Title : Secure Academic Data Commons - TEMPLATE for participating institutions.

Approval Period: **Draft**

Protocol Director

Name		**Degree (Program/year if student)**		**Position, e.g. Assistant Professor, Resident, etc.**
Lesley Park		PhD		IMPORTANT PI
Department	**Mail Code**	**Phone**	**Fax**	**E-mail**
Med/Primary Care and Population Health	5560	(650) 723-2513		lesley.park@stanford.edu
CITI Training current				N

Admin Contact

Name		**Degree (Program/year if student)**		**Position, e.g. Assistant Professor, Resident, etc.**
Isabella Chu		MPH		Assoc Director - Data Core
Department	**Mail Code**	**Phone**	**Fax**	**E-mail**
Medicine - Med/General Internal Medicine	5411	(650) 723-2513		isabella.chu@stanford.edu
CITI Training current				Y

Investigator

Name		**Degree (Program/year if student)**		**Position, e.g. Assistant Professor, Resident, etc.**
Department	**Mail Code**	**Phone**	**Fax**	**E-mail**
CITI Training current				

Other Contact

Name		**Degree (Program/year if student)**		**Position, e.g. Assistant Professor, Resident, etc.**
Department	**Mail Code**	**Phone**	**Fax**	**E-mail**
CITI Training current				

Academic Sponsor

Name		**Degree (Program/year if student)**		**Position, e.g. Assistant Professor, Resident, etc.**
Department	**Mail Code**	**Phone**	**Fax**	**E-mail**
CITI Training current				

Other Personnel

Participant Population(s) Checklist	**Yes/No**
• Children (under 18)	Y

Title : Secure Academic Data Commons - TEMPLATE for participating institutions.
Approval Period: **Draft**

• Pregnant Women and Fetuses	N
• Neonates (0 - 28 days)	N
• Abortuses	N
• Impaired Decision Making Capacity	N
• Cancer Subjects	N
• Laboratory Personnel	N
• Healthy Volunteers	N
• Students	N
• Employees	N
• Prisoners	N
• Other (i.e., any population that is not specified above)	Y

Study Location(s) Checklist

	Yes/No
• Stanford University	Y
• Clinical & Translational Research Unit (CTRU)	Y
• Stanford Hospital and Clinics	Y
• Lucile Packard Children's Hospital (LPCH)	Y
• VAPAHCS (Specify PI at VA)	
• Other (Click ADD to specify details)	

General Checklist

Multi-site

	Yes/No
• Is this a multi-site study? A multi-site study is generally a study that involves one or more medical or research institutions in which one site takes a lead role.(e.g., multi-site clinical trial)	N

Collaborating Institution(s)

	Yes/No
• Are there any collaborating institution(s)? A collaborating institution is generally an institution that collaborates equally on a research endeavor with one or more institutions.	Y

Institution Name	Contact Name	Contact Phone	Contact Email	Permission?	Engaged?
OTHER UNIVERSITY	someone	650-555-1212	someone@university.edu	Y	N

Cancer Institute

	Yes/No
• Cancer-Related Studies (studies with cancer endpoints), Cancer Subjects (e.g., clinical	N

trials, behavior/prevention) or Cancer Specimens (e.g., blood, tissue, cells, body fluids with a scientific hypothesis stated in the protocol).

Clinical Trials	**Yes/No**
• Investigational drugs, biologics, reagents, or chemicals?	N
• Commercially available drugs, reagents, or other chemicals administered to subjects (even if they are not being studied)?	N
• Investigational Device / Commercial Device used off-label?	N
• IDE Exempt Device (Commercial Device used according to label, Investigational In Vitro Device or Assay, or Consumer Preference/Modifications/Combinations of Approved Devices)	N
• Will this study be registered on# clinicaltrials.gov? (See Stanford decision tree)	N
• Is Stanford responsible for ClinicalTrials.gov registration? (See Stanford decision tree) NCT#	N

Tissues and Specimens	**Yes/No**
• Human blood, cells, tissues, or body fluids (tissues)?	N
• Tissues to be stored for future research projects?	N
• Tissues to be sent out of this institution as part of a research agreement? For guidelines, please see https://sites.stanford.edu/ico/mtas	N

