Commonwealth of Massachusetts
Center for Health Information & Analysis (CHIA)
Non-Government Application for MA APCD Limited Data Set
[Exhibit A: Data Application]

This form is required by all Applicants, except Government Agencies as defined in 957 CMR 5.02. All Applicants must also complete the Data Management Plan, attached to this Application. The Application and the Data Management Plan must be signed by an authorized signatory of the organization. This Application and the Data Management Plan will be used by CHIA to determine if your organization may receive CHIA data. Please be sure the documents are completed fully and accurately. You may wish to consult the Evaluation Guide that CHIA will use to review your documents. Prior to receiving CHIA Data, the organization must execute the Data Use Agreement. You may wish to review that document as you complete these forms. This application should be completed by the Primary Investigator, and must be signed by a party with authority to bind the organization seeking CHIA Data for the purposes described herein.

NOTE: In order for your Application to be processed, you must submit the required application fee. Please consult the fee schedules for MA APCD data for the appropriate fee amount. A remittance form with instructions for submitting the application fee is available on the CHIA website.

All attachments must be uploaded to IRBNet with your Application. All applications documents can be found on the CHIA website in Word and/or PDF format.

I. GENERAL INFORMATION

APPLICANT INFORMATION	
Applicant Name: (Primary Investigator)	Naomi Bardach
Title:	Associate Professor
Organization Requesting Data: (Recipient)	The Regents of the University of California, on behalf of its San Francisco campus
Project Title:	IMPLEMENT Aim 1
IRBNet ID:	975842-1
Address, City/Town, Zip Code	3333 California St. Suite 265, SF CA 94118
Telephone Number:	415-476-9188
Email Address:	Naomi.bardach@ucsf.edu
Names of Co-Investigators:	Michael Cabana (PI); Chuck McCulloch (biostats)
Email Addresses of Co-Investigators:	michael.cabana@ucsf.edu; charles.mcculloch@ucsf.edu
Original Data Applicant Submission Date:	12/06/2016
Dates Data Application Revised:	
Project Objectives (240 character limit):	We will test new pediatric quality measures for sickle cell disease, asthma, and suicidality. By understanding how to refine and apply these measures, we can potentially improve pediatric care and outcomes for children with these diseases.
Project Research Questions (if applicable) or Business Use Case(s):	1. What is the variation in performance on these measures? At several levels of clustering: state level, county, state region, health plan, 3-digit zip code,

	hospital service area, medical group, clinic, and individual physician level. 2. Assess whether performance differs by patient or cluster characteristics.

II. PUBLIC INTEREST & PROJECT SUMMARY

1. Briefly explain why completing your project is in the public interest.

This work is part of a larger project testing and implementing pediatric quality measures, to improve the health care delivered to children, with the goal of improving their health. The populations we focus on (asthma, sickle cell, patients with suicidality) are high risk populations and the measures are potentially important to their health.

2. Has an Institutional Review Board (IRB) reviewed your project?

☒ Yes, a copy of the *approval letter* and *protocol* must be included with the application package on IRBNet.
☐ No, this project is not human subject research and does not require IRB review.

3. **Research Methodology**: Applicants must provide a written description of the project methodology (typically 1-2 pages), which should state the project objectives and/or identify relevant research questions. This document must be included with the application package on IRBNet, and must provide sufficient detail to allow CHIA to understand how the data will be used to meet objectives or address research questions. Applications that do not include this methodology statement cannot be reviewed or approved.

III. DATA FILES REQUESTED

1. Please indicate the MA APCD databases from which you seek data, the year(s) of data requested, and your justification for requesting *each* file. Please refer to the MA APCD Release Data Specifications for details of the file contents.

