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

I. INSTRUCTIONS

This form is required for 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 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. ORGANIZATION AND INVESTIGATOR INFORMATION

Project Title:	**Effects of Utilization Patterns and Coordination of Care on Outcomes**
IRBNet Number:	
Organization Requesting Data:	University of Massachusetts (Amherst)
Organization Website:	https://www.umass.edu/
Authorized Signatory for Organization:	
Title:	
E-mail Address:	
Address, City/Town, State, Zip Code:	
Primary Investigator:	**Kimberley Geissler**
Title:	Assistant Professor of Health Policy and Management
E-mail Address:	kgeissler@umass.edu
Telephone Number:	415-847-6781
Names of Co-Investigators:	Keith Ericson, Amanda Starc
E-mail Addresses of Co-Investigators:	kericson@bu.edu, amanda.starc@kellogg.northwestern.edu

III. FEE INFORMATION

1. Consult the Fee Schedule for All-Payer Claims Database data and select one of the following options:

☒ Researcher
☐ Other
☐ Reseller

2. Are you requesting a fee waiver?

☒ Yes
☐ No

3. Complete and submit the Fee Remittance Form. If requesting a fee waiver, submit a letter stating the basis for your request (if required). Please refer to the Fee Schedule (effective Feb 1, 2017) for fee waiver criteria.

IV. PROJECT INFORMATION

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
☐ Reference tool   ☒ Research studies   ☐ Severity index tool
☐ Surveillance   ☐ Student research   ☐ Utilization review of resources
☐ Inclusion in a product   ☐ Other (describe in box below)

2. Provide a summary of the specific purpose and objectives of your Project. This may include research questions and/or business use Projects.

We investigate the relationship between coordination of care and utilization patterns and cost, access, quality, utilization, and health outcomes. This includes questions such as 1. What are the impacts of smaller physician networks and other aspects of health insurance plans on coordination of care and utilization? 2. What impact does coordination of care between treatment settings have on outcomes including quality, cost, and health outcomes? 3. What do professional networks of shared patients among physicians look like,and how do such networks vary by type of insurance plan (e.g. HMO v PPO vs. Medicaid)? 4. What is the relationship between networks of physicians, resource use, and patient outcomes?

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

☒ Yes [*If yes, a copy of the approval letter and protocol <u>must</u> be included with the Application package on IRBNet.*]
☐ No, this Project is not human subject research and does not require IRB review.

4. <u>**Research Methodology**</u>: 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. *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.*

> In this project, we will examine the relationship between coordination of care and utilization patterns and cost, access, quality, utilization, and health outcomes. We will use enrollment and claims data to examine different measures of coordination of care and utilization patterns, including standard coordination measures from the health services literature, social network analysis, referral concentration, time to treatment, distance to treatment, and other characteristics of treatment patterns. We will examine influences on these coordination/utilization pattern measures as well as the influence of coordination/utilization pattern measures on outcomes.

VI. DATASETS REQUESTED

1. Specify below the dataset(s) and year(s) of data requested for this Project, and provide your justification for requesting *each* dataset.

☒ **Medical Claims**
☒ 2011 ☒ 2012 ☒ 2013 ☒2014 ☒2015
Describe how your research objectives require Medical Claims data: We require medical claims to determine individual level coordination of care, spending, quality measures, and other outcome measures.
☒ **Pharmacy Claims**
☒2011 ☒2012 ☒2013 ☒2014 ☒2015
Describe how your research objectives require Pharmacy Claims data: We require pharmaceutical claims to understand full medical spending and potential impacts of coordination of care on pharmaceutical prescribing patterns.
☒ **Dental Claims**
☒2011 ☒2012 ☒2013 ☒2014 ☒2015
Describe how your research objectives require Dental Claims data:

We require dental claims to determine individual level coordination of care, spending, and process measures, particularly for children and the elderly.

☒ **Member Eligibility**
☒2011 ☒2012 ☒2013 ☒2014 ☒2015

Describe how your research objectives require Member Eligibility data:

We require Medical Eligibility files to determine member eligibility during claims periods, length of enrollment, member characteristics, and to identify individuals without medical claims.

☒ **Provider**
☒2011 ☒2012 ☒2013 ☒2014 ☒2015

Describe how your research objectives require Provider data:

 We require provider files to obtain specialty information about providers, links within organizations, and to link medical claims to other datasets for claims without NPI in the original claim.

☒ **Product**
☒2011 ☒2012 ☒2013 ☒2014 ☐2015

Describe how your research objectives require Product data:

Information about product will be used for linkages to other datasets as well as for obtaining information about insurance plan enrollment.

---

2. All-Payer Claims Database data are refreshed and updated periodically and made available in Release Versions that contain the most recent five calendar years of data. As certain Project objectives may require future years of data not yet available, CHIA will consider requests for additional Release Versions of the *same data (i.e., same elements and files)* without the need to submit a new application. Please note that approved requests will be subject to applicable terms in the Data Use Agreement and fees for additional data. Please indicate below whether this is a one-time request, or if the described Project will require future Release Versions of data and if so, which Versions

☒ One-Time   **OR**   ☐ 2016 ☐ 2017 ☐ 2018 ☐ 2019 ☐ 2020

VII. DATA ELEMENTS 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 additional elements listed below for inclusion in their analyses. Requests for additional elements 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 elements you are requesting in addition to the "Core" LDS, provide your justification for requesting each element.

