Thursday, August 10, 2023

MARYLAND Health Care Commission

Application for Standard Analytic Files from the Maryland Medical Care Data Base (Non-Governmental Applicants)

INSTRUCTIONS

Non-Governmental Applicants may request MD APCD Standard Analytics Files by submitting this completed Application, including attachments and the Data Management Plan. MHCC will review the application package to determine whether the request meets the criteria for data release pursuant to COMAR 10.25.05. Review data availability here then calculate the applicable fees here. Review Important notes:

- Incomplete applications will be returned to the Applicant and the request may be delayed.
- All applications require a non-refundable application fee, payable at the time of submission.
- All application attachments will be incorporated into the Data Use Agreement (DUA) that must be signed prior to any MCDB data being transmitted. A draft DUA will be provided to the applicant after this Application is submitted, so that the Applicant can review the terms and conditions.
- COMAR 10.25.05.07A requires that all completed applications be published on the Commission's website while the application is under review, without the data management plan and security measures.
- Requests that include Maryland Medicaid Managed Care data and Medicare Fee for Service data require special consideration that may increase the review timeline.

Data Fee Calculator available to estimate the fee for your data sets. The Data Fee Waiver is available to support those who are unable to access the data for financial hardship. If completing a Data Fee Waiver, please attach it to this application under Attachment H.
This application should only be completed and submitted for Standard Analytic Files. All requests for Custom Data Files should be sent directly to MHCC at mhcc.datarelease@maryland.gov.

List of Required Forms

ATTACHMENT A: PROJECT SCOPE
ATTACHMENT B: MD APCD DATASET REQUESTED
ATTACHMENT C: ADDITIONAL DATA SOURCES AND LINKAGE
ATTACHMENT D: DATA MANAGEMENT PLAN
ATTACHMENT E: USE OF CONTRACTORS AND/OR CONSULTANTS (External Entities)
ATTACHMENT F: APPLICANT QUALIFICATIONS
ATTACHMENT G: ATTESTATION
ATTACHMENT H: INSTITUTIONAL REVIEW BOARD AND DATA FEE DOCUMENTS

PROJECT INFORMATION

Project Title	Diabetic Ketoacidosis Trends and Resource Utilization at Diagnosis of Type 1 Diabetes in the United States
Scheduled Project Start Date	Present
Scheduled Project End Date	June 30, 2024

Project Overview

Diabetic ketoacidosis (DKA) at onset of type 1 diabetes (T1D) is life-threatening and expensive to treat. DKA events at diagnosis have increased to an alarming 60% among all children diagnosed with T1D in Colorado over the past decade. There is emerging evidence on the clinical benefits of pre-symptomatic T1D screening and education programs. However, two major obstacles remain for adoption of screening programs around the United States: first, the current and best available evidence around DKA rates at diagnosis of T1D is limited single tertiary centers. Second, the evidence gaps around population level DKA rates and DKA management costs before and after events limit the external validity of cost-effectiveness analyses on T1D screening, ultimately limiting the uptake of screening programs. We propose to fill these gaps with real-world evidence from an EMR registry and all-payer claims databases (APCDs) from around the United States. Our goal is to validate the use of claims data to identify DKA at diagnosis and estimate trends in DKA events and resource utilization at T1D diagnosis across all available private and public insurers in multiple states. We plan to address this goal by first developing an algorithm to measure DKA events at onset of T1D. Second, we plan to estimate trends and variation in DKA at diagnosis among patients with T1D across 5 additional states beyond Colorado. Finally, we will estimate the resources and expenditures used surrounding a DKA event. Achieving our goal will provide public health decision makers accurate estimates of the scale of DKA events at T1D onset across the country. each aim provides unique information for researchers and public health decision makers and will inform efficient allocation of screening, monitoring, and educational awareness programs. Our study will also inform future research around predicting DKA events before diagnosis and inputs to identify cost-effective screening strategies.

Applicant

(Agency, Academic Institution, Research Organization, Company, Individual, etc.)

