Commonwealth of Massachusetts
Center for Health Information & Analysis (CHIA)
Non-Government Agency Application for Data

This application is to be used by all applicants, except Government Agencies, as defined in 957 CMR 5.02.

***NOTE*: In order for your application to be processed, you must submit the required application fee. Please consult the fee schedules for APCD and Case Mix data for the appropriate fee amount. A remittance form with instructions for submitting the application fee is available on the CHIA website.**

I. GENERAL INFORMATION

APPLICANT INFORMATION	
Applicant Name:	Ariel Pakes (primary applicant)
Title:	Professor of Economics
Organization:	Harvard University and the National Bureau for Economic Research
Co-Investigator:	Katherine Ho
Title:	Associate Professor of Economics
Organization:	Columbia University and the National Bureau for Economic Research
Co-Investigator:	Mark Shepard
Title:	Doctoral Student
Organization:	Harvard University
Co-Investigator:	Karen Stockley
Title:	Doctoral Student
Organization:	Harvard University
Project Title:	Prices, Incentives and Hospital-Physician Integration in Health Care
Date of Application:	February 28, 2014
Project Objectives (240 character limit)	To evaluate the implications of recent changes in provider affiliations for negotiated payment rates, referral patterns and treatment intensity for different insurer-determined payment incentives for providers to reduce costs and improve quality.
Project Research Questions (if applicable)	1. To what extent have ownership and contractual relationships between providers, including physician groups, hospitals, and outpatient facilities, changed in recent years? 2. How have these changes in provider affiliations affected the payment rates negotiated between providers and insurers?

	3. Do employers serve as effective agents for their employees in their choice of insurance plans, particularly in how they value the network inclusion of affiliated provider groups that are able to negotiate higher pament rates? 4. Did physician referrals and the intensity of treatment respond to these changes in provider affiliations and payment rates? If so, did the response differ by the type of reimbursement method? 5. What are the implications of these changes in provider consolidation for cost and quality of care and how should that inform antitrust policy with respect to vertical and horizontal mergers?

Please indicate if you are a Researcher, Payer, Provider, Provider Organization or Other entity and whether you are seeking data pursuant to 957 CMR 5.04 (De-Identified Data), 957 CMR 5.05 (Direct Patient Identifiers for Treatment or Coordination of Care), or 957 CMR 5.06 (Discretionary Release).

x	Researcher	957 CMR 5.04 (De-identified Data)
	Payer	957 CMR 5.05 (Direct Patient Identifiers)
	Provider / Provider Organization	957 CMR 5.06 (Discretionary Release)
	Other	

II. PROJECT SUMMARY

Briefly describe the purpose of your project and how you will use the requested CHIA data to accomplish your purpose.

We plan to investigate the changes in ownership and contractual relationships between providers (physician groups, outpatient facilities and hospitals) in Massachusetts in recent years, their relationship with provider payment mechanisms, and their impact on provider payment rates, referral patterns, treatment intensity and costs. This requires that we establish links between plans, the providers in their networks, medical care utilization and payments to providers. We will begin by using the CHIA data, linked with the SK&A office-based physician dataset, to identify changes in ownership and affiliation between providers over time and describe how they interact with the introduction by private insurers of physician incentives to control costs. We will then document the extent to which patient referrals to particular hospitals, treatments chosen for particular conditions, and payment rates to physicians and hospitals are associated with changes in provider ownership and differential provider incentives.

The second stage of the project will estimate models of patient referrals to hospitals and outpatient units, and physicians' treatment choices for their patients conditional on diagnosis and severity, that take into account the affiliations and incentives of referring physicians. We will then estimate models of the price negotiations between insurers and providers to help understand how the size and scope of provider organizations affects their ability to negotiate high prices. As an input into the price negotiation model, we will also estimate employer and individual demand for insurance. These insurance demand

estimates will also allow us to investigate the extent to which employers serve as effective agents for their employees in their choice of insurance plans, particularly in how they value the network inclusion of affiliated provider groups that are able to negotiate higher pament rates. Together the analyses will provide a framework in which we can analyze the impact of both different incentive schemes and provider mergers on costs and quality of care that will be informative to policy makers. They will also help us understand providers' private incentives to consolidate given the cost-control incentives set by insurers.

These analyses will require a dataset that identifies patient visits, diagnoses, physician group affiliation, prices paid to hospitals and physician incentives (e.g. capitation, P4P, shared savings). The APCD will be the primary dataset for our analyses.

III. FILES REQUESTED

Please indicate the databases from which you seek data, the Level(s) and Year(s) of data sought.

