

**Commonwealth of Massachusetts
Center for Health Information & Analysis (CHIA)
Non-Government APCD Request for Data**

This form 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 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
Mailing Address:	Division of Emergency Medicine Boston Children's Hospital 300 Longwood, Mailstop BCH 3187 Boston, MA 02115
Telephone Number:	617-355-4145
Email Address:	Kenneth_mandl@harvard.edu
Names of Co-Investigators:	Karen L Olson, Mei-Sing Ong, Alon Geva
Email Addresses of Co-Investigators:	Karen.olson@childrens.harvard.edu Mei-sing.ong@childrens.harvard.edu Alon.geva@childrens.harvard.edu
Original Data Request Submission Date:	4-10-2015 5-16-2014
Dates Data Request Revised:	10-20-2015
Project Objectives (240 character limit)	To examine how health outcomes/costs are related to network-derived measures of coordination for the team of providers that care for a patient. APCD captures a major portion of providers' practices, enabling us to build 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 have: <ol style="list-style-type: none"> 1. Better health outcomes (e.g. appropriate asthma and diabetes care, lower hospital readmission rates)? 2. Higher rates of proper medication utilization (e.g. no overlapping fills for interacting drugs)? 3. Lower healthcare costs, either overall or by category (inpatient, outpatient, pharmacy)?

II. 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 master provider ID would be helpful. However, we can use other provider ids that exist in the data to construct our own master id until one is available from CHIA. Using data from a single insurer, we have already developed methods to build provider 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 master patient ID is critical to accurately count the number of unique patients in providers' panels.

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 IDs 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. Month and year of birth will be used to calculate approximate 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 would benefit from a master provider ID that is not yet available in the APCD, much can be accomplished without it. To the extent possible, we plan to use the National Provider Identifier (NPI) as a master id. For claims that do not include the NPI, we will attempt to complete that missing variable with a correct NPI derived from the provider file or claims that do include it.

III. FILES REQUESTED

Please indicate the databases from which you seek data, and the year(s) of data requested.

ALL PAYER CLAIMS DATABASE	Single or Multiple Use	Year(s) Of Data Requested Current Yrs. Available 2009 – 2013
<input checked="" type="checkbox"/> Medical Claims	Single	<input checked="" type="checkbox"/> 2009 <input checked="" type="checkbox"/> 2010 <input checked="" type="checkbox"/> 2011 <input checked="" type="checkbox"/> 2012 <input checked="" type="checkbox"/> 2013
<input checked="" type="checkbox"/> Pharmacy Claims	Single	<input checked="" type="checkbox"/> 2009 <input checked="" type="checkbox"/> 2010 <input checked="" type="checkbox"/> 2011 <input checked="" type="checkbox"/> 2012 <input checked="" type="checkbox"/> 2013
<input type="checkbox"/> Dental Claims	n/a	2009 <input type="checkbox"/> 2010 <input type="checkbox"/> 2011 <input type="checkbox"/> 2012 <input type="checkbox"/> 2013 <input type="checkbox"/>
<input checked="" type="checkbox"/> Member Eligibility	Single	<input checked="" type="checkbox"/> 2009 <input checked="" type="checkbox"/> 2010 <input checked="" type="checkbox"/> 2011 <input checked="" type="checkbox"/> 2012 <input checked="" type="checkbox"/> 2013
<input checked="" type="checkbox"/> Provider	Single	<input checked="" type="checkbox"/> 2009 <input checked="" type="checkbox"/> 2010 <input checked="" type="checkbox"/> 2011 <input checked="" type="checkbox"/> 2012 <input checked="" type="checkbox"/> 2013
<input checked="" type="checkbox"/> Product	Single	<input checked="" type="checkbox"/> 2009 <input checked="" type="checkbox"/> 2010 <input checked="" type="checkbox"/> 2011 <input checked="" type="checkbox"/> 2012 <input checked="" type="checkbox"/> 2013

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. FEE INFORMATION

Please consult the fee schedules for APCD data) and Case Mix data, available at http://chiamass.gov/regulations/#957_5, and select from the following options:

APCD Applicants Only

- Academic Researcher
- Others (Single Use)
- Others (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.

VI. 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. Including those insured by Medicaid, as well as those with private insurance, makes any findings more generalizable and noteworthy.

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	No filters	
Pharmacy Claims	No filters	
Dental Claims		
Membership Eligibility	No filters	
Provider	No filters	
Product	No filters	

IX. 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.

X. 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 the Principal Investigator for a 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.

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.

Alon Geva, MD, MPH is a Fellow at Boston Children’s Hospital and Harvard Medical School, whose interests are critical care medicine and biomedical informatics; specifically, how “big data,” modeling, and other computational methods might be used to analyze complex systems and improve the quality of care delivered to critically ill infants and children.

