Health Associates (AHA) is the #1 mobile diagnostics provider in the United States, serving 31% of Skilled Nursing Facilities (SNFs) in the country. AHA operates 20 full-service laboratories and employs nearly a thousand mobile phlebotomists who are dispatched daily. The patient population served by AHA is unique, as long-term-care facilities cater to a vulnerable, diverse, and frail demographic. AHA's data set covers millions of observations over many years, with an average of five observations per patient over time, linking demographic data with laboratory results. This data is ideal for studying national practice patterns in long-term care, physician characteristics associated with underuse or overuse of testing, diagnostic quality, diagnostic safety, public health surveillance, chronic condition monitoring, end-of-life care, and more. Under the HD4A CFP, AHA plans to share de-identified, longitudinal, multi-payer data sets of lab results with individual patient-level data to researchers, which may be linkable to other data sources with advance notice.

- A more detailed description of the data is available here.
- A video recording of the data set is available here.
- A data dictionary is available here.
- For questions about the AHA data, please contact Nirav Shah, *nshah@ahalabs.com*

COVID-19 Research Database

The COVID-19 Research Database is a research and public health community created by HHS Technology Group (HTG) and its research database partners in an attempt to provide real-world data and related software tools to researchers from any sector; including non-profit, grant-funded, and both industry and commercially funded research and public health activities concerning COVID-19. HHS Technology Group is an international software development company founded on the principles of agility, innovation, transparency, and predictability of results. The database is a partnership between HHS Technology Group, Datavant, Mathematica, Change, AnalyticsIQ, HealthJump, Office Ally, Snowflake, and many others.

We are releasing a common data schema that will allow researchers to access linkable data sets including claims, electronic health records and consumer data covering 316 million patients with 85 billion records.

Primary studies could focus on the effects of COVID-19 on specific cohorts of patients based on underlying conditions and demographics. For example, the effects of long COVID on Alzheimer and dementia patients or predicting the impact of COVID-19 on patients with chronic kidney disease.

Access to the data is provided via HTG's Discover Your Data Platform residing on AWS and leverages Snowflake.

- A more detailed description of the data is available here.
- A data dictionary is available here and here.
- A video recording of the data set is available here.
- For questions about the COVID-19 Research Database, please contact Brett Furst, *Brett.Furst@hhstechgroup.com*

HealthShare Exchange

HealthShare Exchange (HSX), a 501(c)(3) non-profit organization, is the regional Health Information Exchange (HIE) for the Greater Philadelphia area with a mission to provide secure access to health information that enables preventive and cost-effective care, improves the quality of care, and facilitates care transitions. HSX membership includes health systems, skilled nursing facilities, home health agencies, a mixture of both owned and independent practice groups, and several accountable care organizations. HSX has integrated with 35 electronic health records (EHRs) and 14,500 clinical providers, exchanging data on more than 10 million patients.

The data provided will be in the form of a HIPAA limited data set with linked tables using randomly assigned patient IDs. Data provided will be based on data in the HSX Clinical Data Repository (CDR), which covers longitudinal health records, including encounter-level data and clinical information about patients across the continuum of cases represented by HSX membership, across the Greater Philadelphia region. The CDR depicts a patient's longitudinal medical record across many contributing organizations.

The CDR contains information such as: patient demographics, medications, allergies, hospital inpatient visits, emergency department encounters, outpatient practice visits, immunizations, diagnoses, procedures, lab results, radiology reports, social history, vitals.

HD4A applicants will receive files extracted from the CDR and anonymized (unless the applicant is an HSX member). If files are anonymized, they will be de-identified with respect to both patients and providers and so will not be combinable with other data sets at this level. It may be possible to combine the data with other data sets about communities that include a field for ZIP code. If patient-level data are provided, only 3-digit ZIP codes will be included, consistent with HIPAA Safe Harbor guidelines. Applicants will need to sign a Data Use Agreement before receiving data. [Ref HIPAA Privacy Rule (45 CFR 164.514]

- A more detailed description of the data is available here.
- A data dictionary is available here.
- A video recording of the data set is available here.
- For questions about the HSX data, please contact Bill Marella, *bill.marella@healthshareexchange.org*

Nebraska Healthcare Collaborative

De-identified and limited data sets are curated from the statewide Health Information Exchange (HIE) for research and public health insights by the Nebraska Healthcare Collaborative, powered by CyncHealth. Most participating facilities in the HIE network permit data use for purposes of research through their Participation Agreements.