MA ALL-PAYER CLAIMS DATABASE FILES	Year(s) Of Data Requested Current Yrs. Available ☐ 2011 ☐ 2012 ☒ 2013 ☒ 2014 ☒ 2015
☒ **Medical Claims**	**Please describe how your research objectives require Medical Claims data:** We need patient level medical claims data to assess whether patient care met the standard of care, assessed using the quality measures we will be testing.
☒ **Pharmacy Claims**	**Please describe how your research objectives require Pharmacy Claims data:** We need patient level pharmacy claims data to assess whether patient care met the standard of care, assessed using the quality measures we will be testing.
☐ **Dental Claims**	**Please describe how your research objectives require Dental Claims data:** NA
☒ **Member Eligibility**	**Please describe how your research objectives require Member Eligibility data:** We need member eligibility files to determine patient eligibility for the quality metric. This is part of the definition of the quality measure.

☒ **Provider**	**Please describe how your research objectives require Provider data:** We need provider files to assess variations in provider performance on quality measures, across all eligible patients for the individual provider or cared for at the site of care (clinic, hospital, etc).
☒ **Product**	**Please describe how your research objectives require Product data:** We need product data to better understand the quality of care for patients cared for under different insurance products, in order to better understand the variation in care across insurance products and plans.

IV. GEOGRAPHIC DETAIL

Please choose *one* of the following geographic options for MA residents. *For releases with 5 digit zip code, CHIA will apply a substance abuse filter which will remove all claims that include a substance abuse diagnosis.*

☒ 3 Digit Zip Code (MA) (standard)	☐ 5 Digit Zip Code (MA)***
***Please provide justification for requesting 5 digit zip code. Refer to specifics in your methodology:**	

V. DATE DETAIL

Please choose *one* option from the following options for dates:

☐ Year (YYYY) (Standard)	☐ Month (YYYYMM) ***	☒ Day (YYYYMMDD) *** [for selected data elements only]
***If requested, please provide justification for requesting Month or Day. Refer to specifics in your methodology:** Measure eligibility relies upon whether the patient has been a member for at least 30 days (see attached technical specifications). In order to assess this, we need to obtain the month and day of enrollment in relation to the month of the qualifying event.		

VI. NATIONAL PROVIDER IDENTIFIER (NPI)

Please choose *one* of the following options for National Provider Identifier(s):

☐ Encrypted National Provider Identifier(s) (standard)	☒ Unencrypted National Provider Identifier(s)***
***If requested please, provide justification for requesting unencrypted National Provider Identifier(s). Refer to specifics in your methodology:** For our second research question, we need to link claims data to external data (e.g., link to the National Plan and Provider Enumeration System) in order to assess the relationships between provider characteristics and measure performance.	

VII. MEDICAID DATA

Please indicate here whether you are seeking Medicaid Data:

☒ Yes
☐ No

Federal law (42 USC 1396a(a)7) restricts the use of individually identifiable data of Medicaid recipients to uses that are directly connected to the administration of the Medicaid program. If you are requesting Medicaid data from Level 2 or above, please describe, in the space below, why your use of the data meets this requirement. Requests for Medicaid data will be forwarded to MassHealth for a determination as to whether the proposed use of the data is directly connected to the administration of the Medicaid program.

This work is funded through the Centers for Medicare and Medicaid services as part of the Pediatric Quality Measurement Program, mandated in the federal CHIPRA legislation of 2009. The RFA for the grant specifically calls for work with the Medicaid population, which is the population of interest in our proposal. The proposal tests feasibility of pediatric quality measurement in the administrative data (Aim 1) as well as ability to improve those measures in the Medicaid population (Aim 2). Aim 3 proposes to submit the measures to the National Quality Forum for use in the Measure Applications Partnership, which chooses measures for potential use by CMS in the Medicaid and Medicare programs. Hence the work we will be doing will be relevant to Medicaid program administration, in relationship to quality measurement and quality improvement programs.

VIII. DATA LINKAGE AND FURTHER DATA ABSTRACTION

Note: Data linkage involves combining CHIA data with other databases to create a more extensive database for analysis. Data linkage is typically used to link multiple events or characteristics within one database that refer to a single person within CHIA data.