Biosafety (APB)	**Yes/No**
• Are you submitting a recombinant DNA vector or Human Gene Transfer investigation using biological agents? If yes, please complete and attach the Gene Transfer Protocol Application Supplemental Questions to section 16 of the eProtocol application.	N
• Are you submitting a Human study using biohazardous/infectious agents? If yes, refer to the http://www.stanford.edu/dept/EHS/prod/researchlab/bio/index.html Administrative Panel on BioSafety website prior to performing studies.	N
• Are you submitting a Human study using samples from subjects that are known or likely to contain biohazardous/infectious agents? If yes, refer to the http://web.stanford.edu/dept/EHS/prod/researchlab/bio/index.html Administrative Panel on BioSafety website prior to performing studies.	N

Human Embryos or Stem Cells	**Yes/No**
• Human Embryos or Gametes?	N
• Human Stem Cells (including hESC, iPSC, cancer stem cells, progenitor cells)	N

Veterans Affairs (VA)	**Yes/No**

Title : Secure Academic Data Commons - TEMPLATE for participating institutions.
Approval Period: **Draft**

• The research recruits participants at the Veterans Affairs Palo Alto Health Care System(VAPAHCS).	N
• The research involves the use of VAPAHCS non-public information to identify or contact human research participants or prospective subjects or to use such data for research purposes.	N
• The research is sponsored (i.e., funded) by VAPAHCS.	N
• The research is conducted by or under the direction of any employee or agent of VAPAHCS (full-time, part-time, intermittent, consultant, without compensation (WOC), on-station fee-basis, on-station contract, or on-station sharing agreement basis) in connection with her/his VAPAHCS responsibilities.	N
• The research is conducted using any property or facility of VAPAHCS.	N

Equipment	**Yes/No**
• Use of Patient related equipment? If Yes, equipment must meet the standards established by Hospital Instrumentation and Electrical Safety Committee (650-725-5000)	N
• Medical equipment used for human patients/subjects also used on animals?	N
• Radioisotopes/radiation-producing machines, even if standard of care? http://www.stanford.edu/dept/EHS/prod/researchlab/radlaser/Human_use_guide.pdf More Info	N

Payment	**Yes/No**
• Subjects will be paid/reimbursed for participation? See payment considerations.	N

Funding	**Yes/No**
• Training Grant?	N
• Program Project Grant?	N
• Federally Sponsored Project?	Y
• Industry Sponsored Clinical Trial?	N

Funding

Funding - Grants/Contracts

Funding - Fellowships

Gift Funding

Dept. Funding

Other Funding

Title : Secure Academic Data Commons - TEMPLATE for participating institutions.
Approval Period: **Draft**

Other Fund Name : Our Grant

Expedited Form

A protocol must be no more than minimal risk (i.e., "not greater than those ordinarily encountered in daily life") AND must only involve human subjects in one or more of the following paragraphs.

Select one or more of the following paragraphs:

1. N **Clinical studies of drugs and medical devices only when condition (a) or (b) is met.**
   a) Research on drugs for which an investigational new drug application (21 CFR Part 312) is not required. (Note: Research on marketed drugs that significantly increases the risks or decreases the acceptability of the risks associated with the use of the product is not eligible for expedited review.)
   b) Research on medical devices for which
      i) an investigational device exemption application (21 CFR Part 812) is not required; or
      ii) the medical device is cleared/approved for marketing and the medical device is being used in accordance with its cleared/approved labeling.

2. N **Collection of blood samples by finger stick, heel stick, ear stick, or venipuncture as follows:**
   a) from healthy, nonpregnant adults who weigh at least 110 pounds. For these subjects, the amounts drawn may not exceed 550 ml in an 8 week period and collection may not occur more frequently than 2 times per week; or
   b) from other adults and children, considering the age, weight, and health of the subjects, the collection procedure, the amount of blood to be collected, and the frequency with which it will be collected. For these subjects, the amount drawn may not exceed the lesser of 50 ml or 3 ml per kg in an 8 week period and collection may not occur more frequently than 2 times per week.

3. N **Prospective collection of biological specimens for research purposes by non invasive means.**

4. N **Collection of data through non invasive procedures (not involving general anesthesia or sedation) routinely employed in clinical practice, excluding procedures involving x-rays or microwaves. Where medical devices are employed, they must be cleared/approved for marketing. (Studies intended to evaluate the safety and effectiveness of the medical device are not generally eligible for expedited review, including studies of cleared medical devices for new indications.)**
   Examples:
   a) physical sensors that are applied either to the surface of the body or at a distance and do not involve input of significant amounts of energy into the subject or an invasion of the subject's privacy;
   b) weighing or testing sensory acuity;
   c) magnetic resonance imaging;
   d) electrocardiography, electroencephalography, thermography, detection of naturally occurring radioactivity, electroretinography, ultrasound, diagnostic infrared imaging, doppler blood flow, and echocardiography;
   e) moderate exercise, muscular strength testing, body composition assessment, and flexibility testing where appropriate given the age, weight, and health of the individual.

