Exploring associations of gut and liver diseases Epidemiology Study Research Protocol

|  |  |
| --- | --- |
| INVESTIGATIONAL PRODUCT(S): | Not applicable |
| STUDY NUMBER(S): | *IRB Number* | 852719 |
| Other Protocol Identifiers  | Not applicable |
| PROTOCOL(S) TITLE: | Exploring Associations of Gut and Liver diseases Epidemiology Study (EAGLES) |
| REGULATORY SPONSOR: | University of Pennsylvania  |
| FUNDING SPONSOR(S): | Investigator-initiated study |
| PRINCIPAL INVESTIGATOR | James D. Lewis, MD |
| ORIGINAL PROTOCOL DATE: | 11/28/2022 |
| VERSION NUMBER: | EAGLES v1.1 |
| VERSION DATE: | 11/7/2023 |

|  |  |
| --- | --- |
| [1 Study Summary 5](#_Toc119572210)[1.1 Synopsis 5](#_Toc119572211)[1.2 Key Roles and Study Governance 7](#_Toc119572212)[2 Introduction and Rationale 7](#_Toc119572213)[2.1 Study Rationale 7](#_Toc119572214)[2.2 Background 7](#_Toc119572215)[3 Risk/Benefit Assessment 8](#_Toc119572216)[3.1.1 Potential Risks 8](#_Toc119572217)[3.1.2 Potential Benefits 8](#_Toc119572218)[3.1.3 Assessment of Potential Risks and Benefits 8](#_Toc119572219)[4 Study Objectives 9](#_Toc119572220)[5 Study Plan 9](#_Toc119572221)[5.1 Study Design 9](#_Toc119572222)[6 Study Population 9](#_Toc119572223)[6.1 Inclusion Criteria 9](#_Toc119572224)[7 Data collection methods 10](#_Toc119572225)[8 Creation of standardized data sets 11](#_Toc119572226)[9 standardized data set storage and access 11](#_Toc119572227)[10 Statistical Considerations 12](#_Toc119572228)[10.1 Statistical Hypotheses 12](#_Toc119572229)[10.2 Sample Size Determination 12](#_Toc119572230)[10.3 Statistical Analyses 12](#_Toc119572231)[10.3.1 General Approach 12](#_Toc119572232)[11 Supporting Documentation and Operational Considerations 12](#_Toc119572233)[11.1 Regulatory, Ethical, and Study Oversight Considerations 12](#_Toc119572234)[11.1.1 Waiver of Informed Consent and Health Insurance Portability and Accountability Act (HIPAA) Authorization 12](#_Toc119572235)[11.1.2 Study Discontinuation and Closure 13](#_Toc119572236)[11.1.3 Subject Confidentiality and Privacy 13](#_Toc119572237)[11.1.4 Future Use of Stored Specimens and Data 13](#_Toc119572238)[11.1.5 Safety Oversight and Monitoring 14](#_Toc119572239)[11.1.6 Data Collection and Management 14](#_Toc119572240)[12 Study Finances 14](#_Toc119572241)[12.1 Funding Source 14](#_Toc119572242)[12.2 Conflict of Interest 15](#_Toc119572243)[13 Publication Plan 15](#_Toc119572244)[14 References 15](#_Toc119572245)[15 APPENDIX 15](#_Toc119572246)[15.1 Appendix 1. Limitations on data use 15](#_Toc119572247)[15.2 Appendix 2. Data Stewardship Best Practices 17](#_Toc119572248) |  |
|  |  |

|  |
| --- |
| PRINCIPAL INVESTIGATOR SIGNATURE |
| STUDY SPONSOR: | James Lewis (Investigator Sponsor) |
| STUDY TITLE: | Exploring Associations of Gut and Liver diseases Epidemiology Study (EAGLES) |
| STUDY IDPROTOCOL VERSION | 852719V1.1.20231107 |
| I have read the referenced protocol. I agree to conduct the study in accordance to this protocol, in compliance with the Declaration of Helsinki, Good Clinical Practices (GCP), and all applicable regulatory requirements and guidelines. |
| Principal Investigator Name  | James D. Lewis, MD |  | Signature |  |
| Affiliation: | University of Pennsylvania |  | Date |  |
|  |  |  |  |  |

