

# 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 <u>apcd.data@chiamass.gov</u> 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 <u>CHIA</u> 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 <u>Fee Remittance Form</u> 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 <a href="mailto:apcd.data@chiamass.gov">apcd.data@chiamass.gov</a>.
- 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:	Ownership and moral hazard in dental health care
IRBNet Number:	2305827-1
Organization Requesting Data (Recipient):	University of Michigan
Organization Website:	https://umich.edu/
Authorized Signatory for Organization:	Michele Quick
Title:	Contract Officer Ref#25-UFA02980
E-Mail Address:	micheleq@umich.edu
Telephone Number:	734-936-1292
Address, City/Town, State, Zip Code:	3003 S. State St., 1st Floor Wolverine Tower
	Ann Arbor, MI 48109-1274
Data Custodian:	Yuqing Gu
(individual responsible for organizing, storing, and archiving	
Data)	
Title:	PhD candidate
E-Mail Address:	guyuqing@umich.edu
Telephone Number:	7348006095
Address, City/Town, State, Zip Code:	701 Tappan Street, Ann Arbor, MI, 48109
Primary Investigator (Applicant):	Yuqing Gu
(individual responsible for the research team using the Data)	
Title:	PhD candidate
E-Mail Address:	guyuqing@umich.edu
Telephone Number:	7348006095
Address, City/Town, State, Zip Code:	701 Tappan Street, Ann Arbor, MI, 48109
Names of Co-Investigators:	Yesim Orhun
E-Mail Addresses of Co-Investigators:	aorhun@umich.edu

## IV. PROJECT INFORMATION

<u>IMPORTANT NOTE</u>: 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 <u>no</u> other purposes. Use of CHIA Data for other purposes requires a separate Data Application to CHIA <u>or</u> written request to CHIA, with approval being subject to CHIA's regulatory restrictions and approval process. Unauthorized use is a material violation of your Organizations'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		
☐ Reference tool	☑ Research studies	☐ Severity index tool (or other derived input)		
☐ Surveillance		☐ Utilization review of resources		
☐ Inclusion in a product	☑ Other (describe in box below)			

Provide policy makers suggestions on health care market regulation to reduce the health care system cost and improve the accessibility.

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.

We want to evaluate the impact of private equity acquisition on dental utilization rate and provider's billing behaviors. Private equity acquisition will make the dental practice more short-term profit oriented, which will lead to increased utilization and upcoding. Upcoding refers to billing for more expensive procedures than needed. And dental healthcare providers will focus such practices on patients whose out of pocket expenses are lower (i.e., they have higher insurance coverage), which is a form of moral hazard. We also want to investigate long-term effects of any such practices.

Nasseh et al. (2022) showed the percentage of dentists and dental practices affiliated with private equity nearly doubled from 2015 to 2021. As far as we can tell from the estimates reported in their paper, there are around 120 dental clinics being acquired by private equity in Oregon and Massachusetts during our sample period which will form the basis of our analysis sample.

Earlier work showed that private equity affiliated physician practices have higher per patient medical spending compared with non-PE affiliated clinics in ophthalmology, dermatology, and gastroenterology (Singh et al. 2024, Singh et al. 2022). This is likely to happen at the dental area as well because the purpose of private equity acquiring different clinics is to seek short term revenue at the cost of over-utilizing procedures. Einav and Finkelstein (2018) documented that patients with higher insurance coverage receive more expensive medical procedures than patients with low insurance coverage in hospitals. We hypothesize a similar effect in dental clinics.

- 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>: 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.

### 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.

My research aims to evaluate the impact of private equity acquision on the health cost and utilization and accessibility of health. Private equity is short term financially incentive and perhaps push more procedures than needed to make high profit in the short term. I will evaluate whether private equity backed dental clinic do more procedure and more expensive procedures on patients. I will also evaluate whether patients with higher dental insurance coverage suffer more from the

over utilization and health cost increase. The private equity acquisition happens a lot in the past 10 years and my research will provide policy makers suggestions on whether and how to regulate the health care market.

### VI. DATASETS REQUESTED

how much more procedures private equity clinic patients do.