Through the HIE, CyncHealth enables the exchange of thousands of messages daily across more than 6,000 providers and connects over 10 million unique lives. CyncHealth uses the United States Core Data for Interoperability (USCDI) as a standardized set of health data classes and constituent data elements for interoperable health data exchange. A USCDI data class is an aggregation of data elements by use case (or common theme), while a data element is data defined by USCDI. USCDI themes include: allergies, plan of care, health concerns, immunizations, labs, clinical notes, clinical tests, demographics, procedures, vital signs.

- A more detailed description of the data is available here.
- USCDI data elements and updates on standards are available here.
- A video recording of the data set is available here.
- For questions about the Nebraska Healthcare Collaborative data, please contact Larra Petersen-Lukenda, *Lpetersen-lukenda@cynchealth.org*

New Jersey Integrated Population Health Data (iPHD)

For each awarded project, iPHD will create an integrated (with person-level linkage) limited data set from five administrative data sources maintained by the New Jersey Department of Health

(NJDOH): Birth records, mortality records, all-payer hospital billing records, COVID-19 surveillance records, and emergency medical services data. Over 20 million individuals are represented across these five data sources.

NJ Birth Data (2000–2021): A longitudinally standardized data set containing live birth data from birth certificates. Data elements include the setting and date/time of delivery, as well as health and demographic attributes of newborn and mother.

NJ Mortality Data (2000–2021): Immediate, intermediate, and underlying causes of death coded in ICD-10, as well as demographic attributes of the deceased and place of death. Includes New Jersey residents regardless of location of death.

NJ Uniform Billing (UB) Hospital Discharge Data (2010–2021): All-payer hospital billing records from inpatient discharges and in-person visits made to hospital emergency and outpatient departments including admission and discharge dates, encounter information, diagnoses, patient characteristics, and charges.

NJ COVID-19 Surveillance Data (2020–2021): A subset of data elements from NJ Communicable Disease Reporting and Surveillance System related to COVID-19 including dates of onset, lab results, and other relevant information collected from public health partners statewide, including local health departments.

NJ Emergency Medical Services (EMS) Data (2017–2022): Details of EMS responses to calls throughout New Jersey. Data elements include location, dates and times, type of transport (e.g., ambulance, helicopter), relevant medical history, and procedures delivered by the dispatched medical response team.

- A more detailed description of the data is available here.
- A data dictionary is available here.
- A video recording of the data set is available here.
- For questions about the iPHD data, please contact Manisha Agrawal, *magrawal@ifh.rutgers.edu*

OCHIN Community Health Equity EHR Database

OCHIN, a nonprofit healthcare innovation center with a core mission to advance health equity, operates the most comprehensive database on primary healthcare and outcomes of traditionally underserved patients in the United States. The OCHIN Community Health Equity EHR Database aggregates electronic health record (EHR) and social determinants of health (SDH) data representing >6 million patients from 170 health systems and 1,600 clinic sites across 33 states (4.6 million patients are 'active,' with a visit in the last 3 years). Approved projects can obtain access to up to 12 years of longitudinal OCHIN Epic ambulatory EHR data, which is research-ready on the PCORnet Common Data Model (CDM). EHR data are supplemented with national mortality data, Medicaid claims (limited states), and geo-linked community level SDH indicators.

Contributing health systems are outpatient community-based healthcare organizations, which deliver comprehensive, culturally responsive, high-quality primary care healthcare services for communities most impacted by health disparities. Most of OCHIN's clinical members are Federally Qualified Health Centers (FQHCs) and other federally supported health centers that provide services regardless of patients' ability to pay. In addition to primary care, health centers often provide on-site services such as dental, pharmacy, mental health, substance abuse treatment, and social work. OCHIN is a national leader in the collection of patient-level social needs data; the database contains nearly 2 million screenings in domains such as food insecurity, housing status, and financial strain.

Overall Inclusion:

Data years available: 2012–2023 (>170 million total encounters)

Patients with one or more ambulatory, telehealth, or dental visit on or after 1/1/2012

Excluded: Records from institutionalized patients and neonates

- A more detailed description of the data is available here.
- A data dictionary is available here.
- A video recording of the data set is available here.
- For questions about the OCHIN data, please contact Megan Hoopes, *hoopesm@ochin.org*

Truveta Studio in Partnership with Mathematica Data Innovation Lab

Truveta offers researchers AI-enabled, complete, timely, and representative EMR-derived clinical data from approximately 100 million covered lives from over 30 health systems, 20,000+ hospitals and clinics, and 50 states. Through a partnership with Mathematica's Data Innovation Lab, Truveta is offering indirect access to their Truveta Studio platform to allow Health Data for Action grantees to advance the mission of saving lives with data. More information on Truveta is available here: *https://www.truveta.com/*