Abbreviations

|  |  |
| --- | --- |
| C Difficile | Clostridioides difficile |
| CPT | Common Procedural Terminology |
| EAGLES | Exploring Associations of Gut and Liver diseases Epidemiology Study |
| GCP | Good Clinical Practice |
| GERD | Gastroesophageal Reflux Disease |
| GLD | Gastro and Intestinal Liver Disease |
| H. Pylori | Helicobacter Pylori |
| HCPCS | Health Care Common Procedure Coding System |
| HIPAA | Health Insurance Portability and Accountability Act  |
| HIV | Human Immunodeficiency Virus |
| IBD | Inflammatory Bowel Disease |
| ICD | International Classification of Disease |
| PHI | Protected Health Information |
| SNOMED | Systematized Nomenclature of Medicine |
| UPHS | University of Pennsylvania Health System |

# Study Summary

## Synopsis

|  |  |
| --- | --- |
| Title:  | Exploring Associations of Gut and Liver diseases Epidemiology Study (EAGLES) |
| Short Title: | EAGLES Study |
| Study Description:  | This is a retrospective study of the epidemiology, comparative effectiveness, comparative safety, natural history and outcomes of the diseases of the gastrointestinal tract, pancreas, biliary system, liver, and spleen. . |
| Objectives:  | To studythe epidemiology, comparative effectiveness, comparative safety, natural history and outcomes of the diseases of the gastrointestinal tract, pancreas, biliary system, liver, and spleen.  |
| Primary Endpoint:  | The primary endpoint is hospital admission.  |
| Secondary Endpoints: | The secondary endpoints can include any outcomes associated with diseases of the gastrointestinal tract, pancreas, biliary system, liver, and spleen. |
| Study Population: | All patients cared for within the University of Pennsylvania Health System can be included in this study. We are particularly focused on those with disease of the gastrointestinal tract, pancreas, biliary system, liver, and spleen. Depending on the association being investigated, we may also study a sample of patients without the disease(s) of interest.  |
| Phase: | Not applicable |
| Description of Sites/Facilities  | Not applicable |
| Enrolling Sites: | Not applicable |
| Study Duration: | 20 years |
| Participant Duration: | Not applicable |
|  |  |

## Key Roles and Study Governance

Dr. James Lewis is the principal investigator of the study. Drs. Marina Serper and Shivan Mehta will be co-principal investigators. Should any of these investigators leave the University of Pennsylvania, the remaining investigators may appoint one or more additional co-principal investigators. This study will be reviewed and approved by the University of Pennsylvania Institutional Review Board and will follow all Good Clinical Practice (GCP) guidelines.

# Introduction and Rationale

## Study Rationale

Gastrointestinal and liver diseases (GLD) are common, reduce quality of life and incur large healthcare expenditures. There is an unmet need for research that will improve our understanding of the epidemiology, comparative effectiveness and safety of therapies, resource utilization, and outcomes of patients with GLD. In addition, quality improvement is an ongoing goal for the Division of Gastroenterology and Hepatology. The objective of this master protocol is to facilitate the use of data gathered within the usual care of patients from the University of Pennsylvania Health System to address these unmet needs.

## Background

GLD include a wide range of diseases. These are generally divided into those of the luminal gastrointestinal tract, the liver, the pancreas and the pancreaticobiliary system (Table 1). It is notable that some GLD seem to be rising in prevalence at rates that suggest a role for environmental factors, such as the epidemic of obesity, fatty liver disease, eosinophilic esophagitis, etc.

Table 1. Examples of common gastrointestinal and liver diseases (not all inclusive)

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Luminal diseases | Liver diseases | Pancreatic diseases | Pancreaticobiliary |
| Acid peptic diseases  | GERD, peptic ulcer disease |  |  |  |
| Motility disorders | Achalasia, gastroparesis, colonic inertia |  |  | Sphincter of oddi dysfunction |
| Infectious diseases | Infectious diarrhea, H pylori, C difficile | Viral hepatitis | Mumps | HIV |
| Inflammatory and immune disorders | Celiac disease, eosinophilic esophagitis, inflammatory bowel diseases | Primary biliary cholangiopathy, autoimmune hepatitis | Pancreatitis | Primary sclerosing cholangitis |
| Premalignant and malignant  | Barrett’s esophagus, gastric metaplasia, cancers | Hepatocellular carcinoma, cirrhosis | Pancreatic cancer | Primary sclerosing cholangitis |
| Metabolic diseases | Lactose intolerance | Inborn errors of metabolism, hemochromatosis, Wilson’s disease, cirrhosis | Endocrine or exocrine pancreatic insufficiency, cystic fibrosis | Gall stones |
| Functional disorders not otherwise defined | Nonulcerative dyspepsia, irritable bowel syndrome |  |  |  |

