

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

I. INSTRUCTIONS

This form is required for all Applicants, Agencies, or Organizations, hereinafter referred to as “Organization”, except Government Agencies as defined in [957 CMR 5.02](#), requesting protected health information. All Organizations must also complete the [Data Management Plan](#), and attach it to this Application. The Application and the Data Management Plan must be signed by an authorized signatory. 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](#). Organizations 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.

The Application and all attachments must be uploaded to IRBNet. All Application documents can be found on the [CHIA website](#).

Information submitted as part of the Application may be subject to verification during the review process or during any audit review conducted at CHIA’s discretion.

Applications will not be reviewed until the Application and all supporting documents are complete and the required application fee is received.

A [Fee Remittance Form](#) with instructions for submitting the application fee is available on the CHIA website. If you are requesting a fee waiver, a copy of the [Fee Remittance Form](#) and any supporting documentation must be uploaded to IRBNet. Please be aware that if your research is funded and under that funding you are required to release raw data to the funding source, you may not receive CHIA Data.

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:	Insurer-provider Competition and Integration in Health Care Markets
IRBNet Number:	Click here to enter text.
Organization Requesting Data (Recipient):	Northwestern University
Organization Website:	https://www.northwestern.edu/
Authorized Signatory for Organization:	Click here to enter text.
Title:	Click here to enter text.
E-Mail Address:	Click here to enter text.
Telephone Number:	Click here to enter text.
Address, City/Town, State, Zip Code:	Click here to enter text.
Data Custodian: (individual responsible for organizing, storing, and archiving Data)	William Thompson; Ambreen Chaudhri; Richard Williams
Title:	Senior Computational Research Consultants
E-Mail Address:	wkt@kellogg.northwestern.edu, ambreen.chaudhri@kellogg.northwestern.edu, richard.williams@kellogg.northwestern.edu
Telephone Number:	847-467-1854
Address, City/Town, State, Zip Code:	2211 Campus Drive, Office 5422, Evanston, IL, 60208
Primary Investigator (Applicant): (individual responsible for the research team using the Data)	David Dranove
Title:	Walter McNerney Distinguished Professor of Health Industry Management
E-Mail Address:	d-dranove@kellogg.northwestern.edu
Telephone Number:	847-491-8682
Address, City/Town, State, Zip Code:	2211 Campus Drive, Office 4243, Evanston, IL, 60208
Names of Co-Investigators:	Anran Li, Hyein Cho, Alexander Boche
E-Mail Addresses of Co-Investigators:	anranli@u.northwestern.edu, hyein.cho@kellogg.northwestern.edu alexander.boche@kellogg.northwestern.edu

IV. PROJECT INFORMATION

IMPORTANT NOTE: Organization represents that the statements made below as well as in any study or research protocol or project plan, or other documents submitted to CHIA in support of the Data Application are complete and accurate and represent the total use of the CHIA Data requested. Any and all CHIA Data released to the Organization under an approved application may ONLY be used for the express purposes identified in this section by the Organization, and for no other purposes. Use of CHIA Data for other purposes requires a separate Data Application to CHIA written request to CHIA, with approval being subject to CHIA's regulatory restrictions and approval process. Unauthorized use is a material violation of your institution's Data Use Agreement with CHIA.

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

- Epidemiological Health planning/resource allocation Cost trends
 Longitudinal Research Quality of care assessment Rate setting

- | | | |
|---|--|---|
| <input type="checkbox"/> Reference tool | <input checked="" type="checkbox"/> Research studies | <input type="checkbox"/> Severity index tool (or other derived input) |
| <input type="checkbox"/> Surveillance | <input checked="" 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) | |

Click here to enter text.

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.