To support grantees focused on specific use cases, Truveta is offering access through Mathematica to their structured clinical data that is linkable to SDOH, mortality, pharmacy, and claims data (see data dictionary). Applications are encouraged to submit a proposal on a topic of their choice but are highly encouraged to focus on one of two approved use cases: (1) equity, disparities, or SDOH-driven variations in health services use; or (2) research on maternity and newborn care. At the application phase, Mathematica will review proposals for analytic feasibility with Truveta. With a goal of peer-reviewed publication, grantees can specify the creation of one analytic cohort to study research questions; a cohort is defined by a set of inclusion and exclusion criteria based on diagnoses, procedures, or medications over a defined period of time. Upon award, Mathematica will review the grantee's study protocol or analysis plan; upon finalization, Mathematica will create the analytic file on the Truveta Studio platform and provide aggregate data to the grantee.

- A more detailed description of the data is available here.
- A data dictionary is available here.
- A video recording of the data set is available here.
- For questions about the Truveta data, please contact Constance Delannoy, *CDelannoy@mathematica-mpr.com*

Claims & Encounter Data

All Payer Claims Database–Colorado

The Center for Improving Value in Health Care (CIVHC) has administered the All Payer Claims Database for the state of Colorado (CO APCD) since 2012. Commercial insurers (with the exception of ERISA self-funded plans) and Medicaid submit claims monthly. Medicare Fee-for-Service claims are received quarterly from the Centers for Medicare and Medicaid Services state agency program. Medicare Part D claims are submitted annually with an approximate 18-month delay. The CO APCD houses claims data from 2012 to 2022 for approximately 8.6 million unique lives.

The CO APCD data set includes data at the de-identified patient level, de-identified payer code, and identified provider information. It contains comprehensive claims data including: facility information, charged and negotiated paid rates, insurance plan paid, and member liability (co-pays, and co-insurance). The data set includes year, but not month or day of service. We include a sequencing algorithm that allows researchers to determine length of stay, follow-up timing, and evaluate certain date dependent scenarios. If needed, CIVHC can provide dates of service or 5-digit ZIP code following additional compliance approvals.

Accessing the CO APCD data requires an application detailing research questions, descriptions of how the research addresses healthcare access, cost, or experience, and an explicit benefit to Coloradans (contributing to generalizable knowledge or potential remote benefits are not sufficient). CIVHC staff can assist applicants with the application if needed. We also require detailed data security protocols and a data management plan. Access to Medicare Fee-for-Service data is dependent on the specific application.

- A more detailed description of the data is available here.
- A data dictionary is available here.
- A video recording of the data set is available here.
- For questions about the CO APCD data, please contact Kimi Landry, *klandry@civhc.org*

All Payer Claims Database–Virginia

The Virginia All-Payer Claims Database (APCD) contains over 1 billion medical and pharmacy claims for roughly 5.5 million Virginia residents. Each claim and corresponding enrollment record submitted to the Virginia APCD includes information such as patient demographics, location of care across all settings, who provided care to the patient, any diagnoses presented by the patient, and the actual allowed amount or “cost” of a particular service. The database includes claims and enrollment data for the following coverage types–all Medicaid Fee-For-Service (FFS) and Managed Care, all Commercial Fully Insured, all Commercial non-ERISA self-insurance and partial ERISA self-insured. Patient data is linked across health insurance carriers and over time and can easily be used to construct episodes of care. Data is available for the timeframe of claims paid between 2016 and 2022.

- A more detailed description of the data is available here.
- A data dictionary is available here.
- A video recording of the data set is available here.
- For questions about the VA APCD data, please contact Nicole Fawcett, *nicole@vhi.org*

Health Care Cost Institute

The Health Care Cost Institute’s (HCCI) multiyear data, which is de-identified and HIPAA compliant, includes the healthcare claims for over 50 million members per year with health insurance through their employer for the years 2012 to 2021.

These data—contributed by Aetna, Humana, and Blue Health Intelligence—consist of fully adjudicated, paid, de-identified medical and pharmacy claims, including patients’ and providers’ geography, provider specialty and encrypted ID, service dates, diagnostic (ICD-9 or ICD-10) and procedure codes (MS-DRG, CPT, HCPCS), prescription national drug codes (NDC), and allowed amounts plus any co-payments, deductibles, or co-insurance paid by the insured person. Members are assigned a unique identifier and can be tracked over time within an insurer.

The HCCI data set is available to researchers in one of two views: data view 1 includes member year of birth and the unit of geography is CBSA; data view 2 includes age band (e.g., 18–34) and five-digit ZIP code. Members' demographic information is limited to age and sex. Preferred data view should be specific in the proposal.