Care for patients with GLD is complex due to the many different diseases and the great advances that have been made in recent decades. For example, hepatitis C has gone from an untreatable disease to one that is nearly 100% curable. Therapy for inflammatory bowel diseases (IBD) has progressed from having 3 classes of medications to now having at least 8 different medication classes with more in development. However, these new drugs all have a unique risk – benefit profile. During the same time, advances in diagnostic testing and endoscopic therapies have changed the way that we diagnose and manage patients with GLD. For example, the advent of MRI enterography has reduced radiation exposure for patients with Crohn’s disease and the development of minimally invasive endoscopic surgery has changed the management of achalasia.

The rapidly evolving landscape of GLD demands a highly efficient research infrastructure to be able to rapidly assess new hypotheses. The EAGLES study is designed to address this need.

# Risk/Benefit Assessment

### Potential Risks

As a retrospective study, the EAGLES study poses minimal risk to patients other than the potential for loss of confidentiality.

### Potential Benefits

As a retrospective study, there will be no direct benefits to patients for participation.

### Assessment of Potential Risks and Benefits

While there is no direct benefit expected for the participants, the balance of the participants’ risk vs. the knowledge to be gained is believed to be favorable.

# Study Objectives

The overall objectives of the EAGLES study are to: 1) improve our understanding of the epidemiology, comparative effectiveness and safety of therapies, resource utilization, direct, indirect and total cost of care and out of pocket costs, and outcomes of patients with GLD, 2)provide a platform for collecting data to support quality improvement initiatives related to GLD, 3) to facilitate the use of data gathered within the usual care of patients from the University of Pennsylvania Health System to address these unmet needs.

# Study Plan

## Study Design

This is a retrospective cohort study utilizing data collected as part of routine care of patients with GLD. This study will use data from multiple sources that are described below:

1. Penn’s Epic inpatient and outpatient electronic medical record including, Epic Clarity data, any Epic registries that have been created, such as the IBD registry and Nutrition registry. This includes all coded data (such as demographic, medical diagnoses, procedures, services, prescriptions and other information on medications, health behaviors, and health related surveys). Free text from encounters and reports of studies may also be used to generate new variables. We will access UPHS SDS/Performance Manager cost accounting system to evaluate direct, indirect, total costs of care associated with all inpatient and outpatient encounters. We will also evaluate patient out of pocket costs associated with inpatient encounters, outpatient encounters, and pharmacy encounters.

1. Provation endoscopy reporting system – the Provation software is used to generate reports of endoscopic procedures using a mix of defined fields and free text. Still images and videos of the endoscopic procedures are also stored in the Provation reporting system.
2. Raw data files (DICOM) and digital images from radiology studies, such as CT scans, MRI, etc.
3. Pathology slides, digital images of pathology slides, and tissue blocks created in the usual care of patients. Note that these samples will not be used to isolate DNA which can be used for genetic testing.
4. Switchboard data
5. Navicare data
6. Carelign data
7. Medview data

# Study Population

## Inclusion Criteria

1. The key inclusion criteria for the GLD cohort are 1) age 18 years or older and 2) being diagnosed with or treated for any GLD or undergoing evaluation for a GLD at the University of Pennsylvania Health System (UPHS) on or after January 1, 2000. Most GLD may be identified from the following using International Classification of Disease (ICD) version 10 (or the corollary version 9) codes, Health Care Common Procedure Coding System (HCPCS), Systematized Nomenclature of Medicine (SNOMED) and Common Procedural Terminology (CPT) codes. Examples of ICD10 and CPT codes related to GDL include, but are not limited, to those listed below:
	1. B15-B19 viral hepatitis
	2. C15-C26 malignant neoplasms of digestive organs
	3. D10-D36 benign neoplasms
	4. D50-D53 nutritional anemias
	5. D60-D64 other anemias
	6. I80-I89 diseases of veins, lymphatic vessels and lymph nodes (includes esophageal varices)
	7. K00-K95 diseases of the digestive system
	8. K70-K77 diseases of the liver
	9. R90-R94 abnormal findings on diagnostic imaging and functional studies
	10. R00-R19 symptoms and signs
	11. R70-79 Abnormal findings on examination of the blood
	12. T15-T19 Foreign bodies
	13. Z80-89 Personal and family history of influencing health status (such as personal history of colon polyps)
	14. Z12.11 Encounter for screening for malignant neoplasms of colonm
	15. G0105, Colorectal cancer screening; colonoscopy on individual at high risk
	16. CPT codes for endoscopic procedures (e.g. 43191-43278)
	17. CPT and ICD codes for bowel, pancreas and liver surgery
2. Patients age 18 years or older who have been hospitalized at any UPHS hospital or seen in the family medicine or general internal medicine clinics on or after January 1, 2000 will serve as a control cohort.