1. Do you intend to link or merge CHIA Data to other datasets?

   ☒ Yes
   ☐ No linkage or merger with any other database will occur

2. If yes, please indicate below the types of database to which CHIA Data be linked. [Check all that apply]
   ☐ Individual Patient Level Data (e.g. disease registries, death data)
   ☒ Individual Provider Level Data (e.g., American Medical Association Physician Masterfile)
   ☒ Individual Facility Level Data level (e.g., American Hospital Association data)
   ☒ Aggregate Data (e.g., Census data)
   ☐ Other (please describe):

3. If yes, describe the data base(s) to which the CHIA Data will be linked, which CHIA data elements will be linked; and the purpose for the linkage(s):

Provider-level data: We will use the NPPES, and NPI number to link data from NPPES to providers in the database.
Individual Facility level data: We will use the AHA data to link to individual facilities, using the Medicare provider ID or the NPI, depending on which has the least amount of missing data in both datasets.
Aggregate data: 1) We will use the American Community Survey, linked to zip codes, to assess for community factors associated with performance. We will use patient zip code to create the linkage. 2) We will use the Dartmouth Atlas Hospital Service Areas to group member records across zip codes, using the crosswalk between the HSAs and member zip codes.

4. If yes, for each proposed linkage above, please describe your method or selected algorithm (e.g., deterministic or probabilistic) for linking each dataset. If you intend to develop a unique algorithm, please describe how it will link each dataset.

We will use a deterministic approach for each linkage, using the individual variable to link records. We will do a cross check on a random sampling of observations for the dataset for which there are more than one variable that overlaps (e.g., the AHA data has hospital zip code and name, so we can check after matching on Medicare provider ID whether the zip code and name match those listed in the APCD data).

5. If yes, please identify the specific steps you will take to prevent the identification of individual patients in the linked dataset.

We will not be using any patient-level individually linkable variables, such as DOB or SSN. We will only be using a de-identified unique identifier to track linked encounters for the same person. The key to the unique identifier will remain at CHIA. All planned linking will be done using aggregate or provider-level data, which will not require any greater level of granularity at the patient level than if we did not do any linking.
Dates will be deidentified, so that they can be used relationally to determine the eligible popluation for the numerator or denominator of the measures, but will not have calendar dates available.

In terms of data security, we keep research datasets secure and protected from improper use and disclosure by using the following approaches:

Physical Security – Keeping data in locked file cabinets, locked offices, locked suites, and physically securing computers and servers.

Electronic Security – Following UCSF minimum security standards for electronic information resources, which includes (but is not limited to): not storing identifiers on portable devices like laptops or flash drives if they are unencrypted, encrypting portable devices, and storing data in password-protected files and on secure networks.

Data handling -- all electronic data will be stored on encrypted computers with password protection. All data will be transferred securely using standard protocols.

Training: All study personnel have received or will receive extensive training and certification in human subjects protection, including specific training in the protection of subject confidentiality. See additional information here: http://irb.ucsf.edu/citi-human-subjectstraining

UCSF staff and faculty, particularly in the Philip R. Lee Institute for Health Policy Studies where Dr. Bardach is appointed and where the grant is being administered, have extensive experience securely recieving, storing, analyzing, and disposing of administrative databases, including CA Medicaid data.

6. Once the linkage is made, what non-MA APCD data elements will appear in the new linked file?

NPPES: Provider variables from the NPPES (gender, degrees, healthcare provider taxomnoy code).
AHA: Hospital variables such as number of psychiatric beds, ownership, insurer types, whether a teaching hospital or not, rural or not, critical access or not, use of an electronic health record or not, hospital volume, and potentially a few other variables related to utilization, payment, and resources available.
ACS: Variables at the zip code level such as percent of race and ethnic categories, average annual income, and other variables of interest regarding SES (job status, education, use of technology).

