

**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:	Kenneth D. Mandl
Title:	Professor
Organization:	Boston Children's Hospital and Harvard Medical School
Project Title:	Network analysis of providers and shared patients
Date of Application:	5/16/2014
Project Objectives (240 character limit)	The goal of this project is to examine how health outcomes and costs are related to measures thought to reflect stability and communication amongst the team of providers that provide care to a patient. Patients often receive healthcare from multiple providers whom we refer to as a team. This team shares one patient by definition but can share more. Providers who share larger portions of their full patient panels may communicate more with each other and better coordinate care for their patients, leading to better health outcomes and lower costs. Providers, as nodes in a network, are linked by the number of patients shared. Using social network methods, standard measures (e.g. size, density) will be created for each team. Relations between team variables and outcomes such as proper medication use, lower hospital readmission rates, and lower overall costs will be studied. The APCD is unique in that it captures a major portion of providers' practices, thus enabling the construction of a more accurate provider network than has been studied to date.
Project Research Questions (if applicable)	Do patients with provider teams characterized by greater stability and care coordination: <ol style="list-style-type: none"> 1. Have better health outcomes (e.g. appropriate asthma and diabetes care, lower hospital readmission rates)? 2. Have higher rates of proper medication utilization (e.g. no overlapping fills for interacting drugs)? 3. Have lower healthcare costs, either overall or by category (inpatient, outpatient, pharmacy)?

I. PROJECT SUMMARY

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

The goal of this project is to examine how health outcomes and costs are related to measures thought to reflect stability and communication amongst the team of providers that care for a patient. We hypothesize that better outcomes and lower costs will occur when teams communicate more and

coordinate patient care. Our measures of stability and communication are indirectly observed and derived using social network analytics. The APCD offers a unique opportunity to capture the majority of providers' panels across payers. Therefore, we can construct more reliable and valid measures for teams than we have been able to do in the past with other data such as claims from a single insurer.

Part of our efforts will go towards constructing a "social network" of providers, which requires that we be able to follow unique providers throughout the data. Therefore, a provider Enterprise ID (EID) will be critical. Using data from a single insurer, we have already developed methods to build these networks and create variables. Individual providers are nodes and nodes are linked when they share patients. Links between pairs of providers are weighted by the number of patients shared. Therefore, a patient EID is also critical.

Our use of social networks differs somewhat from other researchers with different goals. Rather than studying the large overall network of providers itself, we examine the many small sub-networks of providers who surround individual patients. These are the "teams" of providers upon whom we focus. Although we do not study the large network, we use it to create variables related to the extent to which individual providers are linked to other providers.

We plan to use most of the large number of variables in the APCD. All medical and pharmacy claims are required to construct the provider social network as well as health outcome and cost variables. The provider and patient EIDs are needed to follow individuals across payers. For each provider, a patient panel will be identified by including all unique patients from whom medical claims were filled. For each patient, a team will be identified by including all unique providers who appear on their medical claims.

Dates are needed to correctly sequence records over time and to construct measures such as length of stay or time between events. Date of birth will be used to calculate age. Diagnosis and procedure codes will be used to construct health covariate and outcome variables. Drug codes and fill dates will be used to construct variables related to medication utilization. Cost variables will be used to create summary cost variables as outcomes. Provider characteristics such as specialty or practice size will be used as covariates. Patient demographics (age, gender) will be used as covariates or for sub-group analyses. Regression analysis, controlling for covariates, will be the primary analytic technique. Statistical interactions between predictor variables will be examined, and stratification by groups or post-hoc comparisons will be performed as appropriate.

It is important to note that while our study requires the provider EID that is not yet available, much work related to selecting patients and assessing health status can be performed before this id is ready. In preliminary work with a claims database from a single insurer, we have already created some health outcome variables. Examples include appropriate medication utilization (e.g. no overlapping fills for interacting drugs), recommended care for asthma or diabetes, and hospital readmission rates after heart failure. We will add to this list of outcomes and also include cost variables. Therefore, our cohorts for outcome analyses will be identified, and our covariate and outcome measures refined by the time the provider EID is ready for us to complete this project.

II. 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
------------------------------	--	---------------------------	---

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.

		2009 - 2012	
<input checked="" type="checkbox"/> Medical Claims	<input type="checkbox"/> Level 1 ³ <input checked="" type="checkbox"/> Level 2	Select...	<input checked="" type="checkbox"/> 2009 <input checked="" type="checkbox"/> 2010 <input checked="" type="checkbox"/> 2011 <input checked="" type="checkbox"/> 2012
<input checked="" type="checkbox"/> Pharmacy Claims	<input checked="" type="checkbox"/> Level 2	Select...	<input checked="" type="checkbox"/> 2009 <input checked="" type="checkbox"/> 2010 <input checked="" type="checkbox"/> 2011 <input checked="" type="checkbox"/> 2012
<input type="checkbox"/> Dental Claims <input checked="" type="checkbox"/> Member Eligibility <input checked="" type="checkbox"/> Provider <input checked="" type="checkbox"/> Product	<input type="checkbox"/> Level 2 <input checked="" type="checkbox"/> Level 2 <input checked="" type="checkbox"/> Level 2 <input checked="" type="checkbox"/> Level 2	Multiple Select... Select... Select...	<input checked="" type="checkbox"/> 2009 <input checked="" type="checkbox"/> 2010 <input checked="" type="checkbox"/> 2011 <input checked="" type="checkbox"/> 2012