# Data collection methods

Data may be collected for this study in multiple ways as described below:

Billing diagnoses and procedures using International Classification of Disease (ICD) version 9 and 10, Health Care Common Procedure Coding System (HCPCS), Systematized Nomenclature of Medicine (SNOMED) and Common Procedural Terminology (CPT) codes will be extracted from Epic Clarity database.

Standardized data elements contained in Smartforms and other Epic tools will be extracted.

Demographic data, social history, family history and problem list data will be extracted from Epic.

All prescriptions and medication data, including infusions, will be extracted from Epic.

Endoscopic procedure data will be extracted from the Provation database.

Free text records in Epic and Provation may be manually searched. Natural language processing may be used to extract data from the free text fields.

# Creation of standardized data sets

To facilitate this research, we will generate a standardized data set that will be updated over time. Elements in the standardized data sets will included the following (each with the associated dates):

* Diagnoses and provider specialty
* All prescribed and over the counter medications as recorded in Epic, including dates of prescriptions and discontinuation
* Hospitalizations and the primary and secondary discharge diagnoses
* Emergency department encounters and the discharge diagnosis
* Endoscopic procedures
* Radiology and other imaging procedures
* GLD related surgeries
* Epic’s demographics information (date of birth, race, ethnicity, sex, gender, occupation)
* Epic’s Past medical history table
* Epic’s Past surgical history table
* Epic’s Family history table
* Epic’s Social history table
* Epic’s Immunization table
* Epic’s Problem list
* Medical record number (this is needed to allow additional chart review to gather data from free text)

We will curate the standardized data sets to create derived variables relevant to the various conditions to be studied.

# standardized data set storage and access

Data will be managed by an honest broker who has received the necessary training consistent with Penn Medicine’s policies. The honest broker will be responsible for creating and maintaining the data sets and providing investigators with access to data cuts needed to support the proposed research and quality improvement projects. When possible, the honest broker will provide the investigator with completely deidentified data sets, in which dates will be offset and all protected health information (PHI) removed. When necessary to complete the project, limited data sets can be provided, including, when necessary, provision of medical record numbers to facilitate data validation and other processes that require review of medical records or related research activities. Any investigator obtaining access to a data set that contains PHI will be added to the study protocol and must have completed the appropriate training in human subject research.

Certain data may not be used for research purposes as outlined in Appendix 1. The honest broker will follow these rules.

# Statistical Considerations

## Statistical Hypotheses

As an infrastructure protocol, there are no predefined hypotheses. Rather, users of the data infrastructure will propose their own hypotheses and leverage the data to answer these.

## Sample Size Determination

There is no set sample size for the cohort. It will continue to grow as time goes forward.

## Statistical Analyses

### General Approach

The nature of this infrastructure protocol is that each investigator will propose a different hypothesis to test. For the given hypothesis, the actual patients included, the design and the analysis will need to be created specifically to support the aims of the investigator.

# Supporting Documentation and Operational Considerations

## Regulatory, Ethical, and Study Oversight Considerations

### Waiver of Informed Consent and Health Insurance Portability and Accountability Act (HIPAA) Authorization

We will not seek informed consent for patients to participate in this study. This is a low risk retrospective study. Moreover, upon agreeing to receive care within the University of Pennsylvania Health System, all patients agree that their data may be used for research purposes. Finally, this study would not be possible without a waiver of informed consent. There are more than 100,000 patients, some of whom are no longer alive. As such, obtaining informed consent from all patients would not be practical. Thus, this study meets the standards for waiver of informed consent based on the following criteria:

* The research involves no more than minimal risk to subjects;
* The research could not be carried out practicably without the waiver or alteration;
* The waiver or alteration will not adversely affect the rights and welfare of the subjects; **and**,
* Where appropriate, the subjects will be provided with additional information about their participation.