Geographic Data

The geographic sub-divisions listed below are available for Massachusetts residents and providers only. Choose one of the following geographic options. [*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 or treatment.*]

☐ 3-Digit Zip Code (standard)	☒ 5-Digit Zip Code***
***If requested, provide justification for requesting 5-Digit Zip Code. Refer to specifics in your methodology: We require 5-digit ZIP code as this allows us to better understand geographic influences and to link to specifics about the local area, including information about area socioeconomic status. It also allows us to better identify provider practice locations, which is important in understanding coordination of care.	

Dates

Choose one option from the following options for dates. [*For releases with YYYYMM or YYYYMMDD, CHIA will apply a substance abuse filter which will remove all claims that include a substance abuse diagnosis or treatment.*]

☐ Year (YYYY) (Standard)	☐ Month (YYYYMM) ***	☒ Day (YYYYMMDD) *** [for selected data elements only]
*** If requested, provide justification for requesting Month or Day. Refer to specifics in your methodology: We require the day dates to construct quality measures and utilization patterns such as time to treatment and 30 day readmissions for hospital stays. This is understanding coordination of care between different treatment settings.		

National Provider Identifier (NPI)

Choose one of the following options for National Provider Identifier(s):

☐ Encrypted National Provider Identifier(s) (standard)	☒ Decrypted National Provider Identifier(s)***
*** If requested, provide justification for requesting decrypted National Provider Identifier(s). Refer to specifics in your methodology:	

We will link provider NPI to additional information about provider practices as available in outside datasources. This will allow us to better understand organizational influences on coordination of care.

VIII. MEDICAID 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 Medicaid Data, 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. CHIA cannot release Medicaid Data without approval from MassHealth. This may introduce significant delays in the receipt of Medicaid Data.

This project will examine important questions related to coordination of care, utilization patterns, and outcomes including cost, quality, and access. These questions relate to the efficiency of care provided by the MA Medicaid program, and can provide evidence to Medicaid policymakers on areas in which Medicaid is excelling, as well as potential areas for improvement. Having full information for all individuals in the state allows for the calculation of much more coherent and cohesive coordination of care patterns, which benefits Medicaid, commercial, and Medicare enrollees. Our results will identify patterns of provider networks that patients value and that deliver effective care, giving more information to the Medicaid program about efficient allocation of providers. Additionally, many studies have examined the effects of limited provider networks for Medicaid –we will look at the structure of these networks and their relationships with patient outcomes, particularly as compared to other types of insurance including HMOs. This will potentially inform Medicaid as to the value of expanding networks (if any) in terms of outcomes including cost and utilization measures such as hospitalizations or emergency department visits.

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 propose to link APCD data to the following aggregate datasets, described here: 1)Hospital linkages – We will link hospitals to the American Hospital Association Annual Survey Database (AHA), Massachusetts Acute Hospital Profiles, and American Hospital Databases for hospital characteristics; to the Medicare Hospital Compare dataset for quality and aggregate health outcome data. We will use information on the service provider billing NPI, name, and location to identify hospitals from the medical claims data, linked with location information from the provider file based on the National Service Provider ID. We will link using the hospital name and location we derive from the APCD fields to hospital name and location in the linkage datasets. This is needed to describe provider networks and model patient choice of hospital. 2)Provider linkages – We will link providers to the American Medical Association Physician Masterfile for provider specialty and demographic data (using the National Service Provider ID; if NPI is missing, we will use the provider file to link to the AMA Masterfile by physician name and service location); to tiering (quality/cost-efficient care) measures for specialist providers participating in the GIC UniCare plans; and to the NPPES NPI registry, Massachusetts licensing data, Massachusetts Health Quality Partners (MHQP) provider dataset, SK&A physician dataset, Physician Compare data, publicly available Medicare utilization and referral pattern files to accurately link providers to practices for determination of practice level measures and to link to quality data, utilization, and coordination measures . To link to provider information, we need a number of provider identifiers. In published analyses and reports, we will not identify providers or report information where deductive disclosure would be possible (e.g., we will mask small cells of fewer than 11 patients, etc.). This is needed to characterize provider networks (e.g. are more efficient doctors more likely to be in the same network) and model patient choice of provider. 3)Ambulatory physician group practice linkages – We will link provider data to MHQP quality data on clinical and patient experience measures for primary care physicians (We will do this using provider IDs, names, and practice locations from the medical claims files and the provider file. We will use the physician group information from the MHQP provider dataset linkage). This is needed to characterize provider networks and model patient choice of provider. 4)Geographic area linkages – We will link member geographic data (ZIP from the member eligibility file) to the corresponding geographic indicator in the Area Resource File and the American Communities Survey/Census data to get information on healthcare supply, socioeconomic status, and regional characteristics. This is needed to account for variation in patient characteristics that might affect patient use of medical care or outcomes; we do not identify individual patients, merely link to characteristics of their ZIP code.5)Carrier and/or insurance plan linkages – We will link carrier and/or insurance plan data to market share and premium data from Mass Connector, as well as to the network definition of plans in the Connector, and Health Insurance Marketplace Files. This is needed to model insurer price setting and how utilization and spending moves with plan generosity. 6) Linkages to other all payer claims databases. We will use the MA APCD in conjunction with other claims databases to improve the generalizability of results. This will involve appending (stacking) the information from MA with that from other states, rather than merging on additional data about MA patients

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.