5. Y **Research involving materials (data, documents, records, or specimens) that have been collected, or will be collected solely for nonresearch purposes (such as medical treatment or diagnosis). (NOTE: Some research in this paragraph may be exempt from the HHS regulations for the protection of human subjects. 45 CFR 46.101(b)(4). This listing refers only to research that is not exempt.)**

6. N **Collection of data from voice, video, digital, or image recordings made for research**

Title : Secure Academic Data Commons - TEMPLATE for participating institutions.
Approval Period: **Draft**

purposes.

7. N **Research on individual or group characteristics or behavior(including, but not limited to, research on perception, cognition, motivation, identity, language, communication, cultural beliefs or practices, and social behavior) or research employing survey, interview, oral history, focus group, program evaluation, human factors evaluation, or quality assurance methodologies. (NOTE: Some research in this category may be exempt from the HHS regulations for the protection of human subjects. 45 CFR 46.101(b)(2) and (b)(3). This listing refers only to research that is not exempt.)**

Resources :

a) Qualified staff.

Please state and justify the number and qualifications of your study staff.

Lesley Park is a Professor of Data Science and has more than 20 years experience in data analyses studies for health research. Lesley Park, Isabella Chu, ADD YOUR CENTRAL RESEARCH IT TEAM ON THE PROTOCOL AND DESCRIBE THEM HERE and Research IT staff are experienced research staff, most with advanced degrees and extensive training in data management and risk mitigation.

b) Training.

Describe the training you will provide to ensure that all persons assisting with the research are informed about the protocol and their research-related duties and functions.

All personnel approved as SPHS personnel or approved to use analytic data sets will have completed required data security, HIPAA and human subject protection training. We will also require all personnel to sign a DUA which outlines the use of the data for their project and attests that they will not attempt to download data, reidentify individuals in de-identified datasets or misuse the data.

c) Facilities.

Please describe and justify.

We have described our Data Management Plan in great detail in the attachment in section 16.

In brief:
Sensitive and restricted data sets are received by the Stanford Principal Investigator (PI) via encrypted disk or whatever mode is preferred by the distributor of the data, and are saved to a secure GCP instance or server that abides by Stanford's computer and network usage policy (http://adminguide.stanford.edu/62.pdf) and information security policy (http://adminguide.stanford.edu/63.pdf). The server is behind a firewall with access restricted to authorized investigators only. Two step authentication is required at each sign on to access the data. Intrusion monitoring software is handled through the central university network office and locally as a backup. Data are encrypted and backed up nightly onsite for both onsite and offsite data storage. On prem servers are housed in a secure location with very limited access. This location is monitored 24/7 via camera and has limited keycard access. In case of power failure both the server and network switches are tied into the natural gas generator to eliminate any down time and/or disruption in service.

All files and curation steps are logged in the PHS Data portal. Logs include the date files are received, file size, variables, count, distribution and missingness of each variable, and other descriptive data.

When individuals use the data, the system records: Project and study PI, project title, IRB approval date, date DUA was signed and received, variables included in the analytic data set and date data were made

available on the server. Data access is routinely audited to ensure that there is no inappropriate access.

2. Shared File System:

All computers and workstations used to access and process analytic data sets including laptops and desktops (both office and home computers) that may be used to access participant data will be password protected, encrypted, and backed-up per Stanford policy.

No person will be permitted to download data onto a computer, laptop or workstation. All analytic data sets will be kept on the server. Investigators will access the server via a secure VPN which requires an active authorized account and two-step authentication.

d) Sufficient time.

Explain whether you will have sufficient time to conduct and complete the research. Include how much time is required.

Lesley Park, Isabella Chu AND OTHER TEAM MEMBERS have 100% of their time protected for research and management of these data.

e) Access to target population.

Explain and justify whether you will have access to a population that will allow recruitment of the required number of participants.

We have ample resources to obtain large data sets for use by Stanford researchers.

These data sets will generally represent large cross sections of the population (ie, 20% Medicare Sample, claims data from 68 million commercially insured people).

f) Access to resources if needed as a consequence of the research.

State whether you have medical or psychological resources available that participants might require as a consequence of the research when applicable. Please describe these resources.

This is a data repository protocol for receipt, storage, curation, de-identification and granting access to analytic data. Not applicable.

g) Lead Investigator or Coordinating Institution in Multi-site Study.