ALL PAYER CLAIMS DATABASE	Level 1[1] or 2[2]	Single or Multiple Use	Year(s) Of Data Requested Current Yrs. Available 2009 – 2012
☑ Medical Claims	☐ Level 1 ☑ Level 2	Single	☑ 2009 ☑ 2010 ☑ 2011 ☑ 2012
☑ Pharmacy Claims	☐ Level 1 ☑ Level 2	Single	☑ 2009 ☑ 2010 ☑ 2011 ☑ 2012
☐ Dental Claims ☑ Member Eligibility ☑ Provider ☑ Product	☐ Level 2 ☑ Level 2 ☑ Level 2 ☑ Level 2	Select... Single Single Single	☑ 2009 ☑ 2010 ☑ 2011 ☑ 2012

CASEMIX	Level 1 – 6	Fiscal Years Requested
Inpatient Discharge	☐ Level 1 – No Identifiable Data Elements ☐ Level 2 – Unique Physician Number (UPN) ☐ Level 3 – Unique Health Information Number (UHIN) ☐ Level 4 – UHIN and UPN ☐ Level 5 – Date(s) of Admission; Discharge; Significant Procedures ☐ Level 6 – Date of Birth; Medical Record Number; Billing Number	1998-2012 Available (limited data 1989-1997)

[1] Level 1 Data: De-identified data containing information that does not identify an individual patient and with respect to which there is no reasonable basis to believe the data can be used to identify an individual patient. This data is de-identified using standards and methods required by HIPAA.

[2] Level 2 (and above) Data: Includes those data elements that pose a risk of re-identification of an individual patient.

Outpatient Observation	☐ Level 1 – No Identifiable Data Elements ☐ Level 2 – Unique Physician Number (UPN) ☐ Level 3 – Unique Health Information Number (UHIN) ☐ Level 4 – UHIN and UPN ☐ Level 5 – Date(s) of Admission; Discharge; Significant Procedures ☐ Level 6 – Date of Birth; Medical Record Number; Billing Number	2002-2012 Available
Emergency Department	☐ Level 1 – No Identifiable Data Elements ☐ Level 2 – Unique Physician Number (UPN) ☐ Level 3 – Unique Health Information Number (UHIN) ☐ Level 4 – UHIN and UPN; Stated Reason for Visit ☐ Level 5 – Date(s) of Admission; Discharge; Significant Procedures ☐ Level 6 – Date of Birth; Medical Record Number; Billing Number	2000-2012 Available

IV. FEE INFORMATION

Please consult the fee schedules for APCD (Administrative Bulletin 13-11) and Case Mix data (Administrative Bulletin 13-09) and select from the following options:

APCD Applicants Only

x Academic Researcher
Others (Single Use)
Others (Multiple Use)

Case Mix Applicants Only

Single Use
Limited Multiple Use
Multiple Use

Are you requesting a fee waiver?

x Yes
No

If yes, please submit a letter stating the basis for your request.

V. REQUESTED DATA ELEMENTS [APCD Only]

State and federal privacy laws limit the use of individually identifiable data to the minimum amount of data needed to accomplish a specific project objective. Please use the APCD Data Specification Workbook to identify which data elements you would like to request and attach this document to your application.

VI. MEDICAID DATA [APCD Only]

Please indicate here whether you are seeking Medicaid Data:

x Yes

No

Federal law (42 USC 1396a(a)7) restricts the use of individually identifiable data of Medicaid recipients to uses that are directly connected with the administration of the Medicaid program. If you are requesting Medicaid data from Level 2 or above, please describe in detail why your use of the data meets this requirement. Applications requesting Medicaid data will be forwarded to MassHealth for a determination as to whether the proposed use of the data is directly connected to the administration of the Medicaid program. MassHealth may impose additional requirements on applicants for Medicaid data as necessary to ensure compliance with federal laws and regulations regarding Medicaid.

This project will identify the impact of provider consolidation on costs and quality of care. It will also provide guidance on ways in which the design of insurance plans (e.g. the cost-control incentives given to physicians and the hospitals and providers included in the network) can affect referrals, costs and quality. The results will assist MassHealth in evaluating the potential benefits of introducing particular physician incentive schemes or expanding or reducing network size.

VII. MEDICARE DATA

Please indicate here whether you are seeking Medicare Data:

Yes

x No

Medicare data may only be disseminated to state agencies and/or entities conducting research projects that are directed and partially funded by the state if such research projects would allow for a Privacy Board or an IRB to make the findings listed at 45 CFR 164.512(i)(2)(ii) if the anticipated data recipient were to apply for the data from CMS directly. If you are requesting Medicare data, please explain how your research project is directed and partially funded by the state and describe in detail why your proposed project meets the criteria set forth in 45 CFR 164.512(i)(2)(ii). Applicants must describe how they will use the data and inform CHIA where the data will be housed. CHIA must be informed if the data has been physically moved, transmitted, or disclosed.