We will not seek HIPAA authorization since no data with direct identifiers will be shared with investigators outside of Penn. We will need to collect the medical record number to facilitate manual data collection from the electronic medical record and other data sources that use the medical record number where required.

### Study Discontinuation and Closure

This study may be temporarily suspended or prematurely terminated by the Principal Investigator (PI) at any time and for any reason. In the event that the PI changes institutions or retires, a new PI will be named and oversight will become the responsibility of the new PI.

### Subject Confidentiality and Privacy

Participant confidentiality and privacy is strictly held in trust by the participating investigators and their staff. This confidentiality is extended to cover testing of biological samples in addition to the clinical information relating to participants. Therefore, the study protocol, documentation, data, and all other information generated will be held in strict confidence. No information concerning the study, or the data will be released to any unauthorized third party.

Representatives of the Institutional Review Board (IRB), and other relevant regulatory agencies may inspect all documents and records required to be maintained by the investigator, including but not limited to, medical records (clinic, or hospital) for the participants in this study. The PI and University of Pennsylvania Health System (UPHS) will permit access to such records.

Users will be instructed to store study participant research data including direct identifiers, for purposes of statistical analysis and scientific reporting, on password protected servers or password protected computers. See Appendix 2 for Data Stewardship Best Practices. Whenever possible, we will provide investigators with deidentified data sets. When direct identifiers are needed to support the research, investigators will follow the protocol as described in 11.1.6.3.

### Future Use of Stored Specimens and Data

No samples will be stored for future use. However, this protocol will allow for use in research of existing residual pathology samples that are retained by the Department of Pathology and were obtained as part of usual care. Use of such residual samples will only occur once the necessary specimens have been used for clinical care purposes. Note that these samples may not be used to isolate DNA which can be used for genetic testing.

The data in this protocol may be linked to data and samples from other studies and biobank/repositories conducted within the University of Pennsylvania, if the other protocol does not preclude such linkage.

Data collected for this study may be shared with other investigators for research purposes only at the discretion of the Principal Investigator. Data with direct identifiers will not be released outside of the University of Pennsylvania. Any research use of the data beyond the scope of this protocol will require a separate IRB approved protocol.

### Safety Oversight and Monitoring

Given that the risks of this study are minimal, we will not have a data safety monitoring committee or designated a safety officer. No clinical monitoring is planned for this study given the minimal risk. The PI and co-PIs will be responsible for assuring the study is conducted according to this protocol.

### Data Collection and Management

This protocol supports two principal data management options as described below.

#### Ad hoc data pulls

Investigators may request an ad hoc data pull consistent with this protocol. The data will be extracted from the databases listed above and provided to the investigator in a secure manner.

#### Centralized standardized data

The study team will create standardized data pulls that will be stored in a EAGLES database that will be managed by the honest broker. These data will be used to create standardized data elements that can be used to expedite the implementation of research using the data. These data tables can be provided to investigators with or without direct personal identifiers as needed based on the research or QI proposal.

#### Access to data

In order to gain access to the study data, an investigator must have an active Penn card and have proof of training in human subject research ethics. Investigators wishing to gain access to data that contains direct identifiers, specifically the medical record number, must complete a protocol specific training that describes appropriate data custody practices (see Appendix 2). Investigators will be instructed to archive on a password protected server or other encrypted electronic storage device or deleted from all devices at the end of the investigator’s sub-study.

# Study Finances

## Funding Source

This study will be funded by the Penn Center for Molecular Studies of Digestive and Liver Diseases.

## Conflict of Interest

Any investigator who has a conflict of interest with this study (patent ownership, royalties, or financial gain greater than the minimum allowable by their institution, etc.) must have the conflict reviewed by a properly constituted Conflict of Interest Committee with a Committee-sanctioned conflict management plan that has been reviewed and approved by the study sponsor prior to participation in this study. All University of Pennsylvania investigators will follow the University conflict of interest policy.

# Publication Plan

The study team hopes to publish the results of this study. All publications and presentations should reference support of the Penn Center for Molecular Studies of Digestive and Liver Diseases.

# References

# **APPENDIX**

## Appendix 1. Limitations on data use

**Data that needs to be removed from research datasets before sharing**

1) All encounters at the Community Connect practices and data documented under such encounters.

