

Non-Government Application for Massachusetts All-Payer Claims Data [Exhibit A]

I. INSTRUCTIONS

This form is required for all Applicants, except Government Agencies as defined in [957 CMR 5.02](#), requesting protected health information. 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 whether the request meets the criteria for data release, pursuant to 957 CMR 5.00. Please complete the Application documents fully and accurately. Prior to receiving CHIA Data, the Organization must execute CHIA's [Data Use Agreement](#). Applicants may wish to review that document prior to submitting this Application.

Before completing this Application, please review the data request information on CHIA's website:

- [Data Availability](#)
- [Fee Schedule](#)
- [Data Request Process](#)

After reviewing the information on the website and this Application, please contact CHIA at apcd.data@state.ma.us if you have additional questions about how to complete this form.

All attachments must be uploaded to IRBNet with your Application. All Application documents can be found on the [CHIA website](#) in Word and in PDF format or on [IRBNet](#) in Word format. If you submit a PDF document, please also include a Word version in order to facilitate edits that may be needed.

Applications will not be reviewed until the Application and all supporting documents are complete and the required application fee is submitted. A [Fee Remittance Form](#) with instructions for submitting the application fee is available on the CHIA website and IRBNet. If you are requesting a fee waiver, a copy of the Fee Remittance Form and any supporting documentation must be uploaded to IRBNet.

II. FEE INFORMATION

1. Consult the most current [Fee Schedule](#) for All-Payer Claims Database data.
2. After reviewing the Fee Schedule, if you have any questions about the application or data fees, contact apcd.data@state.ma.us.
3. If you believe that you qualify for a fee waiver, complete and submit the [Fee Remittance Form](#) and attach it and all required supporting documentation with your application. Refer to the [Fee Schedule](#) (effective Feb 1, 2017) for fee waiver criteria.
4. Applications will not be reviewed until the application fee is received.
5. Data for approved Applications will not be released until the payment for the Data is received.

III. ORGANIZATION & INVESTIGATOR INFORMATION

Project Title:	Clinical and Economic Performance of Health Systems
IRBNet Number:	
Organization Requesting Data (Recipient):	National Bureau of Economic Research
Organization Website:	www.nber.org
Authorized Signatory for Organization:	Alterra Milone
Title:	Director, Research and Grants Management, Corporate Secretary
E-Mail Address:	alterra@nber.org
Address, City/Town, State, Zip Code:	National Bureau of Economic Research 1050 Massachusetts Avenue Cambridge, MA 02138
Data Custodian: (individual responsible for organizing, storing, and archiving Data)	Mohan Ramanujan
Title:	Sr. Unix Systems Administrator
E-Mail Address:	mohan@nber.org
Telephone Number:	617-588-0367
Address, City/Town, State, Zip Code:	National Bureau of Economic Research 1050 Massachusetts Avenue Cambridge, MA 02138
Primary Investigator (Applicant): (individual responsible for the research team using the Data)	David Cutler
Title:	Otto Eckstein Professor of Applied Economics
E-Mail Address:	dcutler@harvard.edu
Telephone Number:	617-496-5216
Names of Co-Investigators:	Nancy Beaulieu
E-Mail Addresses of Co-Investigators:	beaulieu@hcp.med.harvard.edu

IV. PROJECT INFORMATION

1. What will be the use of the CHIA Data requested? [Check all that apply]

- | | | |
|---|--|--|
| <input type="checkbox"/> Epidemiological | <input type="checkbox"/> Health planning/resource allocation | <input checked="" type="checkbox"/> Cost trends |
| <input checked="" type="checkbox"/> Longitudinal Research | <input checked="" type="checkbox"/> Quality of care assessment | <input type="checkbox"/> Rate setting |
| <input type="checkbox"/> Reference tool | <input checked="" type="checkbox"/> Research studies | <input type="checkbox"/> Severity index tool |
| <input type="checkbox"/> Surveillance | <input type="checkbox"/> Student research | <input type="checkbox"/> Utilization review of resources |
| <input type="checkbox"/> Inclusion in a product | <input type="checkbox"/> Other (describe in box below) | |

2. Provide an abstract or brief summary of the specific purpose and objectives of your Project. This description should include the research questions and/or hypotheses the project will attempt to address, or describe the intended product or report that will be derived from the requested data and how this product will be used. Include a brief summary of the pertinent literature with citations, if applicable.

