

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 <u>957 CMR 5.02</u>, requesting protected health information. All Applicants must also complete the <u>Data Management Plan</u>, 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 <u>Data Use Agreement</u>. 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 <u>CHIA website</u> in Word and in PDF format or on <u>IRBNet</u> 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 <u>Fee Remittance Form</u> and attach it and all required supporting documentation with your application. Refer to the <u>Fee Schedule</u> (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:	The welfare effects of convenience clinics in the
	Massachusetts health care market
IRBNet Number:	2000023579
Organization Requesting Data (Recipient):	Yale University
Organization Website:	https://www.yale.edu
Authorized Signatory for Organization:	Jennifer Rawlings
Title:	Contract Manager, Corporate Contracts Yale University, Office of Sponsored Projects
E-Mail Address:	jennifer.rawlings@yale.edu
Address, City/Town, State, Zip Code:	25 Science Park - 3rd Floor, 150 Munson St. P.O.B. 208327 New Haven CT 06520-8327
Data Custodian:	Oscar Soler Sanchez
(individual responsible for organizing, storing, and archiving Data)	
Title:	PhD candidate in Economics
E-Mail Address:	Oscar.solersanchez@yale.edu
Telephone Number:	475-655-4319
Address, City/Town, State, Zip Code:	27 Hillhouse Ave, Room B4. New Haven, CT. 06511-8948
Primary Investigator (Applicant):	Jason Abaluck/Oscar Soler Sanchez
(individual responsible for the research team using the Data)	
Title:	Associate Professor of Economics Yale School of Management/PhD candidate in Economics
E-Mail Address:	jason.abaluck@yale.edu/Oscar.solersanchez@yale.edu
Telephone Number:	475-655-4319
Names of Co-Investigators:	
E-Mail Addresses of Co-Investigators:	

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		☐ Utilization review of resources		
☐ Inclusion in a product	☐ 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.

This project is based on a main research question	n, with multiple objectives:	

- 1. What are the effects of convenience clinics (retail clinics and urgent care centers) on the social welfare?
 - a. Determine the causal effects of the entry of convenience clinics in the demand for low-value conditions.
 - b. Identify if there are substitution effects of convenience clinics on Emergency Department usage and primary care.
 - c. Characterize the effects on health care costs.
 - d. Determine the effects of convenience clinics on health care outcomes.
 - e. Characterize the effects of convenience clinics in referrals to specialists and Primary Care Physicians.
 - f. Determine the effects on providers' profits and consumer surplus of the entry of convenience clinics.

3. Has an Inst	itutional Review Board	(IRB) reviewed yo	ur Project?			
⊠ Yes [<i>If yes,</i>	a copy of the approva	l letter and protoco	ol <u>must</u> be included with	the Application	package on IF	RBNet.]

□ 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. 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 rapid expansion of Urgent Care Centers (UCCs) and retail clinics (RCs) in the last 20 years represent an important change in the treatment of emergency low-value conditions (LVCs) in the US health care market. Between 2007 and 2016, approximately 8,000 UCCs appeared in the US (Merchant,2016), and its entire industry reached an annual income of nearly \$25 billions in 2017, treating on average 12,000 patients per year per clinic (Urgent Care Association of America (UCAOA), IBISWorld). However, the welfare effects of the entry of UCCs and RCs in the health care market are unclear. They can reduce waiting times and spending per visit, which can translate in lower costs and higher quality for patients, but also convenience clinics can create over-utilization since patients do not pay the full price of treatment. In addition, if convenience clinics reduce utilization of primary care, it is possible that patients with chronic conditions get worse health care outcomes.

Then, the evidence generated by this research project will be useful to determined if convenience clinics: help reduce health care spending; improve health care outcomes; and benefit patients, providers or both. As a result, this evidence will allow us to inform public policy, and public and private payers if incentivizing the use of UCCs and RCs is an adequate policy, and what mechanisms explained those results.

