

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:	Oral health utilization in the state of Massachusetts
IRBNet Number:	1291226-1
Organization Requesting Data (Recipient):	Harvard Medical School/ Harvard School of Dental Medicine
Organization Website:	https://hsdm.harvard.edu/
Authorized Signatory for Organization:	Christopher J. Davies
Title:	Grants & Contracts Officer
E-Mail Address:	spacontracts@hms.harvard.edu
Address, City/Town, State, Zip Code:	25 Shattuck Street, Suite, 509, Boston, MA 02115
Data Custodian: (individual responsible for organizing, storing, and archiving Data)	Nathan Palmer
Title:	Director, Healthcare Data Science Program
E-Mail Address:	Nathan_palmer@hms.harvard.edu
Telephone Number:	781-424-1710
Address, City/Town, State, Zip Code:	10 Shattuck Street, Boston MA 02115
Primary Investigator (Applicant): (individual responsible for the research team using the Data)	Shenam Ticku
Title:	Instructor
E-Mail Address:	Shenam_Ticku@hsdm.harvard.edu
Telephone Number:	617-432-5271
Names of Co-Investigators:	Nathan Palmer, Jane Barrow, Deborah Gordon, Yara Halasa- Rappel
E-Mail Addresses of Co-Investigators:	nathan_palmer@hms.harvard.edu jane_barrow@hsdm.harvard.edu deborah_gordon@hks.harvard.edu Yara_Halasa-Rappel@hsdm.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 checked="" 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 checked="" type="checkbox"/> Surveillance | <input checked="" type="checkbox"/> Student research | <input checked="" 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.

The goal of our research is to assess oral health utilization across categories of geography, practice type, coverage type, patient demographics, services and cost. Our research questions will look at :

1. a) Variations in the percentage of children at high risk who are receiving preventive care
 - b) Variation in care by geography, or by plan,
 - c) Variation in utilization by dentist, region , or demographics such as age
 - d) Variations in utilization by cost
2. Evaluate oral health utilization and overall health outcomes and expenditures among people with chronic diseases
3. Evaluate variations in prescription drugs across
 - a) geography b) practice type c) coverage type d) patient demographics, e) services and f) provider type for the treatment of oral health conditions

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.*

Dental disease is largely preventable, yet 25 percent of the population in the United States has untreated decay, with higher rates among low- income and minority populations. It is estimated that one in ten minority kindergartners is in pain from a toothache. Low- income adults report an average of five incidences of severe tooth pain over the last five years. In 2000 the Surgeon General of the United States called oral disease “ The Silent Epidemic”.

Despite significant disease burden, nearly a quarter of the US population has no dental insurance. According to the ADA Health Policy institute 40 percent of all spending on dental care in 2015 was attributed to out of pocket cost. Americans therefore have both substantial financial exposure for their dental care, and significant purchasing power. However, due to lack of transparency and a payment landscape dominated by a fee for service model which rewards volume, there is potential for waste.

Exploring Massachusetts dental claims is a first step to look at dental trends and engaging consumers in their dental care purchasing and decision making. Also, our research will highlight evidence for policymakers and industry to widen the gamut for public funding for oral health services and include oral health interventions in chronic disease management plans

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

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.

<input checked="" type="checkbox"/> Medical Claims
Describe how your research objectives require Medical Claims data: Our project needs to look at oral health services provided at medical offices
<input checked="" type="checkbox"/> Pharmacy Claims
Describe how your research objectives require Pharmacy Claims data: We are evaluating variations in prescription drug utilization for oral diseases across the state
<input checked="" type="checkbox"/> Dental Claims
Describe how your research objectives require Dental Claims data: Our project will assess the utilization of dental services across geography. We will need several data elements, including but not limited to CDT Code, Cost of the service (Copay Amount, Coinsurance Amount, Deductible Amount), Product ID Number, Tooth Number/Letter, Dental Quadrant and provider zip code.
<input checked="" type="checkbox"/> Member Eligibility

Describe how your research objectives require Member Eligibility data:

We will require member eligibility files to compute rates, determine length of enrollment, age and 5 digit patient residence zip code

Provider

Describe how your research objectives require Provider data:

Product

Describe how your research objectives require Product data:

The product file is needed to distinguish type of insurance product for each beneficiary.

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.

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 need to use the 5 digit zip code as a proxy for obtaining socio-economic characteristics of the population. We will also need a 5 digit zip code to compare utilization across geography

Date Resolution

Select one option from the following options.