Delivery of medical care requires a diverse set of providers: physicians, allied health personnel (such as nurse practitioners), hospitals (community and tertiary), outpatient facilities, and post-acute providers. The relationships between these providers vary substantially, ranging from small organizations (e.g., solo practitioners) to large integrated delivery systems. This organizational structure may have far-reaching effects. Larger organizations may be better able to share information and coordinate care. On the other hand, diseconomies of scale and scope may impede efficiencies in large organizations, and they may use market power to charge higher prices. Given the rapid transformation of the health care system, with organizations forming new relationships as they react to new incentives for accountability and cost reduction, understanding how system traits affect outcomes is crucially important.

The goal of this research is to systematically examine the delivery system traits that predict greater delivery of PCOR-based evidence, and better clinical and economic outcomes. We believe that gathering and analyzing the right data will allow us to determine which, if any, organizational traits are associated with superior performance. We will examine delivery system performance in several distinct areas: by looking across a variety of delivery systems; by considering specific populations where evidence-based care involves tailoring care to the needs of the population, such as oncology, children, and post-acute care; and by considering the outcome of corporate integration in a number of settings. Three cross-cutting aims span all of our research:

- Specific Aim 1: Characterize the current delivery system and track changes in delivery system structure across space and over time.
- Specific Aim 2: Gather data on use of PCOR-based evidence and related clinical and economic outcomes and merge those data with delivery system information.
- Specific Aim 3: Use these various data sources to examine how differences in delivery system structure are related to the use of PCOR-based evidence and related clinical and economic outcomes.

This study will use APCD data to calculate clinical and economic outcomes. This information along with characteristics of health systems will increase understanding of what systems' attributes impact clinical quality, and costs within Massachusetts and across other states.

CHIA APCD data will be used to calculate clinical and economic outcomes among Massachusetts healthcare providers. We will merge our enhanced systems database (EDB) of healthcare providers, systems, and their characteristics to the CHIA APCD data in order to investigate the delivery system traits that predict greater delivery of PCOR-based evidence and better clinical and economic outcomes.

3. Has an Institutional Review Board (IRB) reviewed your Project?

- Yes [If 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.

4. **Research Methodology:** Applicants must provide either the IRB protocol or 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.

V. PUBLIC INTEREST

1. Briefly explain why completing your Project is in the public interest. Use quantitative indicators of public health importance where possible, for example, numbers of deaths or incident cases; age-adjusted, age-specific, or crude rates;

or years of potential life lost. *Uses that serve the public interest under CHIA regulations include, but are not limited to: health cost and utilization analysis to formulate public policy; studies that promote improvement in population health, health care quality or access; and health planning tied to evaluation or improvement of Massachusetts state government initiatives.*

This study will identify health system characteristics associated with higher quality and lower costs in Massachusetts and across the nation. This knowledge will be disseminated among interested stakeholders including policy makers to encourage the adoption of activities and programs likely to facilitate progress toward the achievement of the triple aim for individuals in Massachusetts and other states across the nation. We will identify and analyze delivery system characteristics associated with lower costs, higher quality and better patient experience. Public and private programs may then be designed to encourage the adoption of delivery system structures and best practices to other health care organizations in the state and beyond.