Individual/Organization Name Robert McQueen

Website https://pharmacy.cuanschutz.edu/

Email Address robert.mcqueen@cuanschutz.edu

Telephone Number 3037099264
Phone Number

Mailing Address 12850 E Montview Blvd
Street Address

City/Town Aurora State CO Zip Code 80045
City **State** **Zip code**

Principal Investigator/Project Manager

(Individual responsible for the research team using the data)

Name Robert McQueen
First Name **Last Name**

Title Associate Professor
Type a label

Organization Name University of Colorado
Type a label

E-mail Address robert.mcqueen@cuanschutz.edu
Email

Telephone Number 303 7099264
Area Code **Phone Number**

Mailing Address 12850 E Montview Blvd
Street Address

City/Town Aurora State CO Zip Code 80045
City **State** **Zip**

Data Custodian
(person responsible for receiving, organizing, storing, and archiving data)

Name Eric Gutierrez
First Name **Last Name**

Title Professional Research Assistant

Organization/Company (if different from Applicant)

E-mail Address ERIC.GUTIERREZ@CUANSCHUTZ.EDU
Email

Telephone Number

Mailing Address

City/Town State Zip Code

Relationship to Applicant (e.g., Contractor)

Project Contact
(person responsible for all communications with MHCC)

Name Robert McQueen
First Name **Last Name**

Title Associate Professor
Type a label

Organization Name University of Colorado
Name

E-mail Address robert.mcqueen@cuanschutz.edu
Email

Telephone Number 3
Area Code

Mailing Address

City/Town State

Zip Code

ATTACHMENT A: PROJECT SCOPE
Project Purpose

a. Describe the specific research question(s) you are trying to answer or problem(s) you are trying to solve with the requested data or describe the intended product or report that will be derived from the requested data. If a research project, please list each individual question or

aim of the analysis.

Our goal is to estimate trends in diabetic ketoacidosis (DKA) events and resource utilization at type 1 diabetes (T1D) diagnosis across all available private and public insurers in multiple states, including Maryland.

b. Briefly describe the purpose(s) for which MD APCD data is sought. Use quantitative indicators of public health importance where possible. For example: variation in costs of care; rates of under or over service utilization; health system performance measures; the effect of public health initiatives, health insurance, etc.

Specific Aim 2: Estimate trends and variation in DKA at diagnosis among people with type 1 diabetes in Maryland. Understanding demographic and geographic differences in DKA event rates is important for public health decision makers.

Specific Aim 3: Estimate resource utilization for people with T1D with a DKA event as compared to people diagnosed with T1D without a DKA event. With sufficient time pre- and post-diagnosis, we will estimate resource utilization before DKA diagnoses to inform future research attempting to predict T1D in advance of a diagnosis using claims data.

c. Explain in detail how the planned project that will use MD APCD data is in the public interest and give specific examples of how the project will serve the public interest.

Our study will provide public health decision makers accurate estimates of the scale of DKA events at T1D onset across the country, with specific local estimates in Maryland. Further, each aim provides unique information for researchers and public health decision makers and will inform efficient allocation of screening, monitoring, and educational awareness programs. Our study will also inform future research around predicting DKA events before diagnosis and inputs to identify cost-effective screening strategies. More specifically, results will inform the incidence of DKA events at diagnosis in children for providers and public health decision makers in Arkansas. This information, paired with potential screening programs (Helmsely Charitable Trust is looking to establish national screening programs), can help reduce the incidence of DKA events. For example, through educating parents and children about symptoms and monitoring children through routine practice, the Autoimmunity Screening for Kids (ASK) program in Colorado reduced incidence of DKA events from 60% to less than 5% (American Diabetes Association 81st Scientific Sessions 2021 virtual; Cost-effectiveness of Type 1 Diabetes Risk Screening – materials provided upon request). This benefit can be applied to Maryland to improve the health of those at risk of T1D. These benefits include reducing DKA events at onset, improved glycemic control after diagnosis, and improved health and well-being of family units. The results will be available to Arkansans through open access journal articles in journals such as Diabetes Care. The study team will also be able to provide direct reporting on data from Maryland to APCD staff for publication on their website. We are open to further suggestions to make sure all Maryland residents have access to the evidence we find in our study.

d. Explain why the planned project could not be practicably conducted without access to and use of protected health information.

