



Application for Maryland Medical Care Data Base

TRACKING TABLE (For MHCC Use Only)

MHCC Data Request Number	
Application Received	
Application Approved	
Data Obtained	

INSTRUCTIONS

This form is required for all Applicants requesting MCDB data. Applicants must also complete all the attachments. The completed Application and the Data Management Plan will be used by MHCC to determine whether the request meets the criteria for data release, pursuant to COMAR 10.26.06. Incomplete Applications will be returned to the Applicant, and the request will be delayed. All Applications require a nonrefundable application fee, payable at the time of submission.

Where to submit documents:

- Completed application packages should be scanned and emailed to: mhcc.datarelease@maryland.gov
- A hard copy Application is acceptable and should be sent, with the application fee, to:
Maryland Health Care Commission
4160 Patterson Avenue,
Baltimore, MD 21215,
ATTN: MHCC Data Release
- Please enclose a cover note page that includes the project title, requesting organization's name, and applicant's name.
- If an invoice is needed, please send a request to: mhcc.datarelease@maryland.gov

Note to Applicants:

- Review [data availability and fees](#)
- All Application attachments will be incorporated into the Approved Data Use Agreement (DUA)

Questions? Email mhcc.datarelease@maryland.gov

TABLE OF CONTENTS

INSTRUCTIONS	1
PROJECT INFORMATION	3
ATTACHMENT A: SCOPE OF WORK	4
ATTACHMENT C: ADDITIONAL DATA SOURCES AND LINKAGE	8
ATTACHMENT D: DATA MANAGEMENT PLAN	10
ATTACHMENT E: USE OF CONTRACTORS AND/OR CONSULTANTS (External Entities)	15
ATTACHMENT F: APPLICANT QUALIFICATIONS	16
ATTACHMENT G: OBLIGATIONS AND ATTESTATION	17

PROJECT INFORMATION

Project Title

Advancing Maryland's Statewide Suicide Data Warehouse to Improve Individual and Population-level Mortality Prediction and Prevention

Scheduled Project Start Date 1-1-20

Scheduled Project End Date 12-31-24

MHCC Staff Approved Pre-Application Number

Project Overview: *Provide an abstract or brief summary (150 words) of the specific purpose and objectives of the Project.*

Despite the loss of over 48k lives a year, suicide death is a rare event and challenging to predict. Additionally, suicide models have mainly focused on predicting attempts rather than deaths. Data linkage is underutilized in suicide death prediction and prevention, leading to our 2018 NIMH funded effort to develop the Maryland Suicide Data Warehouse (MSDW). The aim is to leverage the unique data types linked by MSDW to improve suicide death predictions among a more generalizable patient population.

Applicant *(principal investigator, project manager, individual responsible for the research team using the data)*

Name Hadi Kharrazi

Title Associate Professor / Co-Director

E-Mail Address kharrazi@jhu.edu

Telephone Number 443-287-8264

Organization Name Center for Population Health IT, Johns Hopkins Bloomberg School of Public Health

Mailing Address 624 N Broadway, Room 606

City/Town Baltimore

State

MD

Zip Code

21205

Requesting Organization *(Agency, Academic Institution, Research Organization, Company, etc.)*

Organization Name Center for Population Health IT, Health Policy and Management Johns Hopkins Bloomberg School of Public Health

Website <https://www.jhsph.edu/research/centers-and-institutes/johns-hopkins-center-for-population-health-information-technology/>

E-Mail Address elasser1@jhu.edu

Telephone Number

Mailing Address 624 N Broadway, Room 600

City/Town Baltimore

State

MD

Zip Code

21205

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

Name Tom Richards

Title CPHIT Technology Director

E-Mail Address tom.richards@jhu.edu

Telephone Number

Organization/Company *(if different from Requesting Organization)*

Mailing Address

City/Town

State

Zip Code

Relationship to Requesting Organization *(e.g., Contractor)*

Project Contact *(person responsible for all communications with MHCC)*

Name Elyse Lasser

Title Research Associate/Center Coordinator

E-Mail Address elasser1@jhu.edu

Telephone Number

Organization Name

Mailing Address 624 N Broadway, Room 600

City/Town Baltimore

State

MD

Zip Code

21205

ATTACHMENT A: SCOPE OF WORK

1. Project Purpose

- a. Describe the specific research question(s) you are trying to answer or problem(s) you are trying to solve with the MCDB data requested (Please list and number the individual questions), or describe the intended product or report that will be derived from the requested data.