VI. DATA REQUESTED

The Massachusetts All-Payer Claims Database is comprised of medical, pharmacy, and dental claims and information from the member eligibility, provider, and product files that are collected from health insurance payers licensed to operate in the Commonwealth of Massachusetts. This information encompasses public and private payers as well as data from insured and self-insured plans. APCD data are refreshed and updated annually and made available to approved data users in Release Versions that contain five calendar years of data and three months of run-out. Data requests will be fulfilled using the most current Release Version. For more information about the most current APCD Release Version, including available years of data and a full list of elements in the release please refer to release layouts, data dictionaries and similar documentation included on [CHIA's website](#).

Data requests are typically fulfilled on a one time basis, however; certain Projects may require future years of data that will become available in a subsequent release. Applicants who anticipate a need for future years of data may request to be considered for a subscription. Approved subscriptions will receive, upon request, the same data files and data elements included in the initial Release annually or as available. Please note that approved subscription request will be subject to the Data Use Agreement, will require payment of fees for additional Data, and subject to the limitation that the Data can be used only in support of the approved Project.

1. List years of data requested (only list years available in the current Release Version): 2012-2016; 2017-2018 when available

2. Please indicate below whether this is a one-time request, or if the described Project will require a subscription.

One-Time Request **OR** Subscription

3. Specify below the data files requested for this Project, and provide your justification for requesting each file.

Medical Claims

Describe how your research objectives require Medical Claims data:

Longitudinal medical claims data are needed to compute measures of spending, utilization, and quality at the patient level in order to generate a broad range of provider-specific cost and quality metrics. We will relate these metrics to provider characteristics to examine which characteristics are associated with high quality and/or low cost.

Pharmacy Claims

Describe how your research objectives require Pharmacy Claims data:

Since pharmacy claims constitute a significant spending component and are often used to construct quality measures, these claims are needed to measure spending, utilization, and quality. We will use pharmacy claims to generate some key measures of clinical quality, such as the appropriate use of antibiotics.

Dental Claims

Describe how your research objectives require Dental Claims data:

Dental care is an often understudied dimension of overall health. Many health systems now include dental practices. Dental claims data will allow us to assess the extent that dental care is integrated to the rest of the health care system and the relationship between such integration and quality outcomes.

Member Eligibility

Describe how your research objectives require Member Eligibility data:

Member eligibility data will be crucial in identifying participation in the Medicaid program, tracking Medicaid patients as they move in and out of Medicaid, and in identifying comparison groups. Demographic and residential location data will be important in controlling for patient characteristics and access to providers. This information will be used to study how providers and health systems vary by the patient population they serve. By accurately identifying patients on Medicaid and comparison groups, we will be able to compare outcomes for Medicaid and commercially insured patients within the state, and to examine how delivery system differences affect one of the populations most at risk of losing access to care.

Provider

Describe how your research objectives require Provider data:

We will use data on providers to study how differences in provider characteristics and delivery system impact the cost and quality of care. We have identified health systems across the nation using a large number of proprietary and government datasets.

These health systems will be matched on provider identification codes (e.g., NPIs and TINs) to providers in the Massachusetts APCD.

Product

Describe how your research objectives require Product data:

Spending (decomposed into utilization and prices) is a central outcome measure in this project. The computation of spending measures requires both insurer and patient payment information so that we can control for any differences in benefit design on utilization. We do not need to know the name of the insurer but would like access to a de-identified insurer codes and plan codes. These codes will help us to infer the networks that patients can access through their plan and to control for variation in benefit generosity in our regression analyses.

VII. DATA ENHANCEMENTS REQUESTED

State and federal privacy laws limit the release and use of Data to the minimum amount of data needed to accomplish a specific Project objective.

All-Payer Claims Database data is released in Limited Data Sets (LDS). All applicants receive the "Core" LDS, but may also request the data enhancements listed below for inclusion in their analyses. Requests for enhancements will be reviewed by CHIA to determine whether each represents the minimum data necessary to complete the specific Project objective.