We will be requesting a limited data set to identify dates of diagnosis and corresponding pre- and post-diagnosis resource utilization. Without the use of dates (e.g., hospitalization date) to set index dates, we cannot make any comparisons to pre- and post-diagnosis. Therefore we will need a limited data set to answer the question.

e. Explain why the planned project could not be practicably conducted without waiving any individual authorization required by 45 CFR § 164.508.

Given the study relies on past data, there is no possible way to contact all participants to ask for consent. The study relies on retrospective data. Further, we are not asking for anything beyond a limited data set.

Project Methodology

a. Describe the project methodology, including project objectives, relevant study questions, analysis methods, software, groupers, and other analytical tools

Specific Aim 2: Estimate trends and variation in DKA at diagnosis among people with T1D in five additional APCDs from different regions in the United States. Understanding demographic and geographic differences in DKA event rates is important for public health decision makers. Based on feasibility and population characteristics, we propose to include Arkansas, Colorado, Massachusetts, Maryland, Oregon, and Utah.

Specific Aim 3: Estimate resource utilization for people with T1D with a DKA event as compared to people diagnosed with T1D without a DKA event. With sufficient time pre- and post-diagnosis, we will estimate resource utilization before DKA diagnoses to inform future research attempting to predict T1D in advance of a diagnosis using claims data.

Achieving our goal will provide public health decision makers accurate estimates of the scale of DKA events at T1D onset across the country. Further, each aim provides unique information for researchers and public health decision makers and will inform efficient allocation of screening, monitoring, and educational awareness programs. Our study will also inform future research around predicting DKA events before diagnosis and inputs to identify cost-effective screening strategies.

Research Design and Methods

Model construction and analyses
The primary outcome in aim 2 will present a descriptive measure of the proportion of children with DKA by year and by other characteristics such as insurance status. We will also identify factors associated with DKA events using logistic regression analyses. Specifically, we will test the univariate and multivariable associations between demographic and clinical characteristics such as insurance status, age, race/ethnicity, year of diagnosis, and rural or urban location. We plan to further characterize location at the neighborhood level by linking plan member zip code to the Area Deprivation Index (ADI) which has been associated with increased utilization of health services and even early death. Odd ratios with 95% confidence intervals will be reported for the multivariable analysis.

To ensure our population is representative of the United States, we will create a set of observable demographic and clinical characteristics from publicly available sources in the United States. For each individual in our sample, we will then calculate the probability that each subject was drawn from our observed national sample based on those observed characteristics. Finally, we will use these weights to scale our individual level data from 5 states to a national level. For example, Colorado has a lower non-white population than the broader United States. This method will allow us to give greater weight to non-white individuals in Colorado so they represent a larger proportion of our sample.

Study population
The study population will be the same as in aim 2, however outcomes will be focused on resource utilization surrounding T1D diagnoses.

Model construction and analyses
The outcomes from aim 3 include differences in diabetes-related and all-cause annualized number of outpatient, ED, inpatient, and pharmacy fills between patients with and without DKA events at diagnosis for + or – 6 months surrounding a diagnosis. Separately we will include expenditures to estimate the paid claims for these resources used.

The arithmetic mean all-cause and diabetes-related healthcare resource use and expenditures per patient will be compared using two methods: (1) a parametric t-test with unequal standard deviations; and (2) bootstrapped-t with 1000 samples taken with replacement from the dataset. Multivariable analyses on utilization and expenditure outcomes will use generalized linear modeling (GLM) with cluster robust standard errors to estimate and compare utilization and expenditures between patients with and without a DKA event at diagnosis. For utilization outcomes, a log link with negative binomial family will be appropriate since the data will be coded as counts. For adjusted all-cause and diabetes-related healthcare expenditures, we will use generalized linear models with a Gamma distribution and log link, controlling for potential confounders at baseline including health care resource utilization. However, we will use the modified Park test to test these assumptions. Differences in adjusted mean costs will be reported with

95% confidence intervals developed from non-parametric bootstrapping methods with 1000 random samples taken with replacement from the dataset. All expenditure related variables will be reported in 2022 US dollars. Analyses will be performed using Stata 15.1.

b. If required by your funding source or home institution to obtain Institutional Review Board review for your project, provide the information regarding the IRB approval below and attach a copy of the current IRB approval on tab "Attachment H".