Research questions

The research project examines the role of competition in determining health outcomes, costs, and consumer welfare in the U.S. healthcare market. Specifically, the project focus on the effects of three different types of competitive relations: competition between insurers, competition between providers, and relations between providers and insurers. The project aims to answer the following questions: (1) How does insurer competition affect insurer incentives to cover and invest in preventive and curative care given that some patients may switch to other insurers over time? And how does this affect patient health outcomes, care costs, and quality? (2) In markets where patients can choose from competing providers, how should insurers set cost-sharing levels to maximize patient welfare? Specifically, what cost-sharing levels would optimally balance patient risk protection with incentives for price shopping between competing providers? (3) Do cost and quality of healthcare differ for members on provider-sponsored health plans compared to members on non-provider sponsored health plans? What are the potential mechanisms behind the difference?

Products

The objective of this project is to produce results to be disseminated via publication in a peer-reviewed academic journal (e.g., American Economic Review, Journal of Health Economics) and seminar presentations. These results will be informative to policy makers making decisions about health care market regulations.

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:** Applications must include 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.

Please see the project methodology attachment for detailed descriptions.

V. PUBLIC INTEREST

1. Briefly explain why completing this 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.

The results of this project shed light on how competition, consolidation, and consumer incentives affect health care quality, costs, and the balancing thereof. The project will simulate potential market outcomes under alternative policy designs that aim to reduce health care costs, improve population health and consumer welfare. The results could provide useful quantitative benchmark for formulating public policies that aim to improve health care quality, affordability, and access.

VI. DATASETS 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. For more information about APCD Release Versions, 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. Projects that 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 requests are subject to the Data Use Agreement, will require payment of fees for additional Data for Non-Government Entities, and subject to the limitation that the Data can be used only in support of the approved Project.

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

One-Time Request **OR** Subscription

2. Select Release Version and years of data requested (Release Versions and years not listed are not available).

<input type="checkbox"/> Release Version 8.0 <input type="checkbox"/> 2014 <input type="checkbox"/> 2015 <input type="checkbox"/> 2016 <input type="checkbox"/> 2017 <input type="checkbox"/> 2018	<input checked="" type="checkbox"/> Release Version 10.0 <input checked="" type="checkbox"/> 2015 <input checked="" type="checkbox"/> 2016 <input checked="" type="checkbox"/> 2017 <input checked="" type="checkbox"/> 2018 <input checked="" type="checkbox"/> 2019 <input checked="" type="checkbox"/> 2020
--	---

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:

Medical claims data will allow us to observe health care service each patient received (e.g., preventive service, treatment), the associated costs, and claims-based outcome (e.g., hospital readmission). Additionally, medical claims will be used to study how patients select from among competing providers and how providers refer patients.

 Pharmacy Claims**Describe how your research objectives require Pharmacy Claims data:**

Pharmacy claims will allow us to view cost and medical health outcome variables of interest (e.g., use of preventive vs. curative drugs). Further, estimating optimal copays for competing providers requires understanding the level of background financial risk enrollees face and its correlation with use of medical provider services. This financial risk includes out-of-pocket costs for pharmaceuticals.

 Dental Claims**Describe how your research objectives require Dental Claims data:**

Dental claims will allow us to observe dental service each patient has received (e.g., preventive service, treatment), the associated costs, and claim-based outcome. Also, it will allow us to study how patients choose among competing dental providers including how they respond to differential cost-sharing.

 Member Eligibility**Describe how your research objectives require Member Eligibility data:**

Member eligibility table provides information on insurance characteristics (e.g., cost-sharing features and premiums). This will also give us member and employer zip codes which will be used as predictors of patient choice of provider (together with map and census data). Also, it is critical to conduct subgroup analysis based on demographic information.

 Provider**Describe how your research objectives require Provider data:**

Provider data provides the national provider identifier which is necessary to link to external data sources of provider information which include quality ratings, hospital affiliation, provider location, organizational structure of physician groups, and hospital cost data. Complete list of external data sources can be found in Section IX.

 Product**Describe how your research objectives require Product data:**

Product data provides insurer name, plan codes, and plan characteristics which allow us to link products to insurers and control for spending differences caused by differential financial characteristics design. Spending (decomposed into utilization and prices) is one of the key outcome measure in this project, and product characteristics is important to control for.