For a full list of elements in the release (i.e., the core elements and additional elements), please refer to [release layouts](#), [data dictionaries](#) and similar documentation included on CHIA's website.

1. Specify below which enhancements you are requesting in addition to the “Core” LDS, provide your justification for requesting each enhancement.

Geographic Subdivisions

The geographic subdivisions listed below are available for Massachusetts residents and providers only. Select one of the following options.

<input type="checkbox"/> 3-Digit Zip Code (standard)	<input checked="" type="checkbox"/> 5-Digit Zip Code***
<p>***If requested, provide justification for requesting 5-Digit Zip Code. Refer to specifics in your methodology:</p> <p>We need 5-digit zip code to impute socio-demographic characteristics (using census data), to assign beneficiaries to health care markets, and to measure distances traveled to access care.</p>	

Date Resolution

Select one option from the following options.

<input type="checkbox"/> Year (YYYY) (Standard)	<input type="checkbox"/> Month (YYYYMM) ***	<input checked="" type="checkbox"/> Day (YYYYMMDD) *** [for selected data elements only]
<p>*** If requested, provide justification for requesting Month or Day. Refer to specifics in your methodology:</p> <p>Month and year of birth will be necessary to compute certain quality measures (e.g. pediatric care) and to compute age for risk adjustment. Month and day on claims will be required to link claims for a single episode (e.g. patients seen in a doctors office and an emergency department on the same day), to analyze the sequence of providers seen by patients, and to analyze quality measures in health care service delivery such as follow-up after discharge and the timeliness and appropriate setting of care.</p>		

National Provider Identifier (NPI)

Select one of the following options.

<input type="checkbox"/> Encrypted National Provider Identifier(s) (standard)	<input checked="" type="checkbox"/> Decrypted National Provider Identifier(s)***
<p>*** If requested, provide justification for requesting decrypted National Provider Identifier(s). Refer to specifics in your methodology:</p> <p>We need decrypted National Provider Identifiers in order to link to external data sources with information on provider characteristics (e.g. physician specialty, license (e.g. nurse practitioner, MD, MSW), years in practice).</p>	

VIII. MEDICAID (MASSHEALTH) DATA

1. Please indicate whether you are seeking Medicaid Data:

- Yes
- No

2. 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 MassHealth Data, please describe, in the space below, why your use of the Data meets this requirement. *Your description should focus on how the results of your project could be used by the Executive Office of Health and Human Services in connection with the administering the MassHealth program.* Requests for MassHealth 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 MassHealth

program. CHIA cannot release MassHealth Data without approval from MassHealth. This may introduce significant delays in the receipt of MassHealth Data.

The goal of this project is to examine how characteristics of health care organizations and systems relates to health care quality and cost. Medicaid data are essential for comparing outcomes for Medicaid and commercially insured patients within the state and for examining how delivery system differences affect one of the populations most at risk of losing access to care. Specifically, we are trying to characterize which types of health care providers and systems provide high quality care to Medicaid beneficiaries. Findings from these analyses will inform plan design and regulation in the MassHealth program, such as regulations related to network design.

IX. DATA LINKAGE

Data linkage involves combining CHIA Data with other data 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 data?

Yes

No linkage or merger with any other data will occur

2. If yes, please indicate below the types of data to which CHIA Data will 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 (e.g., American Hospital Association data)

Aggregate Data (e.g., Census data)

Other (please describe):

3. If yes, describe the dataset(s) to which the CHIA Data will be linked, indicate which CHIA Data elements will be linked and the purpose for each linkage.

We will conduct this study in full concordance with the Health Insurance Portability and Accountability Act (HIPAA). All of the basic processing of the raw data and all data linkages required for individual research projects are conducted centrally through NBER's full-time data management personnel, subject to our highest data security standards.