IRB Approval End Date

06-30-2025

Date

IRB Name and Location

Colorado Multiple Institutional Review Board, Aurora, Colorado

blanks

Publication and Dissemination

a. Do you anticipate that the results of your analysis will be published or made publicly available? If yes, how do you intend to disseminate the results of the study (e.g., publication in a professional journal, poster presentation, newsletter, web page, seminar, conference, statistical tabulation, etc.)?

Yes, published through journals such as Diabetes Care and presented at the American Diabetes Association Scientific Sessions.

b. All public displays of MD APCD data, regardless of the medium, must comply with MD APCD's cell size suppression policy, as set forth in the Data Use Agreement. Describe how you will ensure that any public display will suppress every cell containing less than 11 observations and suppress percentages or other mathematical formulas that result in the display of every cell with less than 11 observations.

We denote cells < 30 as "CR" or cannot report any frequencies, percentages, or other output.

c. Identify the lowest geographical level of analysis of data you will present for publication or presentation (e.g., state level, city/town level, zip code level, etc.). Will maps be presented? What methods will be used to ensure that individuals cannot be re-identified in this publication or presentation?

zip code. No maps will be presented. No other information below the zip code level will be presented.

If you answer "yes" to any of the following questions, describe the types of products, software, services, or tools and what the corresponding fees will be for such products, software, services, or tools.

a. Will the MD APCD data be used for consulting purposes?

No

b. Will report(s), website(s), or statistical tabulation(s) using MD APCD data be shared or sold?

No

c. Will a software product using MD APCD data be shared or sold?

No

d. Will MD APCD data be used as input to develop a product (i.e., severity index tool, a risk adjustment tool, a reference tool, etc.)?

No

e. Will MD APCD data be sold or shared in any format not noted above? If yes, in what format and who are the purchasers of the data?

No

f. Will the project result in disclosing MD APCD data, or any data derived or extracted from such data, in any paper, report, website, a statistical tabulation, seminar, or another setting that is not disseminated to the public?

No

g. Will the results from the project be used for the purpose of price transparency?

No

h. Will health care providers be individually identified? If yes, describe your protocol for informing health care providers prior to publication of this data/report.

No

Funding Sources

a. What is the source of funding that supports this project? Provide detailed information about potential and actual sources.

Leona M. And Harry B. Helmsley Charitable Trust funded this project to the University of Colorado.

b. Describe any data sharing or other requirements imposed by the above funding sources as a condition for receipt of funding?

The funder is not involved in the study design, analysis, or dissemination of the research. There are no conditions imposed by the funder.

c. Please upload documentation that includes grant number, a budget, and any information that shows the available funding for this project.

June Helmsley_Grant_Agree... .pdf

Data Security

a. Explain how your use of data will involve no more than a minimal risk to the privacy of individuals. As part of your response, please address how you will protect the data from improper use or disclosure and assure that the data will not be reused or disclosed to any other person or entity, except as required by law, for authorized oversight of the research for which the data was requested, or for other research for which the use or disclosure of PHI would be permitted under 45 CFR 164.512(i)(2)(ii).

We will use a limited data set with only dates relevant to the research question. We will include no other information to prevent any risk to persons in the data. data will be stored within OIT Data Center on the Anschutz Medical Campus in the Dell EMC Isilon clustered storage system, developed to store and secure

large amounts of data, on self-encrypting drives. The University of Colorado Office of Information Technology Data Centers on the Anschutz Medical Campus and on the Colorado Springs Campus have restricted physical access. The Data Centers are accessible only via badge access badge access in addition to a PIN for entry. Badge access is given only to essential personnel by senior Office of Information Technology personnel. Logical access is managed and authorized by the department. Access to the data is reserved for the PI and listed team members of the approved project through the Colorado Institutional Review Board (COMIRB). Any changes to staff will be sent to ACHI along with the amended COMIRB application. Data is only accessible to the Principal Investigator and 1 data analyst as part of the research team. All computers are encrypted with required passwords with lockout requirements. See link for more information: 5005 - Information Systems Networked Resource Passwords (ucdenver.edu). All staff and faculty are required to take information security training upon initial employment and renewed during employment at various intervals (depends on position). Firewall penetration requests are routinely conducted. Vulnerabilities are identified during this process. Patches are updated quarterly. Every computer is encrypted and anti-virus software, Microsoft Defender, is used on every computer as well.