Identifying the plan will also allow us to bring in external information on provider networks and provider tier lists, which indicates differential cost-sharing levels for different providers. This information will be an important predictor of how patients choose from among competing providers.

VII. DATA ENHANCEMENTS REQUESTED

State and federal privacy laws limit the release and use of CHIA 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 Organizations 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.

a. 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 Codes (standard)	<input checked="" type="checkbox"/> 5-Digit Zip Codes***
<p>***If requested, provide justification for requesting 5-Digit Zip Code. Refer to specifics in your methodology: 5-digit zipcode is necessary to define markets and construct price index per zip code. It will also be used to predict how patients choose between competing providers based on distance from home or work to provider. 3-digit zipcodes would introduce bias as it covers a large geographical area.</p>	

b. 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: In order to construct price/ cost measure for each episode of care, we would have to bridge claims scattered across different days. We need exact date information in order to construct correct price/ cost measures. Also, analyzing the sequence of providers is central to understand the quality of care, which requires month and day on claims.</p>		

c. National Provider Identifier (NPI)

Select one of the following options.

<input type="checkbox"/> Encrypted National Provider Identifiers (standard)	<input checked="" type="checkbox"/> Decrypted National Provider Identifiers***
<p>*** If requested, provide justification for requesting decrypted National Provider Identifier(s). Refer to specifics in your methodology: We would need the decrypted National Provider Identifiers (NPI) in order to link the APCD to external data sources such as NPPES, AMA master file, and the AHA annual survey. This linked data to NPI’s will</p>	

provide information about providers used to predict how patients choose among competing providers as well as quality information about the providers.

In addition, decrypted NPIs are needed in order to aggregate to the level of the provider in cases where a single provider has multiple NPIs. For example, most large hospitals have multiple institutional NPIs, such as separate NPIs for the emergency department and the inpatient wards. If these multiple NPIs are not aggregated to a single provider before running the provider choice regression, then the estimates will be incorrect. CMS's public NPES file will be used to combine NPIs.

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.

[Click here to enter text.](#)

3. Organizations approved to receive Medicaid Data will be required to execute a [Medicaid Acknowledgment of Conditions](#). MassHealth may impose additional requirements on applicants for Medicaid Data as necessary to ensure compliance with federal laws and regulations regarding Medicaid.

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): CMS Insurance Product Attributes Public Use Files

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.

CHIA Data elements will be linked to the following datasets:

(1) Individual provider level data: NPES, AMA physician Masterfile, the Massachusetts Registration of Provider Organization (RPO) data (provider organization and physician roster file collected by the Massachusetts Health Policy Commission), Physician Compare, IRS 990, CMS Open Payment, National Survey of Healthcare Organizations and Systems (NSHOS), and National Electronic Health Records Survey (NEHRS). The linking variables we use is the NPIs.

(2) Individual facility level data: AHA annual survey data, Medicare hospital cost reports, Hospital Compare, IRS990, Irving Levin Health Care Acquisition Reports, NPES, CMS Open payment, [Rand/CMS HCRIS](#) hospital cost data, and CHIA Multi-Source Acute Hospital Financial Dataset. The linking variables we use is the NPIs. Hospital tier and out-of-pocket copay for different plans are taken from public plan documents such as Mass Group Insurance Commission (GIC) documents. This will be linked based on facility NPI and plan ID.

(3) Aggregate data: Census data on zipcode-level socio-economic characteristics, regional classification as urban/ suburban/ rural, and map data to compute distance between centroids of different zip codes to provider addresses or zip codes. The linking variables we use are patient and employer zip-codes as well as provider addresses or zip codes.

(4) Other: CMS Public Use Files on Insurance Product Attributes and Enrollment Statistics and [Mass Group Insurance Commission \(GIC\) documents](#). The linking variables we use is HIOS ID or insurer name for product attributes, and [zip-code for enrollment](#).