To address the challenges of suicide death prediction, in 2018 NIMH funded our team to develop the Maryland Suicide Data Warehouse (MSDW). The aim was to link a diverse set of data sources and use various layers of risk factors to predict suicide death on a generalizable population. We have since secured additional funding through an R01 grant from NIMH and expanded our aims to include (1) a refresh of existing data and acquiring new data from different sources and replication of descriptive findings, (2) develop and assess hypothesis-driven techniques to predict suicide death, (3) explore hypothesis-generating use cases of the MSDW, and (4) evaluate the generalizability of data sources and methods.

- b. Briefly describe the purpose(s) for which MCDB data are 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.

Suicide is the 10th leading cause of death in the US. Despite the loss of over 48k lives a year, suicide death is a rare event and challenging to predict. Moreover, suicide does not have individual risk factors with high positive predictive value for death that are captured in routine clinical data. Thus, studies have often struggled to integrate multiple risk factors to predict suicide death across generalizable patient populations. This has left most research into predicting suicide deaths dependent on models developed for specific subpopulations such as veterans, arm soldiers, and psychiatric patients. The goal of this study is to leverage data from multiple sources including that from the Maryland Health Care Commission, hospital discharges, and office of medical examiner (OCME) data. Linking these data sources will allow for novel secondary data analyses which will enhance our understanding of risk of suicide death and allow outreach and interventions to be deployed in a timely manner. The predictive models could provide insight into the timing and nature of predictors as well as gaps in care.

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

We have previously shown that linking data is underutilized in youth suicide prevention. Data from payer claims database, mortality data, surveillance data, and electronic health records could be linked to data from suicide prevention efforts to help identify which interventions are most effective in preventing suicidal behaviors, and who benefits from specific interventions and what under what conditions. Our predictive models of suicide death will be in the public domain and can be adopted by various health systems and insurers.

2. Project Methodology

- a. Provide a written description of the project methodology, state the project objectives, the protocol, software and/or identify relevant study questions and analysis method to allow MHCC to understand how the MCDB Data will be used to meet project objectives or address research questions.

Data Sources & Linkage: This is a retrospective study using limited data from various data sources including MHCC MCDB data. We are requesting MCDB (a.k.a., APCD) files at an individual level between 2012 and 2020 for Maryland residents. We understand that 2019 and 2020 data will become available later. CRISP (Maryland's HIE) will receive the MCDB data (if not already available to them), assign an encrypted research ID (EID), and share a limited dataset with the research team at Johns Hopkins. The limited data will include age in years, dates of service, and census block group information. The research team will not receive any protected health information (PHI) except the limited data. The EID is used by the research team to link MCDB data to OCME (office of medical examiner) records. CRISP has also anonymized and tagged OCME data with EID before sharing it with the research team. The EID will be unique to this study and not linked to other data sources outside of this study. Any PHI data provided to CRISP, including MHCC ID, will be used only by CRISP to identify individuals and assign the appropriate study EID as needed.

Study Aims & MCDB: This study aims to better identify patients who may commit suicide. We will develop a series of suicide prediction models to achieve this. MCDB data will provide the research team with clinical data (e.g., diagnostic codes, medications, encounters) that have shown to be predictive of suicide death. The limited data elements (i.e., age, dates of encounter, census block group info) are also vital to the construction of the predictive modeling process (e.g., better construct the trajectory of a temporal predictive model; better understand the neighborhood effect on suicide outcome).

Methods: We will compare individuals who have died by suicide to those who have died by other means as specified in the OCME data. We are also requesting the living patients from MHCC to build a complete control population as recommended by the National Institute of Mental Health (NIMH). Having this control population will allow us to develop a more generalizable prediction model. Once the data is linked, we will develop the predictive analytics using statistical approaches to identify suicide deaths. Using the linked data, we will construct a retrospective case-control study exploring suicide as recorded by the medical examiner and other causes of death. We will first use logistic regression to develop a basic model for suicide risk and compare to other models. We will build the model by using stepwise and LASSO Penalization selection including possible pairwise interactions between variables. Cross-validation will be conducted to fine tune the parameters and avoid over-fitting, especially among the living population. Other more sophisticated models will be used once the data is refreshed and models are explored. Analysis will be stratified by race, sex, insurance type and other attributes as needed.

- b. Do you have approval for your project from an Institutional Review Board (IRB)? All Applicant must obtain review by an Institutional Review Board that has been or will be recognized by the MHCC.
- Yes (attach a copy of the approval to this application)
- No

Enter date submitted to IRB Name of the IRB

3. Publication and Dissemination

Briefly (1-3 sentences) explain any "Yes" answer.