XI. PUBLICATION / DISSEMINATION / RE-RELEASE

1. Describe your plans to publish or otherwise disclose CHIA Data, or any data derived or extracted from such CHIA Data, in any paper, report, website, statistical tabulation, seminar, conference, or other setting. All publication of CHIA Data must comply with CHIA's cell size suppression policy, as set forth in the Data Use Agreement. Please explain how you will ensure that any publications will not display a cell less than 11, and no percentages or other mathematical formulas will be used if they result in the display of a cell less than 11.

We plan to present and publish analyses from these data for the purpose of dissemination of the scientific research questions we are asking in the proposal (see summary above). However, we will only publish aggregate data on these measures, without disclosing any individual provider or hospital names. No publications will display a cell less than 11, nor any percentages or other mathematical formulas will be used if they result in the display of a cell less than 11. We have experience using other administrative databases (e.g., HCUP databases) which have the same restrictions. To ensure we comply with this, the PI and co-investigators will proof the results and any tables for publication to check for cell sizes of <11.

2. Do you anticipate that the results of your analysis will be published and/or publically available to any interested party? Please describe how an interested party will obtain your analysis and, if applicable, the amount of the fee, that the third party must pay.

We will make the results of our analysis publicly available through publication in a scientific, peer-reviewed journal. Jourbals have variable fee structures, but through the NIH open access policies, a near final draft of the paper will be made publicly available through NCBI/Pubmed.

3. Will you use CHIA Data for consulting purposes?
   - ☐ Yes
   - ☒ No

4. Will you be selling standard report products using CHIA Data?
   - ☐ Yes
   - ☒ No

5. Will you be selling a software product using CHIA Data?
   - ☐ Yes
   - ☒ No

6. Will you be reselling CHIA Data in any format?
   - ☐ Yes
   - ☒ No

If yes, in what format will you be reselling CHIA Data (e.g., as a standalone product, incorporated with a software product, with a subscription, etc.)?

7. If you have answered "yes" to questions 4, 5 or 6, please describe the types of products, services or studies.

8. If you have answered "yes" to questions 4, 5, or 6, what is the fee you will charge for such products, services or studies?

X. APPLICANT QUALIFICATIONS

1. Describe your qualifications (and the qualifications of your co-investigators) to perform the research described.

Naomi S. Bardach, MD, MA
Dr. Bardach is an Associate Professor in the Department of Pediatrics. She is member of the Core Faculty at the Philip R. Lee Institute for Health Policy Studies the University of California, San Francisco (UCSF).

Dr. Bardach's areas of research expertise include quality measurement and quality improvement. She has extensive experience analyzing large datasets to measure quality, using complex quantitative design. Her previous studies have included analysis of claims data to apply quality measures, identify outlier hospitals, and identify of patient and provider characteristics associated with performance, which are all applicable to the work proposed in the grant. She has a track-record of successfully leading national collaborations focused on topics in quality measurement and assessment of quality measures. Dr. Bardach collaborated with the previous Pediatric Quality Measures Program (PQMP) program during Phase 1 of the pediatric measure development with the Center of Excellence on Quality of Care Measures for Children with Complex Needs (COE4CCN). This collaboration led to published work on hospitalizations for pediatric mental health disorders. In addition, Dr. Bardach led the submission of three of the COE4CCN measures for National Quality Forum (NQF) endorsement, preparing all materials, presenting the measures at the NQF measure review committee, and responding to questions for review committee members and the public. She has experience with the use of All-Payer Claims Databases (APCDs) from her work as co-I on subcontract with the Agency for Healthcare Research and Quality (AHRQ) on "Inventory and Prioritization of Measures to Support the Growing Effort in Transparency Using All-payer Claims Databases." The contract includes work assessing feasibility of quality measurement using APCDs from multiple states.
Dr. Bardach will serve as the lead for the IMPLEMENT Measure Feasibility Team and serve on the IMPLEMENT Steering Committee. Dr. Bardach will oversee all aspects of Aim 1, and the data acquisition and analyses for PQMP measures in Aim 2. Dr. Bardach and the Project Analyst will coordinate the analyses of data from the California and Oregon Medicaid Programs. She will lead the analyses of the Massachusetts All-Payer Claims Database and she will also work with Dr. Valerie Harder in the analyses of the Vermont All-Payer Claims Database. Dr. Bardach will also work with the Advisory Core for Measures. As appropriate, Dr. Bardach will work with Dr. Cabana to prepare and submit measures to National Quality Forum/HHS-Consensus Based Entity to leverage field testing results.