Diagnosis variables (MC039, MC041-2, MC143-53, MC071, MC083-88) will be linked to [CMS DRG weights](#) and [AHRQ diagnosis classifications](#) to proxy for the relative costs of different medical services depending on diagnosis; for dental care, Current Dental Terminology lists will be used in a similar way.

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.

For (1) and (2), we would link using the decrypted National Provider Identifiers. For (3), we would link using 5-digit zipcode. For (4), we would use HIOS ID or insurer name for product attributes, and **zip-code for enrollment**; diagnosis variables will be matched deterministically based on their code.

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.

(1) Provider level data: Provider organization, specialty, primary care physician status, pediatrician status, hospitalist status, employed status, practice address, practice characteristics, years in practice, medical/dental school, affiliation with academic medical center, organizational structure of provider organization structure, investment, and profit.

(2) Facility level data:

- Available service, organizational structure, affiliation with other facilities, staffing, investment, and profit.
- Total costs from inpatient services, cost-to-charge ratio (Rand/CMS HCRIS hospital cost data).
- Number of beds, case mix index, weighted average licensed beds, weighted average available beds, inpatient days, discharges, percentage occupancy, average daily census, average length of stay (CHIA hospital cost data).
- Grouping of NPI's into a single unit for providers that have multiple NPI's (CMS public NPPES file).

(3) Zipcode: Income level, population, education attainment, age distribution, employment, racial composition, and urban/ suburban/ rural classification. Patient expected distance to each provider, based on 5-digit zip code and map information (e.g. zipcode tabulation area (ZCTA) distance database).

(4) Other:

- Insurance product premium, metal level, cost sharing, deductible, out-of-pocket maximum, total number of enrollees. **Hospital tier and out-of-pocket copay for different plans.**
- Diagnosis categories and cost weights from CMS, AHRQ, Current Dental Terminology.

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

All external dataset sets linked to will be aggregate- or provider-level data. No external patient level data will be linked to. Thus linking will not increase risk of patient identification. Results will always be aggregated to a level to be presented so that identification is impossible.

X. PUBLICATION / DISSEMINATION / RE-RELEASE

1. Do you anticipate that the results of your analysis will be published or made publicly 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 plan to disseminate the results of the study through publication in a peer-reviewed academic journal. The results we intend to publicize primarily include effect sizes, such as "physicians integrated with insurers spend x% less money every year", or "a one standard deviation increase in the level of competition leads to a y% increase in diabetes care utilization". These results will be displayed for the entire patient and provider population, or subgroups larger than 11 individuals. No results with subgroup less than 11 individuals will be disseminated.

We also intend to publish descriptive statistics of the data that help readers understand the underlying setting, for example, the demographics or medical spending distributions of all patients or providers in the MA health care markets. No results with subgroup less than 11 individuals will be disseminated.

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 intend to give seminar presentations about this research. The presentations are typically to an audience of between 5 and 30 scholars and students. The presentations will consist of similar results to those that will be published in peer-reviewed academic journal, as described above.

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 level of geographic analysis that we expect to include our results is **county** level. If there is such a county with fewer than 11 individuals in the study sample, we will omit results for that county. We will not present maps.

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?

Click here to enter text.

9. If you have answered “yes” to questions 5, 6, 7 or 8, please provide the name and a description of the products, software, services, or tools.

Click here to enter text.

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?

Click here to enter text.

XI. 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.

David Dranove (Principal Investigator) is a health care economist and a nationally recognized expert in antitrust policy. Dranove originated workhorse approaches to measuring consumer demand for networks (Capps, Dranove and Satterthwaite 2003), worked on the relationship between capacity and prices (Dranove, Satterthwaite, and Sindelar 1986), and written extensively in other areas of health economics. He has also served as the lead economics expert in several high-profile healthcare antitrust cases, including *FTC v. St. Luke's Healthcare System* and *U.S. v. Anthem and Cigna*.

Anran Li is an economics PhD student at Northwestern University, where she works under the supervision of Professor David Dranove. She has graduate training in econometrics, statistics, machine learning and economic theory. Anran's research focuses on the industrial organization of health care markets, particularly the implications of policy designs and competition between insurers. She has four years of experience working with claims data from Medicaid and commercial markets from other states. Her paper on quality provision in health insurance market has won the best third year paper award from Northwestern University.