Researchers who wish to merge external data files into the HCCI data may do so with approval from HCCI. Providers are identified in the HCCI data by encrypted National Provider Identifiers (NPIs). External data files that have NPIs can be encrypted for merging with the HCCI data for an additional fee, for which the research team will be responsible. Planned merges and/or comparisons between the HCCI data and other data sources must be clearly specified in research proposals.

Researchers at academic and noncommercial policy institutions are eligible to license the HCCI data. Researchers based at or affiliated with health insurance plans, consulting companies, or provider organizations are not eligible.

- A more detailed description of the data is available here.
- A data dictionary is available here.
- A video recording of the data set is available here.
- For questions about the HCCI data, please contact Aditi Sen, *asen@healthcostinstitute.org*

Medicare Advantage Enrollment and Encounter Data

CareJourney would provide access to a variety of research files in a machine-readable format. These would be data that were aggregated within the Virtual Research Data Center (VRDC) and subsequently exported for external research purposes. The data are derived from Medicare Advantage (MA) enrollment and encounter data. All the available data are linked longitudinally and include all MA enrolled beneficiaries (~30 million in 2022). The enrollment data includes geographic data (beneficiary state, county, and CBSA), demographic information (race, age, sex, dual status, and various socio-economic factors) and beneficiary level MA plan/contract enrollment on a month-by-month basis. The encounter data includes all MA encounters that pass through the traditional healthcare system. This includes encounters with hospitals, physicians, post-acute care facilities, durable medical equipment, and hospice.

We have run multiple analyses against these enrollment and encounter data and exported them in a way that would make them directly accessible to researchers via machine readable files. Additional tables are likely to become available.

These are not patient-level tables. Rather, they are aggregated to various levels. These tables include, but are not limited to:

- Number of beneficiaries attributed by Provider and Plan.
- Trends in MA enrollment.
- Physicians and organizations that patients are seeing.
- MA provider trends in diagnoses and procedures counts.
- Changes in HCC scores and diagnoses following updates in the MA rate notice, from v24 HCCs to v28 HCCs.
- MA Switchers (churn) between FFS and MA, as well as between MA plans.
- Opportunity landscape to understand variation in key performance and utilization metrics among various patient populations.

These tables, independently and collaboratively, provide a wealth of information that could be used for a number of important research questions.

- A more detailed description of the data is available here.
- A data dictionary is available here.
- A video recording of the data set is available here.
- For questions about the Medicare Enrollment and Encounter data, please contact Nate Smith, *nathan.smith@carejourney.com*

Transformed Medicaid Statistical Information System (T-MSIS) Analytic Files (TAF), from Mathematica Data Innovation Lab

Mathematica is a mission-driven organization offering aggregate data to support research questions on geographic variations in care for two populations: **Maternity and Home-and Community-Based Services (HCBS).**

The aggregate files are derived from the Transformed Medicaid Statistical Information System (T-MSIS) Analytic Files (TAF) from 15 states with usable data during the 2019 through 2021 time period; depending on timing, summaries from 2022 may be available to grantees. The resulting data delivered to grantees at the most granular level is at the county-year-subpopulation level.

Leveraging the analytic file described in the data dictionary, Mathematica will help grantees develop aggregate data for their research question based on their specification, as well as the feasibility of the analysis given the data. Grantees are encouraged to review the Medicaid Data Quality Atlas before applying to understand the limitations of the data. Grantees can specify inclusion/exclusion criteria and subgroups to produce summaries of utilization. The underlying analytic file includes information on beneficiary demographics, program eligibility and enrollment, health conditions based on AHRQ's Clinical Classifications Software Refined for Diagnoses (CCSR), and health services use (e.g., inpatient stays). Furthermore, the maternity cohort has information on severe maternal morbidity and type of delivery, and the HCBS and ID/DD populations have indicators on the use of HCBS and potentially avoidable hospitalizations. During the application process, Mathematica will assess feasibility of requested data based on subgroup sample size and data quality.

- A more detailed description of the data is available here.
- A data dictionary is available here.
- A video recording of the data set is available here.
- For questions about the T-MSIS data, please contact Constance Delannoy, *CDelannoy@mathematica-mpr.com*

Transparency in Coverage Data

Applicants requesting Transparency in Coverage (TiC) data should select “Transparency in Coverage Data” in the application to be assigned to one of the following four data providers after submission. If an applicant prefers one of the following TiC providers, that can be indicated in the application. To help applicants understand the different TiC offerings, the data providers below will hold office hours in July and are available for questions. The dates/times and sign-up information for each office hours session are included below.