Applicants seeking Medicare data must complete a Medicare Request Form.

Applicants approved to receive Medicare data will be required to execute an Addendum to CHIA's standard data use agreement, containing terms and conditions required by CHIA's data use agreement with CMS.

N/A

VIII. DIRECT PATIENT IDENTIFIERS[3]

State and federal privacy laws may require the consent of Data Subjects prior to the release of any Direct Patient Identifiers. If you are requesting data that includes Direct Patient Identifiers, please provide documentation of patient consent or your basis for asserting that patient consent is not required.

IX. REQUESTS PURSUANT TO 957 CMR 5.04

Payers, providers, provider organizations and researchers seeking access to Level 1 (de-identified) data are required to describe how they will use such data for the purposes of lowering total medical expenses, coordinating care, benchmarking, quality analysis or other administrative research purposes. Please provide this information below.

N/A

X. FILTERS

If you are requesting APCD elements from Level 2 or above, describe any filters you are requesting to use in order to limit your request to the minimum set of records necessary to complete your project. (For example, you may only need individuals whose age is less than 21, claims for hospital services only, or only claims from small group projects.)

APCD FILE	DATA ELEMENT(S) FOR WHICH FILTERS ARE REQUESTED	RANGE OF VALUES REQUESTED
Medical Claims		
Pharmacy Claims		
Dental Claims		
Membership Eligibility		
Provider		
Product		

XI. PURPOSE AND INTENDED USE

1. Please explain why completing your project is in the public interest.

The high cost of health care is an important issue of concern for public policy. Recent policy changes (e.g. the introduction of Accountable Care Organizations) have created incentives for mergers of provider organizations. This project will identify ways in which ownership and contractual relationships between providers, in tandem with the incentives faced by physicians, affect costs, referrals and the quality of care. Our results will provide guidance for policy-makers on both the effect of different reimbursement methods and the antitrust issues raised by mergers between providers. They will also help identify the magnitude of cost savings likely to result from initiatives like Accountable Care Organizations, taking account of follow-on effects like provider mergers and changes in provider prices.

2. **Attach** a brief (1-2 pages) description of your research methodology. (This description will not be posted on the internet.)

[3] Direct Patient Identifiers. Personal information, such as name, social security number, and date of birth, that uniquely identifies an individual or that can be combined with other readily available information to uniquely identify an individual.

3. Has your project received approval from your organization's Institutional Review Board (IRB)?
   x Yes, and a copy of the approval letter is attached to this application.

   No, the IRB will review the project on _ _____.

   ☐ No, this project is not subject to IRB review.

   ☐ No, my organization does not have an IRB.

XII. APPLICANT QUALIFICATIONS

1. Describe your qualifications to perform the research described or accomplish the intended use of CHIA data.

Ariel Pakes is the Thomas Professor of Economics in the Department of Economics at Harvard University. Before coming to Harvard he was the Charles and Dorothea Dilley Professor of Economics at Yale University. He received his doctorate degree in economics from Harvard University. He is a research associate at the National Bureau of Economics and a fellow of both the American Academy of Arts and Sciences and the Econometric Society. He is pre-eminent in the field of Industrial Organization, having developed or co-developed many of the now-standard methodologies used to estimate models of demand and pricing in this field. We plan to use these methods in the second stage of our analysis.

Katherine Ho is an Associate Professor of Economics at Columbia University. She holds a PhD in Economics from Harvard University and a BA and MA in Mathematics from Cambridge University. She is also a research associate at the National Bureau of Economic Research. She specializes in the industrial organization of the medical care sector and has considerable experience working with confidential data in empirical health economic research.

Mark Shepard and Karen Stockley are economics doctoral students at Harvard University. Both have completed course work in health care economics and are pursuing research projects on insurance markets for their dissertations. Mark has experience analyzing claims data from the Massachusetts CommCare market for his dissertation project. Karen has co-authored numerous peer-reviewed articles and policy briefs analyzing changes in the health insurance coverage, access to health care, and health care use of Massachusetts residents following the 2006 health reform using sensitive survey data.

Pakes and Ho have co-authored a series of papers studying the hospital referral choices of physicians under different incentive schemes (2014 American Economic Review, 2012 International Journal of Industrial Organization, 2014 NBER Working Paper). All investigators have worked with sensitive health data in past projects.