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 fully-insured and self-insured plans. APCD data are refreshed and updated annually and made available to approved data users. For more information about APCD Data, including available years of data and a full list of elements in the database please refer to 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 <u>same data files and data elements</u> 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.	CHIA is currently supporting requests for claims data from 2016 to 2022. Requests made outside of these years may not be supported by CHIA and will be considered on a case-by-case basis. Please specify the years of data that are being requested:2018-2022
3.	Specify below the data files requested for this Project, and provide your justification for requesting <i>each</i> file.
	] Medical Claims
	escribe how your research objectives require Medical Claims data:
C	Click here to enter text.
	] Pharmacy Claims
	escribe how your research objectives require Pharmacy Claims data:
C	Click here to enter text.
X	Dental Claims
D	escribe how your research objectives require Dental Claims data:
	want to evaluate whether private equity backed clinic do more procedures and more expensive procedures on patients compared vith non private equity backed clinics. I need the dental claim data to form difference in difference analysis to evaluate whether and

### **☑** Member Eligibility

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

I want to test whether patients will switch to higher coverage plan after they are exploited by PE clinics. If so, this will further increase the health cost. I need the member eligibility data to identify which patient has dental coverage and whether the patients start to purchase dental insurance plan or switch to higher dental coverage plan after visiting PE clinic.

#### **▼** Provider

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

I want to evaluate the differences in physician billing behaviors between private equity backed clinics and non private equity backed clinics. I will link the provider data with the American Medical Association physician masterfiles by the provider NPI to identify the name and address of the dental practice each physician is associated with. Then I can link the merged file with Pitchbook data throught the name and address of the dental practice. The Pitchbook data contains the name and address of the practice that has been acquired by private equity. Then I can identify whether a physician work in a private equity backed dental clinic through the merger.

#### **☒** Product

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

I want to evaluate the moral hazard issue related to the private equity evasion: whether patients with higher insurance coverage bear more overutilization and cost increase than patients with lower insurance coverage. If so, this will further increase the health cost. I need the insurance plan market and product benefit type to group members into different coverage plan and compare differences between patients group with higher coverage and patient group with lower coverage.

### 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 <u>release</u> <u>layouts</u>, <u>data dictionaries</u> 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 <u>each</u> enhancement.

### a. Geographic Subdivisions

ZIP code and state geographic subdivisions are available for Massachusetts residents and providers only. Small population ZIP codes are combined with larger population ZIP codes. One ZIP Code per person (MEID) per year has been assigned based on the ZIP code/state reported in the member eligibility record's earliest submission year month. If the record does not have an MEID, assignment is based on distinct OrgID/Carrier Specific Unique Member ID.

Non-Massachusetts ZIP codes and state codes except for CT, MA, ME, NH, NY, RI, and VT are suppressed.

Select *one* of the following options.

☐ 3-Digit Zip Codes (standard)	☑ 5-Digit Zip Codes***
***If requested, provide justification for requesting 5-Digit Z We want to evaluate whether private equity backed cl neighborhood level SES metrics like income through t patients.	

#### b. Date Resolution

Select *one* option from the following options.

☐ Year (YYYY) (Standard) ☐ Month (YYYYMM) *** ☐ Day (YYYYMMDD) *		☑ Day (YYYYMMDD) ***	
		[for selected data elements only]	
*** If requested, provide justification for	requesting Month or Day. Refer to specifi	cs in your methodology:	
I want to test whether a clinic become	ne more short term profit oriented and	push more procedures and charger	
higher prices after being acquired by private equity. I need the within clinic level variation and compare the			
utilization and price for the same cli	nic before and after being acquired by	PE. I know the exact date when	
the clinic ownership changed from t	he Pitchbook data and I need the date	of the claim so that I can form the	
before and after comparison.			

### c. National Provider Identifier (NPI)

Select *one* of the following options.

☐ Encrypted National Provider Identifiers (standard)	☑ Decrypted National Provider Identifiers***
*** If requested, provide justification for requesting decrypted	National Provider Identifier(s). Refer to specifics in your
methodology:	
My project aims to test the differences between private equity back	ked clinics and non private equity backed clinics. And the way to
identify whether a provider work in private backed or not is by linl	king the provider NPI data with the American Medical
Association physician Masterfile to get the name and address of th	e practice where the provider works and then merge the practice
address and name with the ownership data stored in the Pitchbook	dataset which records the ownership change for every business

by business name and address. I need the NPI data data so that I can make the link and identify private equity backed clinics.