Clarify Health

Clarify has an expansive patient-level data set of over 300+ million lives coupled with the most advanced technologies to ensure accuracy in data published by payer file sources. This has allowed us to create a solution that delivers insights from a geographic level down to procedural level with varying options to cut analysis based on user preference. Moreover, we have intentionally built our capabilities to be transferable and flexible so that each customer can work directly with the data in an iterative way to derive the deepest insights available.

Our data ingestion process goes through rigorous data integrity checks, normalization and standardization applications and is cross-checked with our vast healthcare claims database for feasibility. This key differentiator ensures outliers are addressed and that researchers or analysts that depend on this data can trust the analytics. The intuitive platform enables on-the-fly analysis, dashboarding, and export capabilities that provide our clients with the flexibility that is required to address the plethora of use cases that emerge in healthcare entities.

This data is available to our customers in flat-file or through our SaaS product, Clarify Rates.

- A more detailed description of the data is available here.
- A data dictionary is available here.
- A video recording of the data set is available here.
- Office hours will be held on July 12, 3:00-4:00 p.m. ET, sign up here.
- For questions about the Clarify Health data, please contact Ryan Taylor, *ryan.taylor@clarifyhealth.com*, and Andrew Garner, *andrew.garner@clarifyhealth.com*

Serif Health

Serif Health is a healthcare data and analytics firm focused on enabling customers across the healthcare ecosystem–from providers, employers, payers to digital health technology firms–easy access to healthcare price transparency data via web app, flat file delivery, and API products.

Serif Health provides access to a national repository of Transparency in Coverage (TiC) price transparency data. TiC data consists of negotiated rate data between insurance payers and in-network providers for commercial fully-insured, self-insured, and marketplace insurance plans publicly posted by payers as machine-readable files (MRF), updated monthly.

Serif Health's complete, aggregated data set contains TiC postings for 150+ payers and over 400 unique plan networks across all 50 states, covering approximately 2.5 million provider entities. Plans and networks available represent over 100 million covered lives. There are no restrictions on joining the data to outside data sources. The Serif Health data includes standardized fields with the payer name, network name and type, provider details (EIN, NPI, taxonomy, location), billing/procedure code (CPT/HCPS, DRG, Revenue Codes), place of service and billing modifiers, rate type (fee schedule, negotiated, percentage of billed charges, etc.), and negotiated rate.

However, given the significant data volumes and file sizes of the MRF data, researchers can get more utility by focusing their proposals to a single state or geographic region (e.g., Washington, Texas, Pacific Northwest), specific payers, or a specific set of procedure codes or specialties.

For HD4A proposals, Serif Health will license data directly to academic institutions or research-focused organizations. Successful applicants are required to complete a Research Data Services Agreement granting access to the Serif Health data for twelve (12) months following file delivery. At the end of 12 months, successful applicants will have the option to renew their data license (at cost to the researcher).

- A more detailed description of the data is available here.
- A data dictionary is available here.
- A video recording of the data set is available here.
- Office hours will be held on July 13, 1:00-2:00 p.m. ET, sign up here.
- For questions about the Serif Health data, please contact Matthew Robben, *matt@serifhealth.com*

Turquoise

Turquoise has aggregated price transparency data from machine-readable files published by Hospitals and Payers in pursuant to new federal regulations called Hospital Price Transparency and Transparency in Coverage. Turquoise has built a data set of negotiated rates data for over 5,300 Hospitals and over 200 Payers. The data from Hospitals includes billed charges, discounted cash rates, and negotiated rates. The data from Payers includes in-network negotiated rates and out-of-network allowed amounts for all contracted providers. Turquoise offers a consolidated data platform that can empower researchers focused on healthcare organizations, including hospitals, insurers, employers, and others, to analyze healthcare prices. Our mission is to eliminate the financial complexity of healthcare.

For HD4A, Turquoise is providing a clean data set from the Payer Machine Readable Files. The data set is organized at the individual rate level (CPT or DRG) in a flat table called “Core Rates”. The flat table is updated monthly and contains institutional and professional rates for institutional providers and the top payers in each state. The Core Rates data set includes rates for Hospitals, Ambulatory Surgery Centers (ASCs), imaging centers, labs etc. Individual practitioners (physicians/APPs/etc) are not included. The table includes information on the payer (name, network product) negotiated rate amount, description and type, modifiers including place of service or billing code, and provider information (address, NPI or EIN, taxonomy)

Researchers will be able to quickly navigate the payer data using SQL without having to run complex queries.