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 <u>same data files and data elements</u> 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): _2012,2013,2014,2015,2016, _2017_
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 <u>each</u> file.
☑ Medical Claims
Describe how your research objectives require Medical Claims data:
In this project, we will examine the effects of convenience clinics on the health care market. Then, we need medical claims to create measures of utilization and spending per patient to estimate the preferences of patients over providers, and how utilization change with the entry of new convenience clinics. Medical claims also provide the prices pay to providers by insurers and patients, which is an input of the model that will be used. Finally, this data allow us to obtain the referral patterns after patients visit convenience clinics, and how those referral are influenced by changes in the network of providers.
☑ Pharmacy Claims
Describe how your research objectives require Pharmacy Claims data:
Similar to medical claims, we require pharmacy claims to generate measures of utilization and spending per patient.
☐ Dental Claims
Describe how your research objectives require Dental Claims data:

☑ Member Eligibility		
Describe how your research objectives red	quire Member Eligibility data:	
We require member eligibility to determine their primary care physician.	e length of enrollment of patients, member of	characteristics and their information about
⊠ Provider		
Describe how your research objectives rec	quire Provider data:	
The provider file is useful to obtained infor	rmation of providers, and to link them within	organizations.
☑ Product		
Describe how your research objectives rec	guire Product data:	
The product file is required to characterize		
The product me is required to characterize	a each plan.	
VII. DATA ENHANCEMENTS REQUESTED)	
·		
State and federal privacy laws limit the r specific Project objective.	release and use of Data to the minimum	amount of data needed to accomplish a
request the data enhancements listed be	ed in Limited Data Sets (LDS). All applica elow for inclusion in their analyses. Req r each represents the minimum data nec	
For a full list of elements in the release (data dictionaries and similar documenta	(i.e., the core elements and additional election included on CHIA's website.	ements), please refer to <u>release layouts,</u>
1. Specify below which enhancements ye requesting <u>each</u> enhancement.	ou are requesting in addition to the "Cor	re" LDS, provide your justification for
Geographic Subdivisions		
• ,	v are available for Massachusetts resider	nts and providers only. Select <u>one</u> of the
☐ 3-Digit Zip Code (standard)	⊠ 5-Digit Zip Code	***
	requesting 5-Digit Zip Code. Refer to specific	
-	s to the providers is a key element that allow ill allow us to accurately locate practices and	· · · · · · · · · · · · · · · · · · ·
Date Resolution Select <u>one</u> option from the following opt	tions.	
☐ Year (YYYY) (Standard)	☐ Month (YYYYMM) ***	☑ Day (YYYYMMDD) ***

IX. DATA LINKAGE

delays in the receipt of MassHealth Data.

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.

program. CHIA cannot release MassHealth Data without approval from MassHealth. This may introduce significant

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

⊠ Yes

Exhibit A: CHIA Non-Government All-Payer Claims Data Application January 2017 v.1.0	
\square 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 w and the purpose for each linkage.	ill be linked
Individual Provider Level Data The purpose to link MA APCD to individual provider level data is to have additional information on Primary Physicians (PCP) and of referred physicians. Additional information about PCP is useful to control for their characteristics, but also to infer information about their referral patterns. Information about refered physicians group practice affiliations) is important to understand and study referral patterns of physicians in Urgent C (UCC). Please see the IRB document for additional information. Database:	cians (as
American Medical Association (AMA) Physician Masterfile CMS National Plan and Provider Enumeration System (NPPES) Files.	
Individual facility level data. The purpose to link MA APCD to individual facility level data is to control for the characteristics of hospitals clinics (RC) and UCC, but also to account in the structural model for possible vertical relationships between For instance, some UCC are owned by hospitals, and this changes incentives between providers. Database:	
CMS National Plan and Provider Enumeration System (NPPES) Files: Publicly available datasets Merchant Medicine Database of UCC and RC. CMS Provider of Services (POS) File: Publicly available datasets American Hospital Association (AHA) Annua Database: Publicly available dataset to be licensed from AHA. Other similar, publicly available datasets.	al Survey
The purpose to link MA APCD to aggregate data is to characterize the demographic areas of patients. In palevels of education, income and the size of the population are important to control for characteristics of the market, but also to create counterfactual analysis in areas without UCC and RC. Databases: US Census: Publicly available datasets	
Internal Revenue Service (IRS) Statistics of Income (SOI) Individual Income Tax Statistics ZIP Code Data: Publicular available datasets. Other similar, publicly available datasets.	olicly

Medical Claims: National Provider ID (MC026), Provider Characteristics (MC027, MC032, MC036, MC037), Provider

Data elements for linkage:

Location (MC034-MC035)

Pharmacy Claims: National Provider ID (PC021, PC048), Pharmacy Location (PC023-PC024), Provider Location (PC054-PC055)

Member Eligibility File: Health Care Home National Provider ID (ME038), Attributed PCP Provider ID (ME124), Organization ID (ME125)

Provider File: National Provider ID (PV039-PV040), Provider Characteristics (PV015)

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 a deterministic one:

American Medical Association (AMA) Physician Masterfile: using unique provider identifiers as National Provider ID (NPI) and Unique Physician Identification Number (UPIN).