The fields to be used for linkages are described previously, but primarily include the service and billing NPIs, plan identification numbers as available, and patient 5-digit ZIP Code identifiers. If NPIs are not available, physician name and entity name will be used to complete the linkages. If plan identification numbers are not available, available plan information will be used to create linkages.

<div style="border:1px solid black; height:80px;"></div>

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

> As these datasets do not increase the ability or likelihood of identification of individual patients in the linked datasets, these linkages would not jeopardize patient confidentiality. As discussed in the data security and integrity section, we will take great care to ensure the confidentiality of the data.

X. PUBLICATION / DISSEMINATION / RE-RELEASE

1. Describe your plans to publish or otherwise disclose CHIA Data, or any data derived or extracted from CHIA Data, in any paper, report, website, statistical tabulation, seminar, conference, or other setting. 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 in the display of a cell less than 11.

> We will mask cells based on fewer than 11 patients – the Investigators will review all potential publications and presentations in advance to ensure that there is no disclosure.

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

> The results will be published and presented in peer-reviewed publications and presentations at local and national conferences. As allowed by journal access policies, we will post full-texts of the publications on our websites; if this is not possible, they will be available upon email request from the authors.

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, by 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?

```
```

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.

Kimberley H. Geissler: Assistant Professor, School of Public Health and Health Sciences
B.A. Williams College 2006, Ph.D. University of North Carolina 2013, Postdoctoral Fellow: Boston University 2013-2014

Keith Ericson: Associate Professor of Markets, Public Policy and Law, Boston University Questrom School of Business
B.A. Williams College 2004, A.M. Harvard University 2008, Ph.D. Harvard University 2011

Amanda Starc: Associate Professor of Strategy at the Kellogg School of Management and a Faculty Research Fellow at the National Bureau of Economic Research (NBER)
BA, Case Western Reserve 2006; PhD Harvard University 2011

We have each conducted substantial quantitative analysis with previous versions of the MA APCD as well as a number of other health insurance claims datasets (e.g., Marketscan data, Medicare data, hospital claims data, etc.)

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 Agency assumes all responsibility for the use, security and maintenance of the CHIA Data by its agents, including but not limited to contractors. The Agecny 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 amendemtn 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:	Commonwealth Computational Cloud for Data Driven Biology (C3DDB) as part of the The Massachusetts Green High Performance Computing Center Inc
Company Website	http://www.mghpcc.org/resources/computer-systems-at-the-mghpcc/c3ddb/
Contact Person:	Joe Pasquini
Title:	IT Manager, UMass School of Public Health and Health Sciences
E-mail Address:	jpasquini@umass.edu
Address, City/Town, State, Zip Code:	100 Bigelow St, Holyoke, MA 01040
Telephone Number:	*413-552-4900*
Term of Contract:	Open ended

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

Data storage and analysis will be done on the SPHHS compute node which is disconnected from the Internet, is not a cloud based storage solution, and has all the correct compensating controls, including but not limited to: Firewall, secure login, public/private key accounts, encryption for storage, and during transfer through a secure SSH/VPN connection.

This SPHHS compute note is a highly secure partition of the C3DDB, which is a secured portion of the MGHPCC. The SPHHS compute node is itself not connected to the internet, and draws resources from the C3DDB when required. The legal structure of the MGHPCC is such that a Business Associate Agreement (BAA; attached) has been signed between UMass SPHHS and the MGHPCC to ensure that this SPHHS compute node is adequately protected for sensitive health data such as the MD MCDB.

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.

> Please see Data Management Plan and BAA.

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.

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 whicih 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:	

<u>Attachments</u>

A completed Application must have the following documents attached to the Application:

☐ 1. IRB approval letter and protocol (if applicable)

☐ 2. Research Methodology (if protocol is not attached)

☐ 3. CVs of 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 Organization's location, off-site server and/or database)

Applications will not be reviewed until they are complete, including all attachments.

TRACKING TABLE (to be completed by CHIA staff only)	
Complete Application Received	
Application Fee Received	
Data Privacy Committee Review	
Data Release Committee Review	
Linkages Approved (as described)	
Approved for additional Release Versions	
Executive Director Approval	
Data Fee Received	
Date of First Audit	
Extract Number:	

Attachment #1 – IRB Approval Letter & Protocol or Research Methodology

Attachment #2 – Data Management Plan(s)