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.)

XI. DATA LINKAGE AND FURTHER DATA ABSTRACTION

Note: Data linkage involves combining CHIA data with other databases to create one extensive database for analysis. Data linkage is typically used to link multiple events or characteristics that refer to a single person in CHIA data within one database.

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

- Yes
- No linkage or merger with any other database will occur

2. If yes, will the CHIA Data be linked or merged to other individual patient level data (e.g. disease registries, death data), individual provider level data (e.g., American Medical Association Physician Masterfile) , facility level (e.g., American Hospital Association data) or with aggregate data (e.g., Census data)? [check all that apply]

Individual Patient Level Data

What is the purpose of the linkage:

What databases are involved, who owns the data and which specific data elements will be used for linkage:

Individual Provider Level Data

What is the purpose of the linkage:

To the extent possible, the NPI will be used as a master id for unique individual providers who are people (not facilities). We hope to establish a cross-reference table that assigns a single NPI to APCD plan provider ids (e.g. PV002). Similar variables (e.g. demographics, specialty, practice location) from the APCD and NPPES Registry will be compared to evaluate the extent of our success at creating this cross-reference table.

What databases are involved, who owns the data and which specific data elements will be used for linkage:

We plan to link the provider NPI (national provider identifier) to the publically available NPPES NPI Registry, which is owned by the Centers for Medicare and Medicaid Services.

Individual Facility Level Data

What is the purpose of the linkage:

To the extent possible, we plan to use the NPI as a master id for unique providers that are facilities. To evaluate our success at this, we will compare similar variables (e.g. city, zip) from the APCD and NPPES Registry. Also, when specific facilities can be associated with individual providers who are people, characteristics of the facility (e.g. location) can be used as additional elements to evaluate the extent to which we are successful at creating a NPI—APCD provider id cross-reference table for unique individual providers.

What databases are involved, who owns the data and which specific data elements will be used for linkage:

We plan to link the facility NPI (national provider identifier) to the publically available NPPES NPI Registry, which is owned by the Centers for Medicare and Medicaid Services.

Aggregate Data

What is the purpose of the linkage:

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.

What databases are involved, who owns the data and which specific data elements will be used for linkage:

United States Census Bureau owns the data and makes it publically available. Zip code will be used as the link.

3. 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 .

NPIs from the APCD will be directly matched to NPIs in the NPPES Registry. 5-digit zip codes from the APCD will be directly matched to 5-digit zip code tabulation areas in the census data.

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 use aggregate US Census data which cannot be linked to individuals. The Census Bureau does not report on very small group sizes so as to prevent identification of individuals. The publically available NPPES Registry is not a patient registry - it is for healthcare providers, both people and facilities.

5. If yes, and the data mentioned above is not in the public domain, please attach a letter of agreement or other appropriate documentation on restrictions of use from the data owner corroborating that they agree to have you initiate linkage of their data with CHIA data and include the data owner's website.

XII. 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 (<http://www.ncbi.nlm.nih.gov/pubmed>). Reprint request 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. Will you be reselling the data?

Yes
 No

If yes, in what format will you be reselling the data (e.g., as a standalone product, incorporated with a software product, with a subscription, etc.)?

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

XIII. 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:	NONE
Contact Person:	
Title:	
Address:	
Telephone Number:	
E-mail Address:	
Organization Website:	

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

- Yes
- No

If yes, please provide information about the agent/contractor’s data management practices, policies and procedures in your Data Management Plan.

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

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

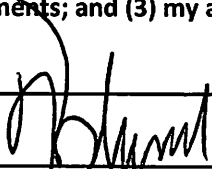
11. Describe your oversight and monitoring of the activity and actions of this agent or subcontractor.

XIV. ASSURANCES

Applicants requesting and receiving data from CHIA pursuant to 957 CMR 5.00 (“Data Recipients”) will be provided with data following the execution of a data use agreement that requires the Data Recipient to adhere to processes and procedures aimed at preventing unauthorized access, disclosure or use of data, as detailed in the DUA and the applicant’s CHIA-approved Data Management Plan.

Data Recipients are further subject to the requirements and restrictions contained in applicable state and federal laws protecting privacy and data security, and will be required to adopt and implement policies and procedures designed to protect CHIA data in a manner consistent with the federal Health Insurance Portability and Accountability Act of 1996 (HIPAA).

By my signature below, I attest to: (1) the accuracy of the information provided herein; (2) my organization’s ability to meet CHIA’s minimum data security requirements; and (3) my authority to bind the organization seeking CHIA data for the purposes described herein.

Signature:	
Printed Name:	Kenneth D. Mandl
Title	Professor
Original Data Request Submission Date:	4-10-2015 5-16-2014
Dates Data Request Revised:	10-20-2015