Hyein Cho is an economics PhD student at Northwestern University, Kellogg school of management (managerial economics and strategy department), where she works under the supervision of Professor David Dranove. She has graduate training in microeconomics, econometrics, and statistics. Hyein’s research focuses on competition and integration, with particular emphasis on health care markets. She has two years of experience working with commercial health claims data (national level).

Alexander Boche is a PhD student at Northwestern University, Kellogg school of management (managerial economics and strategy department) under the supervision of David Dranove. He has experience working, as a research assistant to Professor Dranove, with inpatient claims data from the state of Florida as well as the Medicare Physician & Other Practitioners by Provider and Service dataset.

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.)

XII. USE OF AGENTS AND/OR CONTRACTORS

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. The Organization 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:	Click here to enter text.
Company Website	Click here to enter text.
Contact Person:	Click here to enter text.
Title:	Click here to enter text.
E-mail Address:	Click here to enter text.
Address, City/Town, State, Zip Code:	Click here to enter text.
Telephone Number:	Click here to enter text.
Term of Contract:	Click here to enter text.

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

Click here to enter text.

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.

Click here to enter text.

3. Will the agent or contractor have access to and 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 #1 INFORMATION	
Company Name:	Click here to enter text.
Company Website	Click here to enter text.
Contact Person:	Click here to enter text.
Title:	Click here to enter text.
E-mail Address:	Click here to enter text.
Address, City/Town, State, Zip Code:	Click here to enter text.
Telephone Number:	Click here to enter text.
Term of Contract:	Click here to enter text.

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

Click here to enter text.

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.

Click here to enter text.

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]

XIII. 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.

Organizations 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) this research is not funded by a source requiring the release of raw data to that source; (3) that the requested Data is the minimum necessary to accomplish the purposes described herein; (4) 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 (5) to my authority to bind the Organization.

Signature: (Authorized Signatory for Organization)	Drag signature image here or delete and physically sign
Printed Name:	Click here to enter text.
Title:	Click here to enter text.
Date:	Click here to enter text.

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.

Data Management Plan for Use of CHIA Data [Attach to Data Application]

I. INSTRUCTIONS

Any Recipients, contractors, or agents receiving CHIA Data (“Data”) must complete and execute this [Data Management Plan](#). Certain CHIA Data includes Protected Health Information (“PHI” as defined under the Health Insurance Portability and Accountability Act [HIPAA] and its implementing regulations) and all CHIA Data contains elements that may be used to identify an individual. The Data Management Plan(s) will be incorporated within the Data Use Agreement that must be executed prior to receipt of the Data. You may wish to refer to the Data Use Agreement as you complete this Data Management Plan. This Data Management Plan should be completed by the Chief Information Security Officer, Chief Privacy Officer, legal counsel or another officer with sufficient knowledge of the Agency or Organization’s data privacy and security practices and who has authority to bind the Agency or Organization.

NOTE: This Data Management Plan is confidential and will not become a part of the public record.

II. GENERAL INFORMATION

Project Title: (as it appears on Data Application)	Click here to enter text.
Primary Investigator: (as it appears on Data Application)	Click here to enter text.
Organization Requesting CHIA Data (Recipient): (as it appears on Data Application)	Click here to enter text.
Organization Holding CHIA Data under this Data Management Plan: (this entity must be either the Recipient or listed as an Agent/Contractor for the Recipient on the Data Application)	Click here to enter text.