Please explain (i) your role in coordinating the studies, (ii) procedures for routine communication with other sites, (iii) documentation of routine communications with other sites, (iv) planned management of communication of adverse outcomes, unexpected problems involving risk to participants or others, protocol modifications or interim findings.

1. Purpose

a) In layperson's language state the purpose of the study in 3-5 sentences.

PLEASE NOTE: THIS IS A TEMPLATE. PLEASE SUBSTITUTE YOUR DEPARTMENT NAME OR THE NAME OF YOUR DATA CENTER FOR PHS. THIS IS ONLY TO SUGGEST LANGUAGE WHICH YOU CAN USE TO SUBMIT TO YOUR IRB.

The goal of the YOUR DEPARTMENT OR DATA COMMONS HERE is to 1) Create a library of data assets that is so compelling and easy to use that it spurs transdisciplinary projects across the university. 2)

Title : Secure Academic Data Commons - TEMPLATE for participating institutions.
Approval Period: **Draft**

Innovation in data hosting, management and enabling cutting edge data science. 3) The creation of new datasets by combining data from disparate sources in novel ways, act as a catalyst for new interdisciplinary research directions, questions, and challenges.

b) State what the Investigator(s) hope to learn from the study. Include an assessment of the importance of this new knowledge.

DEPARTMENT hopes to learn the best way to make data and analytic tools available to a wide variety of researchers. This protocol is not for research per se, but a data repository.

c) Explain why human subjects must be used for this project. (i.e. purpose of study is to test efficacy of investigational device in individuals with specific condition; purpose of study is to examine specific behavioral traits in humans in classroom or other environment)

This research necessarily involves human subjects as the research questions will focus on questions of human activity, health and well-being.

2. Study Procedures

a) Please SUMMARIZE the research procedures, screening through closeout, which the human subject will undergo. Refer to sections in the protocol attached in section 16, BUT do not copy the clinical protocol. Be clear on what is to be done for research and what is part of standard of care.

Each dataset has different inclusion/exclusion criteria. Most of the datasets are administrative data that have been passively collected for other purposes (eg, EMRs and claims). However, some datasets have been collected expressly for research purposes and we will make an effort to make these readily available to researchers (eg, HRS).

DESCRIBE EXAMPLES OF YOUR MOST HIGH VALUE DATASETS. NO NEED TO DESCRIBE ALL OF THEM. BE SURE TO INCLUDE MAJOR CATEGORIES.

Examples of data Sets:

1. 20% Medicaid Sample or Research Identifiable files. I have attached the specifications worksheet in Section 16 which describes which data sets and years we are requesting.

2. Optum data: The Optum data are a large dataset of commercial claims for over 150 million unique members.

3. MarketScan data: MarketScan® Databases are the largest of their kind in the industry with data on more than 200 million unique patients since 1995. These data contain claims from large commercially insured populations.

4. Our Data Management Plan is attached in Section 16. This outlines the procedures for receipt, curation, de-identification and granting access to data.

5. Datasets from international partners. These data will vary in content and sensitivity. All data will be stored and disbursed in accordance with the Data Management Plan attached in Section 16.

6. All datasets available can be viewed on our portal at phsdata.stanford.edu or datacommons.stanford.edu.

Other data curation activities include:

Data Overlay: Data may be combined by some aggregating variable such as time or place as part of our data curation and preparation procedures. For example, data containing ER visits for respiratory distress or

asthma may be combined with data on particulate matter and forest fires and poverty or income rates by zip code or census tract. We will engage in ongoing monitoring of reidentification risk and ensure that all outputs removed from our secure environment are aggregated in such a way that individuals are not identifiable. We will also work with data scientists to innovate and automate this risk assessment.

Data Linkage: Where possible or permissible, we will enrich and link data by individual in order to allow investigators to answer research questions around exposures and behaviors and health. This linkage and variable identification will take place under the PHS Data Core protocols. In most cases, we will remove explicit identifiers following linkage as described in the Data Management Plan attached in Section 5.1. Individuals working with PHI or PII will be required to obtain IRB approval.

Innovation in Data Management and hosting: Data management, risk assessment, de-identification and curation are still highly manual processes. Using data over which we have provenance, we will work with experts to find ways to more efficiently, accurately and excellently accomplish these tasks. For example, at present, the declaration of risk relies largely on the investigator. This process should be automated.

Additionally, the costs of data hosting are still considerable, innovations in computational technologies will allow substantial cost savings to institutions over time. Stanford PHS will work with partners to develop better technologies to manage risk and costs associated with data hosting.