Community Connect practices only lease our Epic instance, so although their data resides therein, it is not Penn’s to use.

If the patient has *only* the visits of this kind, the patient needs to be removed from any patient list that you might generate.

PAT\_ENC JOIN CLARITY\_DEP ON CLARITY\_DEP.DEPARTMENT\_ID = PAT\_ENC.DEPARTMENT\_ID

AND CLARITY\_DEP.SERV\_AREA\_ID IN (5001, 5002, 5003)

2) Patients with special chart protections / VIP (all data)

PATIENT\_TYPE.PATIENT\_TYPE\_C = 1002

3) Patients who opted out of research (all data)

PATIENT\_FYI\_FLAGS.PAT\_FLAG\_TYPE\_C = '1079'

4) Patients with 'Do Not Solicit' flag (all data)

PATIENT\_TYPE.PATIENT\_TYPE\_C = 16

5) Genetic test results

NOTE: all data *below* may be included in de-identified datasets, as per the Privacy office.

6) Data of pediatric patients (age < 18)

7) HIV/AIDS information

a) diagnoses

b) all data generated at HIV clinics

c) lab tests (e.g. viral count)

8) Behavioral health information

a) diagnoses

b) *encounter info* at behavioral health locations

9) Substance use information

a) diagnoses

b)*all data* generated at Substance Use Disorder treatment program (a.k.a. Part 2) locations (within: Charles O’Brien Center, PPMC Inpatient Services, PPMC Outpatient Services, and Princeton House)

**Restricted departments**

Definition includes clinics/centers/units for: Behavioral health / psych / crisis response, Infectious diseases, Opioid recovery / Substance use disorders, Maternal medicine / fertility, Pain management, Eating disorders, Plastic surgery

Clinic encounters:

SELECT \* FROM PAT\_ENC JOIN CLARITY\_DEP ON CLARITY\_DEP.DEPARTMENT\_ID = PAT\_ENC.EFFECTIVE\_DEPT\_ID WHERE CLARITY\_DEP.RESTRICTED\_DEPT\_YN = 'Y'

Hospital visits:

SELECT \* FROM PAT\_ENC\_HSP JOIN CLARITY\_DEP ON CLARITY\_DEP.DEPARTMENT\_ID = PAT\_ENC\_HSP.DEPARTMENT\_ID WHERE CLARITY\_DEP.RESTRICTED\_DEPT\_YN = 'Y'

**Sensitive diagnoses**

Definition includes HIV, substance abuse, behavioral/developmental issues

SELECT \* FROM CLARITY\_EDG JOIN X\_EXCLUSION ON X\_EXCLUSION.RECORD\_ID = CLARITY\_EDG.DX\_ID

**NOTE**: NO NEED TO EXCLUDE: **lab tests** and **medications** related to the sensitive condition (e.g. HIV tests and behavioral health meds) documented at non-restricted care locations (primary care, family practice, internal medicine, etc.)

## Appendix 2. Data Stewardship Best Practices

The data provided under the EAGLES protocol will be used for retrospective study of the epidemiology, comparative effectiveness, comparative safety, natural history and outcomes of the diseases of the gastrointestinal tract, pancreas, biliary system, liver, and spleen.

The purpose of this document is to promote sound data governance. Maintaining the security of the data acquired under this protocol is a paramount concern. Data will be provided to investigators in a secure manner which may include upload of the data on PennBox.

All data obtained under this protocol and any datasets derived from those data, must be stored on a password protected server (preferably a Penn server) or other encrypted electronic storage device. Qualified servers must meet all HIPAA Privacy Rules. It is the responsibility of each investigator to ensure that the server is qualified. Once analyses are complete and the data are no longer in use, the data must be destroyed/removed from the investigator’s server or encrypted device.

The data obtained under this protocol will not be shared with anyone else. The data obtained under this protocol may only be used for the purpose under which the data were requested. All users of the data requested under this protocol must have human subjects research ethics training. Members of the study team working with the data for this purpose, must abide by the best practices listed here. Study team members must be employed by the University of Pennsylvania or the University of Pennsylvania Health System. Team members employed outside of Penn are not eligible to receive or view raw data. Summary data may be shared with team members outside of Penn. Any investigator obtaining access to a data set that contains PHI will be added to the study protocol and must have completed the appropriate training in human subject research.

Any research use of the data beyond the scope of this protocol will require a separate IRB approved protocol.