- A more detailed description of the data is available here.
- A data dictionary is available here.
- A video recording of the data set is available here.
- Office hours will be held on July 21, 3:00-4:00 p.m. ET, sign up here.
- For questions about the Turquoise Health data, please contact Gus Roman, *gus@turquoise.health*

Mathematica Data Innovation Lab

At Mathematica, we use data, analytics, and technology to address pressing social challenges, from the effects of climate change on communities around the world to disparities in healthcare, education, and employment across the U.S. We're an employee-owned and mission-driven company, with a deep bench of expertise in both data and social science.

Since 2022, Mathematica has downloaded, processed, and analyzed health plan data following the Transparency in Coverage rule that came into effect July 1, 2022. This rule requires that most group health plans and health insurance issuers disclose price and cost-sharing information to participants, beneficiaries, and enrollees. The purpose of this rule is to "reduce the secrecy behind healthcare pricing with the goal of bringing greater competition to the private healthcare industry."[1]

Under the HD4A CFP, Mathematica is making available health plan pricing data for 2022, and possibly 2023 depending on the timing of the project. Data will be research-ready and will cover a sample of regional plans and may be tailored to specific regions, providers, and billing codes of interest.

In addition, data may include several options for merge-on variables, such as provider details available in NPPES[2], group-practice, hospital and system-level information, and billing-code level data such as Medicare rates. Additional linkages are possible to COVID RDB data and to Medicare claims.

Proposals may focus on a range of topics, including, but not limited to, the following: rate comparisons between commercial and Medicare rates for specific billing codes and plans; rate comparisons in between plans for specific billing codes; rate comparisons within a plan across different providers; rate comparisons across plans in a specific region for distinct billing codes.

- A more detailed description of the data is available here.
- A data dictionary is available here.
- A video recording of the data set is available here.
- Office hours will be held on July 20, 2:00-3:00 p.m. ET, sign up here.
- For questions about the Mathematica data, please contact Constance Delannoy, *CDelannoy@mathematica-mpr.com*

FUNDING DETAILS

Data Access Award:

- **Type of Award:** Awards funded under this opportunity will be structured as grants. Successful applicants will be given access to the selected data set at no cost to the grantee.
- **Number of Awards:** Up to 20 awards will be funded solely for data access, and no specific breakdown by data set is prescribed.
- **Amount of Award:** No financial award will be given with these study awards.
- **Award Duration:** Data access will accommodate studies of up to 12–24 months. Data providers have varying lengths for data access (described herein) that should be considered by applicants in determining their project lengths.

First-Time Investigator Awards:

- **Type of Award:** Awards funded under this opportunity will be structured grants. Applicants who are first-time investigators on a grant award at the time of application will be awarded funds to conduct studies using the selected data set, and no specific breakdown by data set is prescribed. Successful applicants will be given access to the selected data set at no cost to the grantee.
- **Number of Awards:** Up to six First-Time Investigator Awards will be funded.
- **Amount of Award:** Up to $600,000 will be available, with recommended project funding up to $100,000 per project.
- **Award Duration:** Data access and project funding will accommodate studies of up to 12–24 months. Data providers have varying lengths for data access (described herein) that should be considered by applicants in determining their project lengths.

- **Use of Funds:** Grant funds for First-Time Investigator Awards may be used for project staff salaries, consultant fees, data collection and analysis, meetings, supplies, project-related travel, other direct project expenses, including a limited amount of equipment essential to the project, and indirect costs to support the applicant organization’s general operations. In keeping with RWJF policy, funds may *not* be used to support clinical trials of unapproved drugs or devices, to construct or renovate facilities, and for lobbying or political activities. Additional budget guidelines are provided in the online application materials.

- **Payment of Awards:** Payment of the First-Time Investigator Awards will be based upon spending against approved budgets or milestones (not invoices).

The grant opportunity outlined in this call for proposals is contingent upon final award confirmation from the Robert Wood Johnson Foundation for such grants.

ELIGIBILITY CRITERIA

- Applicant organizations must be based in the United States or its territories.
- Preference will be given to applicants that are either public entities or nonprofit organizations that are tax-exempt under Section 501(c)(3) of the Internal Revenue Code and are not private foundations or Type III supporting organizations. The Foundation may require additional documentation.
- Projects may be generated from disciplines including health services research; economics; sociology; program evaluation; political science; public policy; public health; public administration; law; business administration; or other related fields.
- Use of the data under this CFP is restricted to noncommercial purposes.
- Consistent with our Equity, Diversity, and Inclusion Commitment, we particularly encourage applications with project directors or team members from backgrounds historically underrepresented in research disciplines as a result of their race, ethnicity, socioeconomic status, disability, or other factors.