- a. Do you anticipate that the results of your analysis will be published or made publicly available?
- Yes
- No

- i. 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.)?

The primary approach of disseminating will be through published peer-reviewed manuscripts and conferences. Only aggregated results will be published. Information on individuals will not be publicly shared. Additional aggregate results may be shared with the funding agency (NIMH) as requested per NIH guidelines.

- ii. All public displays of MCDB data, regardless of the medium must comply with MCDB's cell size suppression policy, as set forth in the Data Use Agreement. Please explain how you will ensure that any public display will not disclose a cell size (suppression) that is less than 11, and percentages or other mathematical formulas that result in the display of a cell size (suppression) that is less than 11.

All tables, figures and reporting of results will adhere to a cell suppression size of 10 or fewer individuals. No reporting or public presentation of proportions, percentages or proportional weights will allow inference of a cell size less than 11 individuals for any product of this work, including records from MHCC. Our team will review internal and public work products continually to ensure appropriate cell suppression and all public releases will be reviewed by the principle investigator and an additional team member to ensure compliance.

- iii. Identify the lowest geographical level of analysis of data you will to 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 identified?

CRISP will share MCDB data for areas with more than 10 individuals using an encrypted ID associated to each individual. The lowest geographical level will be census block groups (with at least more than 10 individuals). Maps will not include any geographies that may represent less than 10 patients.

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

- i. Will the MCDB data be used for consulting purposes? Yes No

- ii. Will report(s), website(s) or a statistical tabulation(s) using MCDB data be shared or sold? Yes No

Internal reports will be shared with National Institute of Mental Health (NIMH) -- which is an NIH institute. No reports/tabulations will be sold.

iii. Will a software product using MCDB data be shared or sold? Yes No

iv. Will MCDB data be used as input to develop a product (i.e., severity index tool, a risk adjustment tool, a reference tool, etc.)? Yes No

v. Will MCDB data be sold or shared in any format not noted above? Yes No

If yes, in what format and who are the purchaser of the data?

vi. Will the project result in disclosing MCDB 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? Yes No

vii. Will the results from the project be used for price transparency? Yes No

viii. Will health care providers be individually identified? Yes No. Please describe your protocol for informing health care providers.

ATTACHMENT B: MCDB DATASET REQUESTED

MHCC collects privately insured data (claims and membership), known as the Medical Care Data Base (MCDB), on a quarterly basis from life and health insurance carriers, health maintenance organizations (HMOs), third party administrators (TPAs), and pharmacy benefits managers (PBMs) that are licensed to do business in Maryland. The MCDB data that is available for release contains eligibility and professional, institutional, and pharmacy claims. Starting in 2015, the Medical Care Data Base (MCDB) excludes private plan data for self-insured ERISA due to the *Gobeille v. Liberty Mutual* Supreme Court ruling.

The data which is refreshed and updated annually contains only privately fully-insured and self-insured non-ERISA health insurance plans for Maryland and non-Maryland residents. The MCDB comprises of about 90-95% of the privately fully insured market and just about 25% - 30% of the self-insured market (post-Gobeille, primarily non-ERISA). To determine the years for which data are available check on the [MHCC website](#). That site also contains information about the most current MCDB Release Version and a full list of elements in the release including the release record layouts, data dictionaries, and supporting documentation.

1. Which MCDB files are you requesting? Provide a brief justification (1-3 sentences) for each one.

Dataset	Year(s)
<input checked="" type="checkbox"/> Institutional Claims	2012-2020
Institutional claims will be used to assess inpatient care utilization and ED events. Diagnostic and procedure data collected in an inpatient stay are shown to be predictive of suicide death. We understand that 2019 and 2020 are still not available for release.	
<input checked="" type="checkbox"/> Professional Claims	2012-2020
Outpatient services claims will be used to assess standard care utilization for patients of different subpopulations. Diagnostic and procedure data collected in outpatient settings are shown to be predictive of suicide death. We understand that 2019 and 2020 are still not available for release.	
<input checked="" type="checkbox"/> Pharmacy Claims	2012-2020
Derived measures from pharmacy claims will allow us to assess facets to a patient's care that are otherwise unavailable, like medication adherence and gaps in medication possession. Medication data has also shown to be predictive of suicide death. We understand that 2019 and 2020 are still not available for release.	
<input checked="" type="checkbox"/> Member Eligibility	2012-2020
The member eligibility file will be used to inform and validate demographic attributes to each patient that otherwise may be unavailable. Age and gender will be used as features of predictive models while race/ethnicity will allow us to test for model bias that may result in disparity in performance for minority populations. Tracking enrollment and eligibility for healthcare insurance will be of additional predictive interest. We understand that 2019 and 2020 are still not available for release.	