### VIII. MEDICAID (MASSHEALTH) DATA

1. Please indicate whether you are seeking Medicaid Data:

X Yes
$\square$ No
2. Federal law (42 USC 1396a(a)7) restricts the use of individually identifiable data of Medicaid recipients to uses that are <u>directly connected to the administration of the Medicaid program</u> . If you are requesting MassHealth Data, please describe, in the space below, why your use of the Data meets this requirement. <i>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</i> . Requests for identifiable 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.
My research aims to enhance Medicaid administration by examining the impact of recent private equity involvement in the dental health industry on the dental expenses and accessibility of dental services for Medicaid patients. The short-term financial goals of private equity firms might lead to increased prices and a higher number of procedures for Medicaid patients, raising the overall costs for the Medicaid system. There is also a possibility that private equity-owned clinics may prefer to accept more or exclusively non-Medicaid patients due to the limited dental coverage offered by Medicaid compared to commercial insurance. This could restrict access to dental care for Medicaid patients, especially in geographies where many dental practices are acquired by private equity.
Specifically, I will analyze differences between private equity-owned dental clinics and those not owned by private equity in terms of: 1) procedure pricing applied to Medicaid and non-Medicaid patients, 2) the frequency of physician-induced procedures for Medicaid and non-Medicaid patients, and 3) the proportion of patients served who are covered by Medicaid.
My research aims to provide valuable insights for the Medicaid administration that will be useful in potential negotiations with dental providers to manage costs and considerations for expanding dental coverage to improve service accessibility for patients. Following a thorough evaluation of these issues, I will write an academic research paper, which will be available to MassHealth.
3. Organizations approved to receive Medicaid Data will be required to execute a <u>Medicaid Acknowlegment of Conditions</u> . 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 Physician Masterfile)  □ Individual Facility Level Data (e.g., American Hospital Association 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.  1. To identify which provider is associated with private equity, I will link the CHIA data to the American Medical Association physician masterfile and Pitchbook data through a two-step process. First, I will link the CHIA data to the AMA physician masterfile using the provider's NPI, which will allow me to obtain the name and address of the practice associated with each provider who field the claim. Next, I will merge the linked CHIA-AMA physician masterfile from the first step with Pitchbook data based on the practice's name and address. The Pitchbook data includes the name and address of private equity-backed clinics; since either the name or the address may be imperfectly recorded, both variables will be used for the merge and verification. This linkage will determine whether a provider is associated with a private equity-backed clinic. After completing these steps, I will delete the provider NPI and the practice's name and address.  2. To test whether private equity clinics select higher income patients, I will link the CHIA data to the American Community Survey (ACS) data using the patient's zip code. This linkage will determine the income level of patients. After completing the linkage, I will delete the zip code data.  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 in
<ul> <li>☑ Individual Provider Level Data (e.g., American Medical Association Physician Masterfile)</li> <li>☐ Individual Facility Level Data (e.g., American Hospital Association data)</li> <li>☑ Aggregate Data (e.g., Census data)</li> <li>☑ Other (please describe):</li> <li>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.</li> <li>1. To identify which provider is associated with private equity, I will link the CHIA data to the American Medical Association physician masterfile and Pitchbook data through a two-step process. First, I will link the CHIA data to the AMA physician masterfile using the provider's NPI, which will allow me to obtain the name and address of the practice associated with each provider who filed the claim. Next, I will merge the linked CHIA-AMA physician masterfile from the first step with Pitchbook data based on the practice's name and address. The Pitchbook data includes the name and address of private equity-backed clinics; since either the name or the address may be imperfectly recorded, both variables will be used for the merge and verification. This linkage will determine whether a provider is associated with a private equity-backed clinic. After completing these steps, I will delete the provider NPI and the practice's name and address.</li> <li>2. To test whether private equity clinics select higher income patients, I will link the CHIA data to the American Community Survey (ACS) data using the patient's zip code. This linkage will determine the income level of patients. After completing the linkage, I will delete the zip code data.</li> <li>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.</li> <li>1. To identify which provider is associated</li></ul>