We will be merging and linking APCD data with other dataset(s) as described below for the purpose of analyses. However, we will not be stacking CHIA APCD data with databases from other states, and we will not be adding CHIA APCD data into our enhanced system database (EDB) itself. The MA APCD will remain a discrete database to which we can make linkages for the purposes of our analyses, but which will be destroyed at the conclusion of the project.

Claim line-level data, including zip code, will not be linked to individual patient-level data.

Provider level linkage: We will merge information at the provider level from the enhanced system database (EDB) we have built as part of this project. The EDB integrates data from a number of sources, including: Medicare Fee-for-Service Claims, Medicare Provider Enrollment, Chain, and Ownership System (PECOS), Medicare Data on Provider Practice & Specialty (MD-PPAS), Medicare Provider of Services (POS), Hospital Compare, Physician Compare, National Plan & Provider Enumeration System (NPPES) NPI Registry, commercial claims, IRS 990, IRS Business Master, SEC 10K filings, SK&A, Irving Levin Health Care Acquisition Reports, S&P Capital IQ M&A Transactions, and the AHA Annual Survey. The linking variables we will use include provider IDs (e.g. NPIs, hospital CCNs); tax IDs (TINs and/or EINs); and zip codes.

We will also merge information at the provider level from CMS Open Payment, the National Survey of Healthcare Organizations and Systems (NSHOS). Linking variables may include hospital/system name, CMS provider number, names, tax ID, or other available identifiers.

Physician-level linkage: We will also link to physician-level data from NPPES and CMS Open Payment to obtain characteristics of individual providers. We will link to CMS Provider Utilization and Payment Data to obtain information on volume of services provided. We will link to CMS Hospital Compare data and CMS Physician Compare data to obtain information on provider quality. Physicians will be matched on NPI (or other available identifiers if NPI is not available) to physicians in the Massachusetts APCD.

Zip code level linkage: We will link patient zip codes to U.S. Census data to obtain aggregated socio-economic characteristics of individuals living in the zip code and to classify regions as urban/suburban/rural.

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 first attempt to link providers and facilities using government IDs (deterministic). In cases where this linkage fails, we will use fuzzy matching techniques to match on names and addresses (probabilistic). We will map zip codes directly to Census data (deterministic).

5. If yes, attach or provide below a complete listing of the variables from all sources to be included in the final linked analytic file.

Providers: specialty, years in practice, medical school, affiliation with academic medical center, practice address, practice characteristics (size, specialty composition, owner, hospital outpatient dept. status, # of practice sites, health system). Hospitals and other facilities: beds by type, services available, satellite facilities, health system membership, affiliation with health systems or other facilities, provision of specific services by staff or contract, staffing, finances, tax status, address of main campus and all satellite facilities, GME and DSH payments.

Zip code: % residents with incomes below federal poverty level, racial composition, urban/suburban/rural classification, education attainment, insurance coverage, age distribution, employment status.

Data Set	Level of Observation	Linkage variable
Enhanced System Database	Provider (hospital or physician)	National Provider ID or Tax ID
CMS Open Payment data	Individual physician	NPI (or other available identifiers if NPI is not available)
CMS Open Payment data	Individual hospital (teaching hospitals only)	CMS provider number, names, tax ID, or other available identifiers
Census data	Zipcode	Zipcode
CMS Provider Utilization and Payment Data	Individual physician	NPI
Publicly available information (from news reports, etc.) on local providers	Hospital, physician groups, or health systems	Hospital name, CMS provider number, names, tax ID, or other available identifiers

National Survey of Healthcare Organizations and Systems (NSHOS)	Hospital, physician groups, or health systems	Hospital/system name, CMS provider number, names, tax ID, or other available identifiers
---	---	--

Data dictionaries for these data sets, including Enhanced System Database (EDB) source files, EDB output files, and the other data sources listed above, can be found here: https://www.dropbox.com/sh/qb9nvhmy8c6lx6v/AAAYZzek_hhklii1XdUOgvZua?dl=0

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

Identification of individual patients may occur at two junctures. First, in the context of the analyses, in very sparsely populated zip codes, it may be possible to use the linked data in combination with additional data available on the web to identify individual patients. Investigators have been trained and certified in human subjects research and ethics and understand this is a violation of patient privacy. The second juncture is at the presentation and publication of research results. All research results will be aggregated to a level that would make it impossible to identify individual patients.