ATTACHMENT C: ADDITIONAL DATA SOURCES AND LINKAGE

1. Medicaid Data

Applications for access to Medicaid Managed Care data for studies comparing the privately insured to Medicaid Managed Care patients can be submitted, but require a separate approval from the Maryland Medicaid Administration. The fields available on the Medicaid MCO data sets have been aligned with MCDB fields to the extent possible.

- a. Indicate whether you are seeking Medicaid data: Yes No
- b. Do you intend to merge or link MCBDB data with Medicaid data? Yes No
If yes, provide brief justification.

A separate request for Medicaid enrollment data is being pursued with the MDH and the Hilltop institute. the MHCC data set will be linked with this export once it has been approved and released through CRISP.

**We have updated our application to include requesting Medicaid data.

- c. 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 you are requesting Maryland Medicaid Data, please describe, in the space below, why your use of the Data meets this requirement.

The requested data from Medicaid/Hilltop will not include individually identifying information. All patient level identifying information will be removed prior to release from CRISP.

2. Other Linkages

Data linkage involves combining MCDB data with other data to create a more extensive database for analysis.

1. Do you intend to merge or link MCBDB Data with other data? Yes No

If Yes:

- a. What are the files to be linked?

MCDB data will be primarily linked to OCME data using CRISP hashed/encrypted ID.

Other external data used in this project (that will not be directly linked with MCDB) include: the Health Services Cost Review Commission (HSCRC), Johns Hopkins Health System, Anne Arundel Medical Center, Peninsula Regional Medical Center, Maryland Department of Health, Veteran's Health Administration of Maryland, and CRISP Insight.

- b. Why is this linkage needed?

The external data are required to provide the study team with the outcome of interest (i.e., suicide death determination), control group markers (e.g., other types of death and/or no death at all), and additional predictors (e.g., neighborhood census variables). The diversity of data sources strengthens the case for generalizability in predicting a rare event such as suicide death.

- c. Which MCDB data elements will be linked to the data elements in the external file?

We are primarily interested in aggregating observations that inform us of patient diagnoses, medical and psychiatric service codes, and medication records. These MCDB data elements will be linked to OCME data to improve the prediction of suicide deaths.

- d. What methodology or algorithm will be used to create this match? If you intend to create a unique algorithm, please describe how it will link each dataset.

We will work closely with CRISP (and MHCC) to assign CRISP EIDs as the primary key to enable linkage of the data sources (e.g., we already have OCME data with CRISP EIDs; i.e., the study IDs).

e. What variables from each of the source files will be included in the final linked analytic file?

We will have multiple analytical files to evaluate the value of various data types in predicting suicide death. The MCDB data will be linked with OCME data, and will include variables such as demographics, diagnoses, procedures, encounters, and medication fill at a patient level (from MCDB) as well as manner of death and medical examiner's clinical data (from OCME).

2. Explain why the linkages are needed.

Linkage across the different data sources is the only way to observe course of care and prominent features that are related to risk of suicide death for known instances of suicide. This project is the first application of this kind for the state of Maryland (funded by NIH) and represents one of the most ambitious efforts to predict suicide death in the general population.

3. Describe the specific steps the Organization will take to prevent the identification of individuals in the linked files.

The data will be stripped by CRISP from essential PHI and only a limited data set will be provided to the research team. The linked files will not be shared with any other research team and will only be used for the purpose of this study. CRISP will also apply additional censoring as directed by MHCC such as limiting the cell size (e.g., geographical areas) to at least 10 individuals. The limited data files, along with CRISP EID, will remain on HIPAA-compliant servers at Johns Hopkins.

ATTACHMENT E: USE OF CONTRACTORS AND/OR CONSULTANTS (External Entities)

Provide the following information for all consultant and contractors who will have access to the MCDB data. The Requesting Organization must have a written agreement with the contractor/consultant to ensure the use of MCDB data to the approved project(s) of this application as well as the privacy and security standards set forth in the Data Use Agreement. MCDB data may not be shared with any third party without prior written consent from MHCC, or an amendment to this Application.