First-Time Investigator Awards

- Applications for the First-Time Investigator Awards will be restricted to an application from an eligible organization with a principal investigator who is a first-time investigator on a grant award at the time of application.
- First-Time Investigators should not have previously served as a principal investigator on any grant award (after completion of their terminal research degree).
  - Principal investigators from backgrounds historically underrepresented in research disciplines are encouraged to apply.
  - A co-principal investigator on the application may have previously served as a principal investigator or co-principal investigator on a grant award before.

SELECTION CRITERIA

Applicants will be evaluated on the following:

- Importance, significance, and policy relevance of the question to be addressed.
- Strength of the proposed methodology.
- Innovative use of the available data.
- Strength, expertise, and qualifications of the full research team.
  - For the First-Time Investigator Award applicants, in addition to the principal investigator, the research team could include at least one senior researcher with similar or complementary research interests to serve as a research mentor and collaborator on the project.
- Ability to inform decisionmakers and other key players (e.g., policymakers, academia, industry) and create timely deliverables for wide dissemination (e.g., issue briefs, webinars, blog posts, fact sheets), in addition to papers suitable for peer-reviewed publication.
- For Data Access Awards: demonstrated resources to carry out the project within the proposed timeline.

OPEN ACCESS

In order to ensure RWJF supported research is made accessible to a wide and diverse audience, grantees who publish findings in peer-reviewed publications must do so in open access journals. To create open access at no cost, grantees can “self-archive,” by depositing a free copy of an electronic document online, either in the author's own institutional repository or another open archive. Visit SHERPA/RoMEO for more information on self-archiving.

HOW TO APPLY

Proposals for this solicitation must be submitted via the RWJF online system. Visit www.rwjf.org/cfp/hd4a4 and use the “Apply Online” link. If you have not already done so, you will be required to register at http://my.rwjf.org before you begin the proposal process.

There are two phases in the competitive proposal process:

Phase 1: Brief Proposals

Applicants must submit a brief proposal of no more than four pages that describes the project. **Applicants should not begin drafting their proposal until they have familiarized**

themselves with the detailed instructions and formatting requirements displayed in the online application system.

The brief proposal should contain the following information about the proposed project, and any specific content requested in the above data descriptions:

- Research question and potential of the study to generate significant and new knowledge to inform health policy.
- Methodology and demonstrated ability to use the available data sources.
- Deliverables and plan for dissemination.
- The applicant's qualifications and expertise.

Brief proposals will be reviewed by staff at RWJF, AcademyHealth, and the data providers.

Phase 2: Full Proposals

Selected Phase 1 applicants will be invited via email to submit a full proposal of no more than 10 pages. Data access only applicants will also need to complete a narrative that describes the research team's time allocation and resources available to support the project. First-Time Investigator Award applicants will need to provide a line-item budget and budget narrative.

Invited full proposals will undergo peer review by external, subject matter experts, as well as staff at RWJF, AcademyHealth, and the data providers. RWJF will make all final grant decisions.

Please note that the data providers may require that successful applicants complete additional materials to gain access to the data sets as described above.

One informational webinar for applicants will be held on **July 18, 2023, from 1–2pm ET** to provide an overview of the CFP and offer an opportunity for potential applicants to ask questions about the program, the funding opportunity, and the grantmaking process. This webinar will be archived and available for review on the AcademyHealth website.

Questions about the available data or participating data providers will not be answered on the applicant webinar. Applicants are encouraged to view the video recordings hyperlinked in the above descriptions prior to reaching out to the data providers with any specific questions not addressed in the CFP, the longer data descriptions, or the recorded overviews. Specific questions about the available data sets should be directed to the contacts named in the above descriptions.

Please direct all other inquiries to *healthdataforaction@rwjf.org.* Be sure to include your phone number. We will make every effort to respond to all inquiries within 24 hours.

All applicants should log in to the system and familiarize themselves with online submission requirements well before the final submission deadline. Staff may not be able to assist all applicants in the final 24 hours before the submission deadline.

PROGRAM DIRECTION

Direction and technical assistance for this program are provided by AcademyHealth, which serves as the national program office located at:

Health Data for Action
AcademyHealth
1666 K Street NW, Suite 1100
Washington, DC 20006

Email: *healthdataforaction@rwjf.org*
Website: *https://academyhealth.org/*

Responsible staff members at the national program office are:

- Megan Collado, MPH, *senior director (Program Director)*
- Rishika Desai, MPH, *senior manager*
- Maura Dugan, *research associate*

Responsible staff members at the Robert Wood Johnson Foundation are:

- Katherine Hempstead, PhD, *senior policy adviser*
- Alonzo Plough, PhD, MPH, MA, *vice president*
- Brian Quinn, PhD, *associate vice president*
- Isabelle Gerard, MPH, *communications officer*
- Mateusz Szalda, CPA, MAcc, *senior program financial analyst*

KEY DATES AND DEADLINES

- **June 21, 2023**

Call for proposals released.