X. PUBLICATION / DISSEMINATION / RE-RELEASE

1. Do you anticipate that the results of your analysis will be published or made publically available? If so, how do you intend to disseminate the results of the study (e.g.; publication in professional journal, poster presentation, newsletter, web page, seminar, conference, statistical tabulation)? Any and 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 disclose a cell less than 11**, and percentages or other mathematical formulas that result in the display of a cell less than 11.

We anticipate publishing the results of our analyses in peer-reviewed journals and working paper series. We will submit to clinical, health policy, and economics outlets such as the New England Journal of Medicine, Journal of the American Medical Association, Health Affairs, Annals of Internal Medicine, Healthcare, JAMA Internal Medicine, and the Journal of Health Economics. We will also disseminate our findings with policy leaders, payers, provider networks, and academic leaders within Massachusetts and at professional meetings such as AcademyHealth and the Society of General Internal Medicine. All tables and figures will be reviewed by investigators and research compliance officers in our department to ensure that no results with a cell size less than 11 are divulged.

2. Describe your plans to use or otherwise disclose CHIA Data, or any Data derived or extracted from such Data, in any paper, report, website, statistical tabulation, seminar, or other setting that is not disseminated to the public.

We have no plans to publish or otherwise disclose CHIA data directly. We will disseminate our findings with policy leaders, payers, provider networks, and academic leaders within Massachusetts and at professional meetings such as American Economics Association Meetings and AcademyHealth.

3. What will be the lowest geographical level of analysis of data you expect to present for publication or presentation (e.g., state level, city/town level, zip code level, etc.)? Will maps be presented? If so, what methods will be used to ensure that individuals cannot be identified?

The lowest geographical level of data analysis we expect to present is the county. There are sufficiently large populations in Massachusetts counties that it will not be possible to identify individuals from the results.

4. Will you be using CHIA Data for consulting purposes?

- Yes
 No

5. Will you be selling standard report products using CHIA Data?

- Yes
 No

6. Will you be selling a software product using CHIA Data?

- Yes
 No

7. Will you be using CHIA Data as in input to develop a product (i.e., severity index tool, risk adjustment tool, reference tool, etc.)

- Yes
 No

8. Will you be reselling CHIA Data in any format not noted above?

- Yes
 No

If yes, in what format will you be reselling CHIA Data?

N/A

9. If you have answered "yes" to questions 5, 6, 7 or 8, please describe the types of products, software, services, or tools.

N/A

10. If you have answered "yes" to questions 5, 6, 7 or 8, what is the fee you will charge for such products, software, services or tools?

N/A

XII. APPLICANT QUALIFICATIONS

1. Describe your previous experience using claims data. This question should be answered by the primary investigator and any co-investigators who will be using the Data.

Investigators Cutler, Beaulieu, and Zhou each have years of experience working with claims data.

Dr. Beaulieu has worked with Massachusetts APCD data to study high-cost patients and with Medicare and Commercial claims data to investigate the organization and performance of health systems.

Ms. Zhou has worked extensively with Medicare claims and other administrative data (e.g. inpatient discharge data from various states, Medicaid claims data from South Carolina) to study the effect of new delivery models and the relationship between provider characteristics and patients' choice of providers.

Dr. Cutler has had nearly 30 years of experience working with various claims data, including Medicare and Medicaid claims, commercial claims, and APCD data from Massachusetts and other states.