III. CERTIFICATIONS

The undersigned certifies and agrees as follows:

- The Data will be **encrypted at rest on storage media (backup tapes, local hard drives, network storage, et al) with at least AES-256 standard or stronger.**
- The Data will be **encrypted in transit consistent with the approved method(s) described in this Data Management plan at section V.3-b.**
- Anti-malware software or service is active on any server or endpoint containing or accessing the Data or other controls mitigating malware infection and propagation are in place
- If a Covered Entity or Business Associate under HIPAA, the Agency or Organization is in full compliance with the privacy and security requirements of HIPAA; trains all staff who access PHI on the requirements of HIPAA; and has Business Associate Agreements with all non-employees who access PHI.
- The Agency or Organization has policies and procedures in place to address:

- The sharing, transmission and distribution of PHI
- The physical removal, transport and transmission of PHI
- The physical possession and storage of PHI
- The destruction of PHI upon the completion of its use
- Confidentiality agreements with all individuals, including contractors, who will access PHI
- Agreements governing the use and disclosure of PHI with all non-employees who will access PHI

IV. RESPONSIBLE PARTIES

Please identify the following individuals within your Agency or Organization:

1. The individual responsible for organizing, storing and archiving the Data. This individual is the Custodian of the CHIA Data required under Article XI of the Data Use Agreement.

Name:	Click here to enter text.
Agency/Organization:	Click here to enter text.
Title:	Click here to enter text.
Phone:	Click here to enter text.
Address, City/Town, State, Zip Code:	Click here to enter text.
Email:	Click here to enter text.
Reports to (name and title):	Click here to enter text.

2. The individual(s) responsible for the research team using the Data, including ensuring each individual (i) has a signed confidentiality agreement, (ii) accesses and uses only the minimal Data necessary to achieve the research purpose, (iii) accesses the Data only on a secured server according to Applicant's policies. This individual is also responsible for maintaining the access log required under Article II, Section 5 of the Data Use Agreement.

Name:	Click here to enter text.
Agency/Organization	Click here to enter text.
Title:	Click here to enter text.
Phone:	Click here to enter text.
Address, City/Town, State, Zip Code:	Click here to enter text.
Email:	Click here to enter text.
Reports to (name and title):	Click here to enter text.

3. The individual responsible for notifying CHIA of any breach of the Data Use Agreement or this Data Management Plan.

Name:	Click here to enter text.
Organization:	Click here to enter text.
Title:	Click here to enter text.
Phone:	Click here to enter text.
Address, City/Town, State, Zip Code:	Click here to enter text.
Email:	Click here to enter text.
Reports to (name and title):	Click here to enter text.

4. The individual responsible for ensuring the Data is destroyed upon termination of the Data Use Agreement, completing the Data Destruction Form and providing that Form to CHIA.

Name:	Click here to enter text.
Organization:	Click here to enter text.
Title:	Click here to enter text.
Phone:	Click here to enter text.
Address, City/Town, State, Zip Code:	Click here to enter text.
Email:	Click here to enter text.
Reports to (name and title):	Click here to enter text.

V. DATA SECURITY AND INTEGRITY

Agents or contractors that will have access to or store the CHIA Data at a location other than the Recipient's location, or in an off-site server and/or database, must complete a separate [Data Management Plan](#).

1. Physical Location of the Data:

- a. Please provide the delivery address for the Data, as well as the full address, including building and floor, of each location where Data will be delivered and stored.

Delivery:

Organization: Click here to enter text.			
Street Address: Click here to enter text.	City: Click here to enter text.	State: Click here to enter text.	ZIP Code: Click here to enter text.
Office Telephone (<i>Include Area Code</i>): Click here to enter text.			

Storage:

Organization: Click here to enter text.			
Street Address: Click here to enter text.	City: Click here to enter text.	State: Click here to enter text.	ZIP Code: Click here to enter text.
Office Telephone (<i>Include Area Code</i>): Click here to enter text.			

- b. Will the Data be stored by the third party on a system in the cloud (reachable via the Internet)?
 - Yes No
- i. If you answered yes to (b): Has this Cloud Service Provider passed a FedRAMP 3PAO assessment *for the specific cloud system* which will host the data?
 - Yes No

- ii. If you answered yes to (b): What is the name of the provider *and* the FedRAMP level the specific cloud system hosting the data is operating at?

Click here to enter text.