ATTACHMENT B: MD APCD DATASET REQUESTED

The MD APCD contains fully processed records for eligibility and professional, institutional, and pharmacy claims for privately fully-insured and non-ERISA self-insured health insurance plans licensed in Maryland for both in-state and out-of-state covered members. Please review the data dictionary before completing this section. Calendar years 2010-2021 are currently avaliable.
MD APCD Data Dictionary:https://mhcc.maryland.gov/mhcc/pages/apcd/apcd_data_release/documents/User_Manual_2019_V1_Codebook.pdf

Standard Analytic Files: Formerly known as the Standard Data Extract. The Standard Analytics Files contain four fixed (i.e., non-customizable) files- the Medicaid Eligibility File, the Professional Services File, and the Pharmacy File. Information about the specific data elements provided within each of the four files can be found in the Data Dictionary. This data set does not include data from Medicare.
Custom Data Sets: A custom data extract can be created based on criteria provided by an Applicant if the data are deemed the minimum amount necessary for an Applicant's proposed use of the data and includes:
a. Indirect individual identifiers that cannot be used to identify indviduals when combined with other information or data; or
b. Aggregate, summary data in which the risk of identifying individuals is minimal.
Custom Data Sets can also include requests for linkage across data sets.

This application should only be completed and submitted for Standard Analytic Files. All requests for Custom Data Files should be sent directly to MHCC at mhcc.datarelease@maryland.gov.

Which MD APCD files are you requesting? Provide a brief justification (1-3 sentences) for each one. Specifically address why this is the minimum necessary data to accomplish the study.

Institutional Claims Medicaid Commercial

Years of Institutional Claims (Specify for Medicaid and Commercial if different) 2014 - 2021

Justification for Requesting Institutional Claims

Because DKA events often end in hospitalizations or ED visits, part of our objective is to characterize the resources used to treat those DKA events. This will allow us to estimate the resources used in various settings to characterize the burden of type 1 diabetes. Many of these patients at initial diagnosis are both in commercial and Medicaid plans.

Professional Claims Medicaid Commercial

Years of Professional Claims (Specify Medicaid and Commercial if different) 2014 - 2021

Justification for Requesting Professional Claims

To identify diagnosis of type 1 diabetes and corresponding DKA events, we need to understand diagnosis codes and CPT codes.

Pharmacy Claims Medicaid Commercial

Years for Pharmacy Claims (Specify Medicaid and Commercial if different) 2014 - 2021

Justification for Requesting Pharmacy Claims

Our objectives include understanding all resource utilization pre and post diagnosis with and without DKA events. That includes pharmacy services.

Member Eligibility Medicaid Commercial

Years for Member Eligibility (Specify Medicaid and Commercial if different) 2014 - 2021

Justification for Requesting Member Eligibility

Understanding race/ethnicity is a big missing data component of type 1 diabetes diagnoses. We plan to attempt to understand demographic differences of those diagnosed with and without DKA events.

ATTACHMENT C: ADDITIONAL DATA SOURCES AND LINKAGE

1. Maryland Medicaid Managed Care Data

Applications for access to Medicaid Managed Care data are sent to the Maryland Medicaid Administration for review and comment. The fields avaliable on the Medicaid MCO data sets have been aligned with Maryland APCD fields to the extent greatest possible. Medicaid Fee for Service data sets are not avaliable.

a. Are you requesting Medicaid data? Yes

b. Do you intend to merge or link Maryland APCD data with Medicaid data? Yes

If yes, provide a brief justification.

Many of those diagnosed with type 1 diabetes are in different plans such as commercial and Medicaid. Part of our objective is to analyze how patients are treated by insurance status - i.e., what resources did they use pre and post diagnosis and by insurance?

c. Describe how the requested Medicaid Managed Care Data meet the minimum necessary standard.

In order for us to analyze outcomes across insurance, we will need access to the Medicaid data.

d. Federal law (42 USC 1396a (a) 7) restricts the use of individually identifiable data of Medicaid recipients to uses that are directly connected to the administration of the Medicaid program. If