CASEMIX	Level 1 - 6	Fiscal Years Requested
Inpatient Discharge	<input type="checkbox"/> Level 1 – No Identifiable Data Elements <input type="checkbox"/> Level 2 – Unique Physician Number (UPN) <input type="checkbox"/> Level 3 – Unique Health Information Number (UHIN) <input type="checkbox"/> Level 4 – UHIN and UPN <input type="checkbox"/> Level 5 – Date(s) of Admission; Discharge; Significant Procedures <input type="checkbox"/> Level 6 – Date of Birth; Medical Record Number; Billing Number	1998-2013 Available (limited data 1989-1997)
Outpatient Observation	<input type="checkbox"/> Level 1 – No Identifiable Data Elements <input type="checkbox"/> Level 2 – Unique Physician Number (UPN) <input type="checkbox"/> Level 3 – Unique Health Information Number (UHIN) <input type="checkbox"/> Level 4 – UHIN and UPN <input type="checkbox"/> Level 5 – Date(s) of Admission; Discharge; Significant Procedures <input type="checkbox"/> Level 6 – Date of Birth; Medical Record Number; Billing Number	2002-2012 Available (2013 available 8/1/14)
Emergency Department	<input type="checkbox"/> Level 1 – No Identifiable Data Elements <input type="checkbox"/> Level 2 – Unique Physician Number (UPN) <input type="checkbox"/> Level 3 – Unique Health Information Number (UHIN) <input type="checkbox"/> Level 4 – UHIN and UPN; Stated Reason for Visit <input type="checkbox"/> Level 5 – Date(s) of Admission; Discharge; Significant Procedures <input type="checkbox"/> Level 6 – Date of Birth; Medical Record Number; Billing Number	2000-2012 Available (2013 available 9/1/14)

³ Please note that Level 1 APCD data is not available as of 4/30/2014. This is scheduled to be available later in 2014.

III. 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

- 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?

- Yes
- No

If yes, please submit a letter stating the basis for your request. Please refer to the fee schedule for qualifications for receiving a fee waiver. If you are requesting a waiver based on the financial hardship provision, please provide documentation of your financial situation. Please note that non-profit status alone isn't sufficient to qualify for a fee waiver.

IV. 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.

V. MEDICAID DATA [APCD Only]

Please indicate here whether you are seeking Medicaid Data:

- 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.

Variables that affect health outcomes and costs are of direct relevance to the administration of the Medicaid program. Understanding how provider team characteristics impact care can lead to

recommendations for care redesign, which is a major emphasis of the Affordable Care Act. Because Medicaid covers a distinct population, including these individuals in the population studied contributes to the accuracy of the measures we construct, and to the reliability and validity of results we report. Of particular importance is the pediatric population. Children are often underrepresented in medical studies, and including those insured by Medicaid, as well as those with private insurance, makes any findings more generalizable and noteworthy.

VI. REQUESTS PURSUANT TO 957 CMR 5.04

If you are a payer, provider, provider organization or researcher seeking access to Level 1 (de-identified) data, please describe how you 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.

VII. 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		

VIII. PURPOSE AND INTENDED USE

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

A major emphasis of the Affordable Care Act is care redesign. Healthcare processes have been traditionally shaped as if care is delivered by individual providers working in isolation. Yet for many patients, especially those with complex conditions, care is instead delivered by teams. A recent Institute of Medicine report concluded that system redesign around team-based care may achieve improved health and care at a sustainable cost. This study will investigate teams as they currently exist and document situations where they do achieve the goal of improving health outcomes and lowering or sustaining costs.

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

3. Has your project received approval from your organization's Institutional Review Board (IRB)? Please note that CHIA will not review your application until IRB documentation has been received (if applicable).

- 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.

IX. APPLICANT QUALIFICATIONS

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

Kenneth D Mandl, MD, MPH is Chair in Biomedical Informatics and Population Health at Boston Children's Hospital and Professor at Harvard Medical School. He runs a large bioinformatics research group, and is PI of an NIH grant to study provider interactions using network analysis.

Karen L Olson, PhD is a research faculty member and statistical analyst at Boston Children's Hospital and Harvard Medical School with extensive experience analyzing large healthcare datasets, including electronic medical records, medical and pharmacy claims from national insurers, and the MA Casemix data. She has expertise on multivariate and network analysis, and has developed methods for this study using other data.

Aurel Cami, PhD is a research faculty member at Boston Children's Hospital and Harvard Medical School. His work focuses on developing mathematical models to address pressing challenges in public health and has developed methods using network analysis to predict unknown drug-drug interactions.

Mei-Sing Ong, PhD is Research Fellow at Boston Children's Hospital with extensive experience in data mining analytics. Her research focuses on the development of computational methodologies to leverage large clinical datasets for knowledge discovery.

2. Attach résumés or curricula 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.)

X. DATA LINKAGE AND FURTHER DATA ABSTRACTION

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

- 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
- 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. Please be specific in describing vvhich data elements will be linked to outside datasets and how this will be accomplished.

Zip codes can be linked to US Census zip code tabulation areas (ZCTA). Variables describing the ZCTA such as median household income or race/ethnicity percentages can be used as proxy variables for demographic variables that are either unavailable (e.g. income) in insurance claims data or incomplete (e.g. race/ethnicity). In some analyses, it can be appropriate to use these variables as covariates.

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

We plan to only use aggregate US Census data which cannot be linked to individuals. The Census does not report on very small group sizes so as to prevent identification of individuals.

XI. 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.

Results of our analyses will be published in academic medical journals and presented at professional meetings. Data will be presented only as summary statistics.

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.

To the extent possible, papers will be published in open-source journals. Publications that result from funding from the National Institutes of Health will be made publically available via a link on the PubMed web site (www.ncbi.nlm.nih.gov/pubmed). Reprint requests for academic publications or presentations will be honored.

3. Will you use the data for consulting purposes?

Yes
 No

4. Will you be selling standard report products using the data?

Yes
 No

5. Will you be selling a software product using the data?

Yes
 No

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