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

XIII. USE OF AGENTS AND/OR CONTRACTORS

By signing this Application, the Agency assumes all responsibility for the use, security and maintenance of the CHIA Data by its agents, including but not limited to contractors. The Agency must have a written agreement with the agent of contractor limiting the use of CHIA Data to the use approved under this Application as well as the privacy and security standards set forth in the Data Use Agreement. CHIA Data may not be shared with any third party without prior written consent from CHIA, or an amendment to this Application. CHIA may audit any entity with access to CHIA Data.

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

AGENT/CONTRACTOR #1 INFORMATION	
Company Name:	Harvard Medical School
Company Website	https://www.hcp.med.harvard.edu/
Contact Person:	Nancy Beaulieu
Title:	Research Associate
E-mail Address:	beaulieu@hcp.med.harvard.edu
Address, City/Town, State, Zip Code:	180 Longwood Ave., Boston, MA 02115
Telephone Number:	617-432-0628
Term of Contract:	Full project duration

1. Describe the tasks and products assigned to the agent or contractor for this Project and their qualifications for completing the tasks.

This project is a collaboration between researchers at NBER and Harvard. Nancy Beaulieu is a co-investigator on this project from the Health Care Policy department at Harvard Medical School. Dr. Beaulieu has worked with Massachusetts APCD data to study high-cost patients and with Medicare and Commercial claims data to investigate the organization and performance of health systems.

2. Describe the Organization's oversight and monitoring of the activities and actions of the agent or contractor for this Project, including how the Organization will ensure the security of the CHIA Data to which the agent or contractor has access.

The CHIA data will be stored at the National Bureau of Economic Research, on NBER's secure servers. The data will be accessed via an authenticated VPN connection. All users must sign the NBER security pledge:

PROCEDURES FOR ASSURING INDIVIDUAL CONFIDENTIALITY IN NBER HEALTH CARE DATA

1. All researchers and staff who receive authorization to use specific data files sign an agreement certifying that they will comply with the electronic and physical security procedures associated with the use of the data; that they will fully protect the confidentiality of individuals in the data; and that they understand the disciplinary procedures associated with non-compliance.
2. Authorization to use specific data files is restricted to the principal investigator and a group of other key investigators and staff who are working under the direct supervision of the principal investigator.
3. All physical tapes/CDROMs/portable media containing the raw data files are secured in locked cabinets and locked offices. Only systems administrators and data managers would have access to them, who will use these for the purposes of loading the data on the servers and will be carefully monitored.
4. Data extracts developed for specific analyses are maintained in restricted access storage drives. Access to these storage devices are restricted to servers dedicated for the analyses of medical data. Only those individuals who have signed the appropriate confidentiality agreements, and received official authorization, can access these servers and files. Electronic security measures (using password and group authorization) prevent any access to unauthorized users, and restrict authorized users to the data of their particular projects.
5. Exclusive disk space in the restricted storage space will be assigned to each study and all the analysis would be conducted only in these allocated spaces. Transfer of data to any other place is prohibited. This includes prohibition of transfer of data from location provided for one DUA to a location provided for another DUA.
6. Only data extracts being used for ongoing analyses are stored in the restricted access disks on the NBER computer network. Data extracts developed for completed studies are removed from the computer network and a backup copy either on tapes or on offline hard drives are made and secured in locked cabinets or locked offices for future analysis.
7. Any violation of these security standards may result in suspension from NBER affiliation and prohibition from participation in or receiving compensation from any NBER funded project for a period to be determined by the NBER Human Subjects Committee. Serious Human Subjects violations will be penalized by permanent expulsion from any future NBER affiliation. Written notification will also be provided to the appropriate funding agencies and data providers documenting the nature of the violation, who may enforce additional disciplinary procedures that are within their legal authority.
8. Any known disclosure of confidential data will be immediately reported to NBER officers, Data Providing Agency and Project Funding Agency.