Michael D. Cabana, MD, MPH
Dr. Cabana is a Professor of Pediatrics, Epidemiology and Biostatistics. He is member of the Core Faculty at the Philip R. Lee Institute for Health Policy Studies the University of California, San Francisco (UCSF) and Chief of the UCSF Division of General Pediatrics.
Dr. Cabana will serve as the Principal Investigator for the UCSF/Vermont collaboration entitled, "IMPLEmenting MEasures NeTwork for Child Health (IMPLEMENT)". He will be responsible for the overall administration of the

IMPLEMENT Collaboration, as well as the conduct and management of the proposed projects. He will work closely and meet routinely with the IMPLEMENT Steering Committee, which includes the lead for the Usability/QI Team, Dr. Judy Shaw, the lead for the Measure Feasibility Team, Dr. Naomi Bardach, the Sickle Cell Disease Clinical Leads, Dr. Marsha Treadwell and Dr. Anne Marsh and the Biostatistics Lead Dr. Chuck McCulloch. Dr. Cabana will also serve as the Clinical Lead for Asthma Care. He will coordinate the efforts of the IMPLEMENT Advisory Cores for Asthma, Sickle Cell Disease, Quality Improvement and Measures. As appropriate, Dr. Cabana will work with Dr. Bardach to prepare and submit measures to National Quality Forum/HHS-Consensus Based Entity to leverage field testing results.

Charles McCulloch, PhD
Dr. McCulloch is a Professor and Head of the Division of Biostatistics and Vice Chair of the Department of Epidemiology and Biostatistics at the University of California, San Francisco (UCSF).

Dr. McCulloch has broad experience in both the development of statistical methodology and the novel application of advanced statistical methods. He has long-standing collaborations and numerous publications that involve implementation and dissemination projects on clinical topics. He has served as the senior statistician for cluster randomized trials and has been involved in a number of other research projects with system level interventions. Another focus has been on methods for correlated data including longitudinal data models for normally or non-normally distributed outcomes, latent variable and latent class models. He has co-authored book length treatments of those topics (Variance Components, Wiley, 1992 with SR Searle and G Casella; Generalized, Linear, and Mixed Models, Wiley, now in its second edition in 2008 with JM Neuhaus and SR Searle). He is co-author of Regression Methods in Biostatistics (Springer, now in its 2nd edition).

Dr. McCulloch will serve as the Biostatistics Lead and serve on the IMPLEMENT Steering Committee. Dr. McCulloch will provide biostatistical support to ensure appropriate research design, statistical analysis and interpretation for the Measure Feasibility Team and the Usability/QI Team. He will work closely with the QI and Feasibility Teams to support data analysis and interpretation.

Robert Thrombley, B.S.
Mr. Thrombley has extensive experience working with large and complex databases, including work conducted within the Institute for Health Policy Studies at UCSF, where this project will be conducted. He has over 10 years of experience with a variety of programming languages including Python, R, SQL and SAS. He has a strong background in information system design across multiple, diverse industries (health care, education, earth sciences, and telecommunications). He has completed the CITI Human Subjects Protection Training and the HIPAA training at UCSF.

Mr. Thombley will serve as the primary data management and measure assessment team member. He will maintain the data securely using secure servers at UCSF per our standard information security protocols. He will perform data cleaning, assessment of eligibility, determination of performance, attribution of members to practices and providers, and merging of files with outside data as noted above.

2. **Resumes/CVs**: Please include with your application package on IRBNet résumés or curricula vitae of the Applicant/principal investigator, and co-investigators. (These attachments will not be posted on the internet.)