<input type="checkbox"/> Year (YYYY) (Standard)	<input checked="" type="checkbox"/> Month (YYYYMM) ***	<input 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>We need the month to validate follow up dates to certain dental procedures, to track utilization through time.</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 will need maximum provider information as we are comparing utilization trends/ differences across providers</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.

Masshealth covers 27.5 percent of Massachusetts' population and has invested significantly to improve the oral health of the Commonwealth. Our research will be incomplete without looking at the utilization patterns of masshealth beneficiaries and providers who serve these beneficiaries. The results of our research will inform masshealth policymakers to understand differences in provider practice and service utilization patterns between mass health enrollees and enrollees of private dental insurance.

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.

Individual Provider Level Data:

CMS NPI Registry data :To obtain additional information about the providers, we will be linking the decrypted NPI element to the publicly available

Aggregate Data:

Census Data: Using patient's zip code in census data, we will get statistics on median income, education etc

Area Resource File: Linked by county to determine healthcare resources, demographics and expenditure

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 linkage will be 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.

CMS NPI Registry: NPI

Census Data: Zip Code

Area Resource file: County

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

Technical: Fields will not contain any identifiable data. Therefore, the the technical merging of the fields will not increase the risk of reidentification, as they are at a larger unit level.

Administrative: All users must covenant to adhere to the terms of the Data Use Agreement and are subject to professional and legal sanctions in the event of attempted misuse or reidentification.

Policy: Implementation of policy under Harvard Research Data Security Policy (RDSP) prohibiting re-identification
When data is presented, it will be aggregated and cell suppressed. In mapping data based on small area 5-digit ZIP code boundaries, methods of data perturbation will be used like introducing statistical noise, controlled rounding, or creating fuzziness by not using exact boundaries but a heat map.

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 intend to disseminate the results of our studies to a variety of sources, including publications in peer reviewed journals, poster presentations, newsletters, seminars, blogs and conferences. Our research protocol is designed to ensure that published data is not identifiable. Scientific review and editorial processes ensure non-identifiability of data. Our results will not divulge any results that would result in the display of a cell less than 11. Senior data science staff must review prior to publication.

Published data is monitored by the responsible researcher as a safeguard.

Scientific journals also have policies in place to prevent publication of identifiable data.

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.

Results will be presented at professional meetings and conferences and available to the public through these avenues.

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?

We will publish geographic data , with maps that will go to the 5 zip code level. If there are instances of fewer than 100 data points we will move to 4 zip code level.

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?

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

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?

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.

Our group as entered into a research collaboration with Aetna, a large commercial health insurance company. Under this agreement, Aetna has shared with us a de-identified copy of all of their commercial medical and pharmacy claims, as well as lab measurements (where they are available) and eligibility information. This data set contains, among other variables, diagnosis codes, procedure codes, plan member zip code and birth year, reimbursement and adjudication information, dates of service, provider details, pharmacy information, NDC drug codes, dispense dates, days’ supply, and health plan coverage details.

The data set contains plan member eligibility and claims from January 2008 through February 2018, comprising tens of billions of diagnosis, procedure and pharmacy facts, all warehoused in a large relational database. As custodians of the data, we have worked with Harvard’s compliance and Information Technology teams to secure the data at rest and in flight with industry standard encryption and firewalls.

We have used this commercial claims data for a number of different epidemiology and public health surveillance studies that have recently been published, notably:

There have been no known data breaches nor publication of identifiable data using these data.

Association of Sex With Recurrence of Autism Spectrum Disorder Among Siblings. Palmer N, Beam A, Agniel D, Eran A, Manrai A, Spettell C, Steinberg G, Mandl K, Fox K, Nelson SF, Kohane I.. *JAMA Pediatr.* 2017 Nov 01; 171(11):1107-1112. PMID: 28973142.

Utilization, Cost, and Outcome of Branded vs Compounded 17-Alpha Hydroxyprogesterone Caproate in Prevention of Preterm Birth. Fried I, Beam AL, Kohane IS, Palmer NP.. *JAMA Intern Med.* 2017 Nov 01; 177(11):1689-1690. PMID: 28973537.

Postsurgical prescriptions for opioid naive patients and association with overdose and misuse: retrospective cohort study. Brat, G., Agniel, D., Beam, A., Yorkgitis, B., Bicket, M., Homer, M., Fox, K., Knecht, D., McMahonill-Walraven, C., Palmer, N., Kohane, I.. *BMJ* 2018; 360 :j5790

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:	
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.

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:	Christopher J. Davies
Title:	Grants & Contracts Officer

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.

[INSERT IRB approval letter and protocol, or research methodology]