□ 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.  1. To identify which provider is associated with private equity, I will link the CHIA data to the American Medical Association physician masterfile and Pitchbook data through a two-step process. First, I will link the CHIA data to the AMA physician masterfile using the provider's NPI, which will allow me to obtain the name and address of the practice associated with each provider who filed the claim. Next, I will merge the linked CHIA-AMA physician masterfile from the first step with Pitchbook data based on the practice's name and address. The Pitchbook data includes the name and address of private equity-backed clinics; since either the name or the address may be imperfectly recorded, both variables will be used for the merge and verification. This linkage will determine whether a provider is associated with a private equity-backed clinic. After completing these steps, I will delete the provider NPI and the practice's name and address.  2. To test whether private equity clinics select higher income patients, I will link the CHIA data to the American Community Survey (ACS) data using the patient's zip code. This linkage will determine the income level of patients. After completing the linkage, I will delete the zip code data.  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.
<ul> <li>☑ Other (please describe):</li> <li>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.</li> <li>1. To identify which provider is associated with private equity, I will link the CHIA data to the American Medical Association physician masterfile and Pitchbook data through a two-step process. First, I will link the CHIA data to the AMA physician masterfile using the provider's NPI, which will allow me to obtain the name and address of the practice associated with each provider who filed the claim. Next, I will merge the linked CHIA-AMA physician masterfile from the first step with Pitchbook data based on the practice's name and address. The Pitchbook data includes the name and address of private equity-backed clinics; since either the name or the address may be imperfectly recorded, both variables will be used for the merge and verification. This linkage will determine whether a provider is associated with a private equity-backed clinic. After completing these steps, I will delete the provider NPI and the practice's name and address.</li> <li>2. To test whether private equity clinics select higher income patients, I will link the CHIA data to the American Community Survey (ACS) data using the patient's zip code. This linkage will determine the income level of patients. After completing the linkage, I will delete the zip code data.</li> <li>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.</li> <li>1. To identify which provider is associated with private equity, in step 1, I will merge CHIA data with the AMA physician masterfile using exact matching, as the NPI is a reliable identifier. In step 2, I will combine the results from step 1 with Pitchbook data using a fuzzy matching algorithm followed by</li></ul>
<ul> <li>☑ Other (please describe):</li> <li>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.</li> <li>1. To identify which provider is associated with private equity, I will link the CHIA data to the American Medical Association physician masterfile and Pitchbook data through a two-step process. First, I will link the CHIA data to the AMA physician masterfile using the provider's NPI, which will allow me to obtain the name and address of the practice associated with each provider who filed the claim. Next, I will merge the linked CHIA-AMA physician masterfile from the first step with Pitchbook data based on the practice's name and address. The Pitchbook data includes the name and address of private equity-backed clinics; since either the name or the address may be imperfectly recorded, both variables will be used for the merge and verification. This linkage will determine whether a provider is associated with a private equity-backed clinic. After completing these steps, I will delete the provider NPI and the practice's name and address.</li> <li>2. To test whether private equity clinics select higher income patients, I will link the CHIA data to the American Community Survey (ACS) data using the patient's zip code. This linkage will determine the income level of patients. After completing the linkage, I will delete the zip code data.</li> <li>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.</li> <li>1. To identify which provider is associated with private equity, in step 1, I will merge CHIA data with the AMA physician masterfile using exact matching, as the NPI is a reliable identifier. In step 2, I will combine the results from step 1 with Pitchbook data using a fuzzy matching algorithm followed by</li></ul>
<ol> <li>Other (please describe):</li> <li>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.</li> <li>To identify which provider is associated with private equity, I will link the CHIA data to the American Medical Association physician masterfile and Pitchbook data through a two-step process. First, I will link the CHIA data to the AMA physician masterfile using the provider's NPI, which will allow me to obtain the name and address of the practice associated with each provider who filed the claim. Next, I will merge the linked CHIA-AMA physician masterfile from the first step with Pitchbook data based on the practice's name and address. The Pitchbook data includes the name and address of private equity-backed clinics; since either the name or the address may be imperfectly recorded, both variables will be used for the merge and verification. This linkage will determine whether a provider is associated with a private equity-backed clinic. After completing these steps, I will delete the provider NPI and the practice's name and address.</li> <li>To test whether private equity clinics select higher income patients, I will link the CHIA data to the American Community Survey (ACS) data using the patient's zip code. This linkage will determine the income level of patients. After completing the linkage, I will delete the zip code data.</li> <li>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.</li> <li>To identify which provider is associated with private equity, in step 1, I will merge CHIA data with the AMA physician masterfile using exact matching, as the NPI is a reliable identifier. In step 2, I will combine the results from step 1 with Pitchbook data using a fuzzy matching algorithm followed b</li></ol>