2. Data Privacy Training and Awareness:

- a. Has every individual who will access the Data received training on the proper handling of protected health information and/or personal data within the last year?
- Yes No

3. Encryption of Data:

- a. Will all CHIA Data at rest be encrypted on storage media (backup tapes, local hard drives, network storage, et al) with **encryption at least AES-256 or stronger**. *All holders of CHIA Data must encrypt data at rest a Data Management Plan indicating otherwise will not be approved and returned for revisions.*

Yes No

- b. Will CHIA Data be transmitted by your Agency or Organization over the Internet?

Yes No

If you answered yes to (b): which of the following if any are used when transmitting data over the internet? If selecting *other* please describe the method in space provided below.

SSL (meets or exceeds TLS 1.1 or TLS 1.2) SFTP other

Click here to enter text.

4. Information Security:

- a. Does your Agency or Organization have published information security policies which are followed and accessible to all staff accessing or handling CHIA Data?

Yes No

- b. Has every individual who will access the CHIA Data received cyber security awareness training in the last year?

Yes No

- c. Has your Agency or Organization experienced a breach of PHI or Personally Identifiable Information in the last seven (7) years?

Yes No

- i. *If you answered yes to (c): how was the breach resolved and what steps were taken to prevent a recurrence?*

Click here to enter text.

5. Technical and Physical Controls:

- a. Are all the user accounts that log on to any machine (server or endpoint) that accesses the Data uniquely assigned to individual users (i.e., the user accounts are not shared)?
 Yes No
- b. Is an audit log maintained of all user log-on attempts to the system hosting the CHIA Data?
 Yes No
- c. What is the minimum password length and character complexity (uppercase, lowercase, numeric, and special characters) required for new passwords on the user accounts logging on to the system accessing the CHIA Data?

Click here to enter text.

- d. Describe any additional authentication technical security controls you employ to defend the system against unauthorized logon, e.g. maximum failed login attempts, lockout period, etc.:

Click here to enter text.

- e. Do you run a current version of a commercial off-the-shelf anti-virus or anti-malware product on the systems that will host or access the CHIA Data?
 Yes No

If you answered no to (e), please provide details within your response in question (g) regarding additional controls implemented to prevent malware infection and propagation.

- f. If the CHIA Data will be on a server or network accessible storage drive, then check all the security features present in the room containing CHIA Data:
- i. Recorded video
 - ii. Access log of all individuals entering the room
 - iii. Secure server rack
 - iv. Access control limiting access only to authorized individuals
- g. What additional specific physical or technical safeguards (not mentioned in prior answers) will be used to *mitigate* the risk of unauthorized access to CHIA Data?

Click here to enter text.

- h. When was the last information security risk assessment performed in your Agency or Organization? Who conducted it?

Click here to enter text.

- i. When was the last IT audit performed in your Agency or Organization? Who conducted it?

Click here to enter text.

VI. DATA DESTRUCTION

The Recipient attests that the CHIA Data and all copies of the CHIA Data used by the Applicant or its employees, contractors, or agents will be destroyed upon Project Completion or termination of the Data Use Agreement. All data destruction must conform to the requirements of [M.G.L. c. 93I](#) and to the Data Use Agreement. Please specify below the technical measures you will use to meet these requirements.

Click here to enter text.

VII. ATTESTATION

By submitting this Data Management Plan, the Agency or 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 Agency or 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 its agents.

By signature below, I attest: (1) to the accuracy of the information provided herein (2) that the Agency or Organization agrees to hold and/or access CHIA Data at all times in compliance with all provisions of this Data Management Plan and the Data Use Agreement; and (3) to my authority to bind the Agency or Organization undersigned as an authorized signatory of the Agency or Organization.

Signature: (Authorized Signatory for Agency or Organization holding CHIA Data)	Drag signature image here or delete and physically sign
Printed Name:	Click here to enter text.
Title:	Click here to enter text.
Agency or Organization Holding CHIA Data:	Click here to enter text.
Date:	Click here to enter text.