- **July 18, 2023 (1 p.m.–2 p.m. ET)**
  Optional applicant informational webinar. Registration is required.
- **August 18, 2023 (3 p.m. ET)**
  Deadline for receipt of brief proposals.*
- **November 3, 2023**
  Applicants notified whether invited to submit full proposals.
- **January 12, 2024 (3 p.m. ET)**
  Deadline for receipt of full proposals.*
- **May 1, 2024**
  Notification of finalists.
- **June 15, 2024**
  Grants start.

Applicant Deadline Policy

*All proposals for this solicitation must be submitted via the RWJF online system. Visit *www.rwjf.org/cfp/hd4a4* and use the “Apply Online” link. If you have not already done so, you will be required to register at *my.rwjf.org* before you begin the proposal process.

All applicants should log in to the system and familiarize themselves with online proposal requirements well before the final submission deadline. Please note that, in the 24-hour period leading up to the proposal deadline, staff may not be able to assist all applicants with any system-related issues. Therefore, we encourage you to submit your proposal well before the deadline so that any unforeseen difficulties or technical problems may be addressed in advance.

Late submissions will not be accepted for any reason. While late submissions will not be accepted, RWJF may choose, at its sole discretion, to extend the proposal deadline for all applicants. Such extensions generally will be granted only in the event of: (1) a verified issue with the RWJF proposal system that prevented completion and submission of proposals; or (2) a disaster, emergency, or significant internet outage that affects one or more regions. For purposes of this policy, a region is generally considered to be one or more states. RWJF strives to give all applicants any support needed to successfully submit their proposal prior to the deadline. Submission is defined as all sections completed, marked finished, the proposal "submit" button used, and the proposal status shows "Submitted." If the deadline is extended for any reason, the extension will be posted on the funding opportunity page at *www.rwjf.org*. In addition, an email will be sent to all individuals that have started a proposal in the RWJF online system.

MONITORING

RWJF monitors the grantees' efforts and careful stewardship of grant funds to assure accountability. Grantees will be required to submit periodic narrative reports. First-Time Investigator Award recipients will be required to submit periodic financial reports, in addition to periodic narrative reports.

APPLICANT SURVEY PROCESS

The principal investigator of the proposal may be contacted after the submission deadline by SSRS, an independent research firm. The principal investigator will be asked to complete a brief, online survey about the proposal process and applicant characteristics. This voluntary questionnaire will take no more than 15 minutes to complete. Responses provided to SSRS will not impact the funding decision for your proposal in any way.

SSRS will protect the confidentiality of your responses. RWJF will not receive any data that links your name with your survey responses.

OUR EQUITY, DIVERSITY, AND INCLUSION COMMITMENT

The Robert Wood Johnson Foundation is committed to building a Culture of Health that provides everyone in America a fair and just opportunity for health and well-being. Achieving this goal requires focus on equity, diversity, and inclusion. To that end, we are committed to fostering diverse perspectives. We recognize that individuals' perspectives are shaped by a host of factors, such as their race, ethnicity, gender, physical and mental ability, age, socioeconomic status, gender identity and expression, sexual orientation, familial status, education, religion, legal status, military service, political affiliation, geography, and other personal and professional experiences.

We know that the presence of diverse perspectives alone is not sufficient. Therefore, we also are committed to creating inclusive environments where all individuals are encouraged to share their perspectives and experiences. We believe that only through valuing our differences and similarities, and remaining vigilant in advancing equity, will we be able to maintain an equitable workplace and actively pursue equity in all aspects of our work. We commit to being continuous learners and working alongside others to cultivate equity, diversity, and inclusion.

ABOUT THE ROBERT WOOD JOHNSON FOUNDATION

The Robert Wood Johnson Foundation (RWJF) is committed to improving health and health equity in the United States. In partnership with others, we are working to develop a Culture of Health rooted in equity that provides every individual with a fair and just opportunity to thrive, no matter who they are, where they live, or how much money they have. For more information, visit *www.rwjf.org*.

Sign up to receive email alerts on upcoming calls for proposals at *www.rwjf.org/manage-your-subscriptions.html*

50 College Road East
Princeton, NJ 08540-6614