CERTIFICATION OF ADHERENCE TO CONFIDENTIALITY STANDARDS FOR PATIENT HEALTH INFORMATION

I, _____ understand that I must adhere to the following agreement in order to be in compliance with the confidentiality standards associated with Personal Health Information:

I will not transfer, sell, or lend access to the data under any circumstances.

Without consent, even within NBER's computing facilities, I will not transfer data or extracts created and stored in folders provided for a DUA to any other location, such as, to a home directory or to folders provided for another DUA.

I will follow the physical and electronic security policies developed by the NBER to prevent access to the data by unauthorized personnel.

I will not seek to identify any of the individuals or establishments in the datasets. If the identity of any person or establishment in the data is discovered inadvertently, I (i) will not make any use of this information, (ii) will safeguard or destroy the information that would identify such

person or establishment, (iii) will not inform anyone of this discovery other than those who need to know in order to protect the data from future identifications, and (iv) will immediately report any known disclosure of confidential data to NBER officers, Data Providing Agency and Project Funding Agency.

Further, I understand that if I fail to comply with these standards I may be suspended from NBER affiliation and prohibited from participating or receiving compensation from any NBER funded project for an extended period to be determined by the NBER Human Subjects Committee. I realize that serious Human Subjects violations will be penalized by permanent expulsion from any future NBER affiliation, and that in all cases the NBER will provide written notification to the data source documenting the nature of the violation, and that the source will be encouraged to impose additional penalties that are within their legal authority.

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

- Yes
- No

4. If yes, a separate Data Management Plan **must** be completed by the agent or contractor.

AGENT/CONTRACTOR #2 INFORMATION	
Company Name:	
Company Website:	
Contact Person:	
Title:	
E-mail Address:	
Address, City/Town, State, Zip Code:	
Telephone Number:	
Term of Contract:	

1. Describe the tasks and products assigned to the agent or contractor for this Project and their qualifications for completing the tasks.

2. Describe the Organization’s oversight and monitoring of the activities and actions of the agent or contractor for this Project, including how the Organization will ensure the security of the CHIA Data to which the agent or contractor has access.

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

- Yes
- No

4. If yes, a separate Data Management Plan **must** be completed by the agent or contractor.

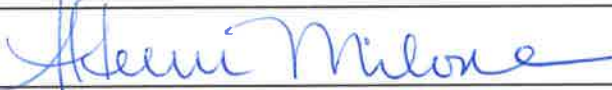
[INSERT A NEW SECTION FOR ADDITIONAL AGENTS/CONTRACTORS AS NEEDED]

IVX. ATTESTATION

By submitting this Application, the Organization attests that it is aware of its data use, privacy and security obligations imposed by state and federal law *and* confirms that it is compliant with such use, privacy and security standards. The Organization further agrees and understands that it is solely responsible for any breaches or unauthorized access, disclosure or use of CHIA Data, including, but not limited to, any breach or unauthorized access, disclosure or use by any third party to which it grants access.

Applicants approved to receive CHIA Data will be provided with Data following the payment of applicable fees and upon the execution of a Data Use Agreement requiring the Organization to adhere to processes and procedures designed to prevent unauthorized access, disclosure or use of data.

By my signature below, I attest: (1) to the accuracy of the information provided herein; (2) that the requested Data is the minimum necessary to accomplish the purposes described herein; (3) that the Organization will meet the data privacy and security requirements described 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) to my authority to bind the Organization.

Signature: (Authorized Signatory for Organization)	
Printed Name:	Alterra Milone
Title:	Director of Research and Grants Management, National Bureau of Economic Research

Attachments

A completed Application must have the following documents attached to the Application or uploaded separately to IRBNet:

- 1. IRB approval letter and protocol (if applicable), or research methodology (if protocol is not attached)
- 2. 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 Organization’s location, off-site server and/or database
- 3. CVs of Investigators (upload to IRBnet)

APPLICATIONS WILL NOT BE REVIEWED UNTIL THEY ARE COMPLETE, INCLUDING ALL ATTACHMENTS.