XI. USE OF AGENTS AND/OR CONTRACTORS

Please note: by signing this Application, the Organization assumes all responsibility for the use, security and maintenance of the CHIA Data by its agents, including but not limited to contractors.

Provide the following information for all agents and contractors who will have access to the CHIA data. *Add agents or contractors as needed.*

Company Name:	N/A
Contact Person:	
Title:	
Address, City/Town, Zip Code	
Telephone Number:	
E-mail Address:	
Organization Website:	

1. Will the agent or contractor have access to or store the CHIA Data at a location other than the Applicant's location, off-site server and/or database?

☐ Yes, a separate Data Management Plan **must** be completed by each agent or contractor

☒ No

2. Describe the tasks and products assigned to this agent for this project; their qualifications for completing the tasks; and the Organization's oversight of the agent, including how the Organization will ensure the security of the CHIA Data to which the agent or contractor has access.

N/A

Company Name:	N/A
Contact Person:	
Title:	
Address, City/Town, Zip Code	
Telephone Number:	
E-mail Address:	
Organization Website:	

1. Will the agent or contractor have access to or store the CHIA Data at a location other than the Applicant's location, off-site server and/or database?

☐ Yes, a separate Data Management Plan **must** be completed by each agent or contractor

☒ No

2. Describe the tasks and products assigned to this agent for this project; their qualifications for completing the tasks; and the Organization's oversight of the agent, including how the Organization will ensure the security of the CHIA Data to which the agent or contractor has access.

N/A

XII. FEE INFORMATION

Please consult the fee schedules for MA APCD Data and select from the following options:
☒ Researcher
☐ Others (Single Use)
☐ Others (Multiple Use)

Are you requesting a fee waiver?
☐ Yes
☒ No

If yes, please refer to the Application Fee Remittance Form and submit a letter stating the basis for your request (if required). Please refer to the fee schedule for qualifications for receiving a fee waiver. If you are requesting a waiver based on the financial hardship provision, please provide documentation of your financial situation. Please note that non-profit status alone isn't sufficient to qualify for a fee waiver.

XIII. ATTESTATION

By submitting this Application, the Data Applicant attests that it is aware of its data use, privacy and security obligations imposed by state and federal law *and* is compliant with such use, privacy and security standards. The Data Applicant further agrees and understands that it is solely responsible for any breaches or unauthorized access, disclosure or use of any CHIA Data provided in connection with an approved Application, including, but not limited to, any breach or unauthorized access, disclosure or use by its agents.

Applicants requesting data from CHIA will be provided with data following the execution of a Data Use Agreement that requires the Data Applicant to adhere to processes and procedures aimed at preventing unauthorized access, disclosure or use of data.

By my signature below, I attest to: (1) the accuracy of the information provided herein; (2) that the requested data is the minimum necessary to accomplish the purposes described herein; (3) the Data Applicant will meet the data privacy and security requirements describe in this Application and supporting documents, and will ensure that any third party with access to the data meets the data use, privacy and security requirements; and (4) my authority to bind the organization seeking CHIA Data for the purposes described herein.

Signature: (Authorized Agent)	
Printed Name :	Mora Fisher Mattingly
Title:	Senior Industry Contracts Officer
Signature: (Applicant/Primary Investigator)	
Name:	Naomi S. Bardach
Title:	Associate Professor of Pediatrics and Health Policy
Original Data Request Submission Date:	12/06/2016

Dates Data Request Revised:	

Attachments. Please indicate below which documents have been attached to the Application and uploaded to IRBNet:

☒ 1. IRB approval letter and protocol (if applicable)
☒ 2. 1-2 page Research Methodology
☒ 3. Resumes of Applicant and co-investigators
☒ 4. Data Management Plan (including one for each agent or contractor that will have access to or store the CHIA Data at a location other than the Applicant's location, off-site server and/or database)
☒ 5. Fee Remittance Form (including any required documentation if a fee waiver is being requested)