CMS National Plan and Provider Enumeration System (NPPES) Files: using unique provider identifiers as NPI. Merchant Medicine Database of UCC and RC: Using the provider identifier NPI.

CMS Provider of Services (POS) File: System (NPPES) Files: using unique provider identifiers as NPI.

US Census: the linkage for subscriber/patient is based on geographical area characteristics (subscriber ZIP code, member ZIP code, employer ZIP code), and demographics as gender and age. For providers the linkage use geographical area characteristics as provider ZIP code, provider census tract, provider census block). Internal Revenue Service (IRS) Statistics of Income (SOI) Individual Income Tax Statistics ZIP Code Data: the linkage for subscriber/patient is based on geographical area characteristics (subscriber ZIP code, member ZIP code, employer ZIP code), and demographics as gender and age. For providers the linkage use geographical area characteristics as provider ZIP code.

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

Source	Variables to be included
American Medical Association (AMA) Physician Masterfile	Variables of education, training and professional certification information on Doctors of Medicine
CMS National Plan and Provider Enumeration System (NPPES) Files	Healthcare Provider Taxonomy Code Variables, Provider License Variables, provider location variables.
Merchant Medicine Database of UCC and RC	Variables of convenience clinics: location, opening dates, ownership structure and services provided.

CMS Provider of Services (POS) File	Variables of organizational structure, personnel, hospital facilities and services, and financial performance. Also, variables that identifies open, closed, and merged hospitals.
US Census	Variables of income and education
Internal Revenue Service (IRS) Statistics of Income (SOI) Individual Income Tax Statistics ZIP Code Data:	Variables of income, wages and salaries.

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

The datasets that will be link to the MA APCD do not increase the likelihood of the identification of individual patients. In addition, as mention in the Data Management plan, we will take all the possible measures to protect the confidentiality of the data.

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.

The objective is to publish the paper coming out of this research to peer reviewed academic journals. In addition, unpublished results may be shared in seminars, conferences, workshops, and working papers. No disclose outcome will use cell sizes of less than 11 observations.

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 aggregated using econometric methods and graphs. No disclose outcome will use cell sizes of less than 11 observations.

ensure that individuals cannot be identified?
The lowest geographical level of analysis of data Results in presentations or publications will be the city level. Some maps will include data at the zip code level, but will omit cell sizes of less than 11 observations.
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 took, 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?

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

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.
Oscar Soler is a PhD student in the Department of Economics at Yale University. Previously, he worked for the Colombian central bank using confidential data and information used to generate policy recommendations to the board of directors. Also, he worked for Quantil, a consulting firm in Colombia, in projects with public and private agencies with projects that required dealing with proprietary and sensitivy data in the sectors of telecommunications, financial, petrochemistry.
Jason Abaluck is an Associate Professor of Economics at the Yale School of Management. Previously, every paper he has ever published uses claims data. More information on his CV.

Exhibit A: CHIA Non-Government All-Payer Claims Data Application January 2017 v.1.0

2. <u>Resumes/CVs</u>: 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 <u>all</u> agents and contractors who will have access to the CHIA Data. [Add agents or contractors as needed.]

AGENT/CONTRACTOR #1		
INFORMATION		
Company Name:	Yale University Information Technology	
Company Website	http://its.yale.edu	
Contact Person:	Rick Kremer	
Title:	Associate Director, Distributed Support	
E-mail Address:	Rick.Kremer@yale.edu	
Address, City/Town, State, Zip Code:	135 College St, New Haven CT 06510-2483	
Telephone Number:	203.737.7131	
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.

Computer Support. All staff assigned have appropriate training including appropriate management of sensitive data.

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.

All system access is logged.

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 Pla	an <u>must</u> 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:	
	ed to the agent or contractor for this Project and their qualifications for
completing the tasks.	
	and monitoring of the activities and actions of the agent or contractor for this vill ensure the security of the CHIA Data to which the agent or contractor has
	·
3. Will the agent or contractor have acces off-site server and/or database? ☐ Yes ☐ No	ss to or store the CHIA Data at a location other than the Organization's location,
_ -	

Exhibit A: CHIA Non-Government All-Payer Claims Data Application

January 2017 v.1.0

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)	Jennifer Rawlings Digitally signed by Jennifer Rawlings Dix cn=lennifer Rawlings, o, ou, email-jennifer.rawlings00@gmail.com, c=US Date: 2018.09.10 13:43:38-0400'
Printed Name:	Jennifer Rawlings
Title:	Contract Manager, Corporate Contracts

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

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