Entity	<input type="checkbox"/> Contractor	<input type="checkbox"/> Subcontractor	<input type="checkbox"/> Consultant
Organization Name			
Title			
Website			
Contact Person			
E-Mail Address			
Telephone Number			
Mailing Address			
City/Town		State	Zip Code
Term of Contract			

1. Describe the tasks and products assigned to this entity for this project.

2. Describe the qualifications of this entity to perform and complete the tasks.

3. Describe the Requesting Organization's oversight and monitoring of the activities and actions of this entity for this project, including how you will ensure the privacy and security of the MCDB data to which the consultant or contractor has access.

4. Will this entity have access to or store the MCDB data at a location other than the data custodian location, off-site server and/or database? Yes No.

If yes, a separate Data Management Plan **must** be completed by this contractor/consultant.

[INSERT A NEW SECTION FOR ADDITIONAL CONTRACTOR/CONSULTANT ENTITIES NEEDED]

ATTACHMENT F: APPLICANT QUALIFICATIONS

1. Describe previous experience using claims data. This question should be answered by the primary investigator/project manager and should encompass the experience of the entire project team who will be using the data.

Hadi Kharrazi MHI MD PhD, PI of the project, is the associate director of the Center for Population Health IT (CPHIT) at the Johns Hopkins School of Public Health (JHSPH) with a joint appointment at the Division of Health Science Informatics at the Johns Hopkins School of Medicine. Dr. Kharrazi will supervise the day-to-day operations of the project and will be involved in the completion of all specific aims including dissemination plans. He will also be involved in the technical design, overall data warehouse architecture, and data linkage process of the proposed solutions. Dr. Kharrazi has an extended experience with HIEs, clinical data, predictive modeling, and suicide data sources.

Dr. Hadi Kharrazi and the research team have 15+ years of experience working with confidential data including the HSCRC case mix data at the census block group. Current work involving HSCRC data includes identifying risk of falls in various populations in Maryland. The team also has extensive experience working with CRISP to link data at individual and geographic levels. Dr. Kharrazi and some members of the research team have been working with the Maryland Department of Health and CRISP linking data sources together at an individual level to identify individuals at risk of opioid overdose and overuse.

The Johns Hopkins Center for Population Health IT (CPHIT) has years of experience handling large scale data warehouses. Under the leadership of Dr. Jonathan Weiner and Dr. Kharrazi, we have developed and maintained a database of novel data sources used to predict opioid overdose and overuse. This project included utilizing a CRISP encrypted ID to link juvenile justice data, HSCRC data, MHCC data, and other data sources. This data is only available to that specific research team and protected using a similar protocol that we are proposing here.

CPHIT coordinated use of MHCC MCDB data for the Johns Hopkins University for a number of years (certificate of data destruction completed in April 2020) with 2010-2016 MCDB data. Tom Richards served as the data custodian/manager and Elyse Lasser as the project coordinator.


2. Resumes/CVs: When submitting your Application package, include résumés or curricula vitae of the principal investigator/project manager and any project team with relevant experience

ATTACHMENT G: OBLIGATIONS AND ATTESTATION

ATTESTATION OF APPLICANT

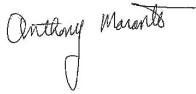

I, Hadi Kharrazi, Applicant, solemnly affirm under penalties of perjury that the information contained in the Application its attachments, and this Attestation, is true and correct to the best of my knowledge, information and belief and that the requested MCDB data is the minimum necessary to accomplish the Project. I accept my obligation to comply with all requirements in this Application and attachments, including:

- (1) Compliance with all data privacy and security obligations;
- (2) Execution of a Data Use Agreement approved by MHCC-staff prior to receipt of the requested data;
- (3) Responsibility for assuring that the data has been destroyed at the conclusion of the project in accordance with the terms and conditions of the Data Use Agreement;
- (4) Responsibility for assuring that specified MHCC staff is notified within 30 days when any person who has access to the MCDB data is removed from or added to the MHCC-approved Project;
- (5) Responsibility for assuring that each required report is sent to the MHCC staff within the time period specified in the Data Use Agreement; and
- (6) Continuing compliance with the Data Management Plan.

Applicant's signature:	
Printed Name:	Hadi Kharrazi
Title:	Associate Professor and Center Co-Director
Date:	Jul 21, 2021

ATTESTATION OF REQUESTING ORGANIZATION

I, Anthony Maranto, Assistant Director of Johns Hopkins University, the Requesting Organization in this Application, have been duly authorized by the Requesting Organization to execute this attestation on its behalf. I solemnly affirm under penalties of perjury that the information contained in the Application, its attachments, and this Attestation, is true and correct to the best of my knowledge, information and belief.

Signature of authorized representative of the Requesting Organization:	 
Printed Name:	Anthony Maranto
Title:	Assistant Director
Date:	3/2/2022