2. Attach résumés or curriculum vitae of the applicant/principal investigator, key contributors, and of all individuals who will have access to the data. (These attachments will not be posted on the internet.)

XIII. DATA LINKAGE AND FURTHER DATA ABSTRACTION

1. Does your project require linking the CHIA Data to another dataset?

x Yes
No

2. If yes, will the CHIA Data be linked to other patient level data or with aggregate data (e.g. Census data)?
Patient Level Data
x Aggregate Data

3. If yes, please identify all linkages proposed and explain the reasons(s) that the linkage is necessary to accomplish the purpose of the project.

We propose to link the CHIA data to the following aggregate datasets:

1) Hospital linkages: we will link hospitals to the American Hospital Association Annual Survey Datase for hospital characteristics, and to the Medicare Hospital Compare dataset for quality and aggregate health outcome data. This is needed to describe provider characteristics and understand physicians' referral choices. The variable used for linkage will be the National Provider Identifier (NPI). For observations where the NPI is incomplete or invalid we will use the Entity Name to match hospitals.
2) Provider linkages: We will link providers to the SK&A office-based physician dataset to accurately link providers to practices and identify affiliations between practices and hospitals. To supplement the information on provider affiliations in the SK&A data, we will also merge on provider ownership relationships extracted from audited financial statements and public charities report forms filed with the Massachusetts Office of the Attorney General. We will not identify providers in published analyses and reports, nor will we report information that would make deductive disclosure possible. We will also link providers to the Americal Medical Association Physician Masterfile for additional information on provider specialty and demographic data. The variables used for linkage will be the NPI and, when the NPI is incomplete or invalid, the License ID. For observations where both the NPI and License ID are not available, we may use the following variables to implement probobalistic matching: Provider First, Middle, and Last name; Gender Code; Provider DOB (year only); Address, Primary Specialty Code, Provider Telephone.
3) Geographic area linkages: We will link member geographic data (ZIP, city, county) to the Area Resource File and the American Communities Survey/Census data to provide information on healthcare supply, socioeconomic status, and regional characteristics. This is needed to account for patient characteristics that might affect medical care utilization or choice of provider. We do not need to identify individual patients, merely to link characteristics of their ZIP code.
4) Employer linkages: We will link employers to publicly available information on employer health benefit offerings from the US Department of Labor's Form 5500 data sets. This will provide information on premiums, broker, and administrative fees for fully insured and self-insured employer health plans. These are needed for our models of insurance demand. We will not identify specific employers in published analyses and reports, nor will we report information that would make deductive disclosure possible. The variable used for linkage will be the Employer EIN.

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

Since we will link the data only to aggregate datasets, which do not increase our likelihood or ability to

identify individual patients, the linkages will not jeopardize patient confidentiality. As discussed in the data security and integrity section, we will take extensive steps to ensure the confidentiality of the data.

XIV. PUBLICATION / DISSEMINATION / RE-RELEASE

1. Describe your plans to publish or otherwise disclose CHIA Data, or any data derived or extracted from such data, in any paper, report, website, statistical tabulation, seminar, conference, or other setting.

We propose to disseminate our work through published peer-reviewed economics journals, working papers, and conferences. We are participants in the National Bureau for Economic Research's industrial organization program and healthcare program, both of which hold conferences that are widely attended by other researchers and policy-makers. The results will include summary statistics and analyses completed using the data; however we will aggregate all statistics to groups (including at least 10 patients) so that no identification of patients will be possible from our published results.

2. Will the results of your analysis be publicly available to any interested party? Please describe how an interested party will obtain your analysis and, if applicable, the amount of the fee.

The results will be available for no fee at the researchers' websites.

3. Will you use the data for consulting purposes?
   Yes
x No

4. Will you be selling standard report products using the data?
   Yes
x No

5. Will you be selling a software product using the data?
   Yes
x No

6. If you have answered "yes" to questions 3, 4 or 5, please describe the types of products, services or studies.

N/A

XV. USE OF AGENTS AND/OR CONTRACTORS

Third-Party Vendors. Provide the following information for all agents and contractors who will work with the CHIA Data.

Company Name:	N/A
Contact Person:	

Title:	
Address:	
Telephone Number:	
E-mail Address:	
Organization Website:	

7. Will the agent/contractor have access to the data at a location other than your location or in an off-site server and/or database?

   Yes
   No

8. Describe the tasks and products assigned to this agent or contractor for this project.

9. Describe the qualifications of this agent or contractor to perform such tasks or deliver such products.

10. Describe your oversight and monitoring of the activity and actions of this agent or subcontractor.