<ol> <li>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.</li> <li>1. To identify which provider is associated with private equity, I will link the CHIA data to the American Medical Association physician masterfile and Pitchbook data through a two-step process. First, I will link the CHIA data to the AMA physician masterfile using the provider's NPI, which will allow me to obtain the name and address of the practice associated with each provider who filed the claim. Next, I will merge the linked CHIA-AMA physician masterfile from the first step with Pitchbook data based on the practice's name and address. The Pitchbook data includes the name and address of private equity-backed clinics; since either the name or the address may be imperfectly recorded, both variables will be used for the merge and verification. This linkage will determine whether a provider is associated with a private equity-backed clinic. After completing these steps, I will delete the provider NPI and the practice's name and address.</li> <li>2. To test whether private equity clinics select higher income patients, I will link the CHIA data to the American Community Survey (ACS) data using the patient's zip code. This linkage will determine the income level of patients. After completing the linkage, I will delete the zip code data.</li> <li>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.</li> <li>1. To identify which provider is associated with private equity, in step 1, I will merge CHIA data with the AMA physician masterfile using exact matching, as the NPI is a reliable identifier. In step 2, I will combine the results from step 1 with Pitchbook data using a fuzzy matching algorithm followed by</li> </ol>
<ol> <li>To identify which provider is associated with private equity, I will link the CHIA data to the American Medical Association physician masterfile and Pitchbook data through a two-step process. First, I will link the CHIA data to the AMA physician masterfile using the provider's NPI, which will allow me to obtain the name and address of the practice associated with each provider who filed the claim. Next, I will merge the linked CHIA-AMA physician masterfile from the first step with Pitchbook data based on the practice's name and address. The Pitchbook data includes the name and address of private equity-backed clinics; since either the name or the address may be imperfectly recorded, both variables will be used for the merge and verification. This linkage will determine whether a provider is associated with a private equity-backed clinic. After completing these steps, I will delete the provider NPI and the practice's name and address.</li> <li>To test whether private equity clinics select higher income patients, I will link the CHIA data to the American Community Survey (ACS) data using the patient's zip code. This linkage will determine the income level of patients. After completing the linkage, I will delete the zip code data.</li> <li>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.</li> <li>To identify which provider is associated with private equity, in step 1, I will merge CHIA data with the AMA physician masterfile using exact matching, as the NPI is a reliable identifier. In step 2, I will combine the results from step 1 with Pitchbook data using a fuzzy matching algorithm followed by</li> </ol>
Medical Association physician masterfile and Pitchbook data through a two-step process. First, I will link the CHIA data to the AMA physician masterfile using the provider's NPI, which will allow me to obtain the name and address of the practice associated with each provider who filed the claim. Next, I will merge the linked CHIA-AMA physician masterfile from the first step with Pitchbook data based on the practice's name and address. The Pitchbook data includes the name and address of private equity-backed clinics; since either the name or the address may be imperfectly recorded, both variables will be used for the merge and verification. This linkage will determine whether a provider is associated with a private equity-backed clinic. After completing these steps, I will delete the provider NPI and the practice's name and address.  2. To test whether private equity clinics select higher income patients, I will link the CHIA data to the American Community Survey (ACS) data using the patient's zip code. This linkage will determine the income level of patients. After completing the linkage, I will delete the zip code data.  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.
To identify which provider is associated with private equity, in step 1, I will merge CHIA data with the AMA physician masterfile using exact matching, as the NPI is a reliable identifier. In step 2, I will combine the results from step 1 with Pitchbook data using a fuzzy matching algorithm followed by
AMA physician masterfile using exact matching, as the NPI is a reliable identifier. In step 2, I will combine the results from step 1 with Pitchbook data using a fuzzy matching algorithm followed by
AMA physician masterfile using exact matching, as the NPI is a reliable identifier. In step 2, I will combine the results from step 1 with Pitchbook data using a fuzzy matching algorithm followed by
<ul> <li>manual verification. This will involve initially performing fuzzy matching based on name and address strings. For any clinics identified as private equity-owned in Pitchbook that do not match the step 1 output, I will conduct a manual search, including online searches for alternative names and addresses, to ensure accurate matching.</li> <li>To test whether private equity clinics select higher income patients, I will use the exact matching when I merge the CHIA data with ACS data as the zip code is a reliable identifier.</li> </ul>

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.

Here is a complete listing of all variables to be included in the final linked analytic file:

- 1. Provider NPI (CHIA Data and AMA physician masterfile)
- 2. The name of the practice (AMA physician masterfile)
- 3. The address of the practice (AMA physician masterfile)
- 4. Whether the clinic is owned by private equity (Pitchbook data)
- 5. Average household income level (ACS data)

6. If yes, pl	ease identify	the specific	steps you v	vill take to	prevent th	ne identific	ation of in	ıdividual 1	patients i	n the
linked datas	set.									

- 1. To identify which provider is associated with private equity, my linkage process does not involve any information about individual patients. Initially, I will extract the provider information into a separate file that contains no patient data. The linkage will be performed using only this provider information. Subsequently, I will merge the indicator column, which specifies whether a provider is affiliated with a private equity-backed clinic, back into the patient-level claim data. Finally, I will remove the provider NPI column to ensure data privacy is maintained.
- 2. To test whether private equity clinics select higher income patients, I will merge the CHIA data with aggregate level household income data and the linkage process does not involve any individual level demographic features.

### 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 final output of this project will be published in a peer-reviewed academic journal. I will not report statistic of any subgroup with less than 11 cells.

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.
The final output paper will be disseminated to the public.
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?
State level. Maps will not be presented.
<ul><li>4. Will you be using CHIA Data for consulting purposes?</li><li>☐ Yes</li><li>☑ No</li></ul>
<ul><li>5. Will you be selling standard report products using CHIA Data?</li><li>☐ Yes</li><li>☒ No</li></ul>
<ul><li>6. Will you be selling a software product using CHIA Data?</li><li>☐ Yes</li><li>☑ No</li></ul>
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
<ul><li>8. Will you be reselling CHIA Data in any format not noted above?</li><li>☐ Yes</li><li>☑ No</li></ul>
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.

The PI, Yuqing Gu worked with marketscan claim data and the co-PI and faculty advisor, Yesim Orhun, worked with various types of claim data including marketscan data and medicare data.

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.)

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

AGENT/CONTRACTOR #1 INFO	ORMATION
Company Name:	Click here to enter text.
Company Website	Click here to enter text.
<b>Contact Person:</b>	Click here to enter text.
Title:	Click here to enter text.
E-mail Address:	Click here to enter text.
Address, City/Town, State, Zip	Click here to enter text.
Code:	
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.

for this Project, including how the Org contractor has access.	ganization will ensure the security of the CHIA Data to which the agent or
Click here to enter text.	
3. Will the agent or contractor have ac Organization's location, off-site server	cess to and store the CHIA Data at a location other than the r and/or database?
□ Yes	
$\square$ No	
4. If yes, a separate Data Management	Plan <u>must</u> be completed by the agent or contractor.
AGENT/CONTRACTOR #1 INFO	ORMATION
Company Name:	Click here to enter text.
Company Website	Click here to enter text.
<b>Contact Person:</b>	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 assi completing the tasks.	gned to the agent or contractor for this Project and their qualifications for
Click here to enter text.	
	tht and monitoring of the activities and actions of the agent or contractor ganization will ensure the security of the CHIA Data to which the agent or
Click here to enter text.	
3. Will the agent or contractor have ac location, off-site server and/or database	cess to or store the CHIA Data at a location other than the Organization's e?
□ Yes	
□ No	

2. Describe the Organization's oversight and monitoring of the activities and actions of the agent or contractor

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

4. If yes, a separate Data Management Plan <u>must</u> 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 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)	Michele Quick  S Verified by signNow 04/02/2025 13:27:37 UTC Drag signature image Scheverssches and physically sign
Printed Name:	Michele Quick
Title:	Contract Officer, ORSP, Ref #25-UFA02980
Date:	Click here to enter text. 04/02/2025

#### 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.