A full list of datasets can be found on our portal at: phsdata.stanford.edu

b) Explain how the above research procedures are the least risky that can be performed consistent with sound research design.

The greatest risk is in this study is data breach. Confidentiality of data will be ensured as described in our data management plan.

Beyond breach of confidentiality, there is no risk to subjects as these data have already been collected.

c) State if deception will be used. If so, provide the rationale and describe debriefing procedures. Since you will not be fully informing the participant in your consent process and form, complete an alteration of consent (in section 13). Submit a debriefing script (in section 16).

Data repository protocol. Not applicable.

d) State if audio or video recording will occur. Describe what will become of the recording after use, e.g., shown at scientific meetings, erased. Describe the final disposition of the recordings.

Data repository protocol. Not applicable.

e) Describe alternative procedures or courses of treatment, if any, that might be advantageous to the participant. Describe potential risks and benefits associated with these. Any standard treatment that is being withheld must be disclosed in the consent process and form. (i.e. standard-of-care drug, different interventional procedure, no procedure or treatment, palliative care, other research studies).

Data repository protocol. Not applicable.

f) Will it be possible to continue the more (most) appropriate therapy for the participant(s) after the conclusion of the study?

Data repository protocol. Not applicable.

g) Study Endpoint. What are the guidelines or end points by which you can evaluate the different treatments (i.e. study drug, device, procedure) during the study? If one proves to be clearly more effective than another (or others) during the course of a study, will the study be terminated before the projected total participant population has been enrolled? When will the study end if no important differences are detected?

Data repository protocol. Not applicable.

Title : Secure Academic Data Commons - TEMPLATE for participating institutions.
Approval Period: **Draft**

3. Background

a) Describe past experimental and/or clinical findings leading to the formulation of the study.

Many of these datasets have been used extensively for epidemiologic research. We expect that these uses will continue and hope (and are focused on) expanding the use of these data into fields where they have not been traditionally used.

b) Describe any animal experimentation and findings leading to the formulation of the study.

Not applicable.

4. Radioisotopes or Radiation Machines

a) List all standard of care procedures using ionizing radiation (radiation dose received by a subject that is considered part of their normal medical care). List all research procedures using ionizing radiation (procedures performed due to participation in this study that is not considered part of their normal medical care). List each potential procedure in the sequence that it would normally occur during the entire study. More Info

Identify Week/Month of study	Name of Exam	Identify if SOC or Research

b) For research radioisotope projects, provide the following radiation-related information:

Identify the radionuclide(s) and chemical form(s).

For the typical subject, provide the total number of times the radioisotope and activity will be administered (mCi) and the route of administration.

If not FDA approved provide dosimetry information and reference the source documents (package insert, MIRD calculation, peer reviewed literature).

c) For research radiation machine projects, provide the following diagnostic procedures:

For well-established radiographic procedures describe the exam.

For the typical subject, identify the total number of times each will be performed on a single research subject.

For each radiographic procedure, provide the setup and technique sufficient to permit research subject dose modeling. The chief technologist can usually provide this information.

For radiographic procedures not well-established, provide FDA status of the machine, and information sufficient to permit research subject dose modeling.

d) For research radiation machine projects, provide the following therapeutic procedures:

For a well-established therapeutic procedure, identify the area treated, dose per fraction and number of fractions. State whether the therapeutic procedure is being performed as a normal part of clinical management for the research participants's medical condition or whether it is being performed because the research participant is participating in this project.

For a therapeutic procedure that is not well-established, provide FDA status of the machine, basis for dosimetry, area treated, dose per fraction and number of fractions.

5. Devices

a) Please list in the table below all Investigational Devices (including Commercial Devices used off-label) to

Title : Secure Academic Data Commons - TEMPLATE for participating institutions.
Approval Period: **Draft**

be used on participants

b) Please list in the table below all IDE Exempt Devices (Commercial Device used according to label, Investigational In Vitro Device or Assay, or Consumer Preference/Modifications/Combinations of Approved Devices) to be used on participants.

6. Drugs, Reagents, or Chemicals

a) Please list in the table below all investigational drugs, reagents or chemicals to be administered to participants.

b) Please list in the table below all commercial drugs, reagents or chemicals to be administered to participants.

7. Medical Equipment for Human Subjects and Laboratory Animals

If medical equipment used for human patients/participants is also used on animals, describe such equipment and disinfection procedures.

Data repository protocol. Not applicable.

8. Participant Population

a) State the following: (i) the number of participants expected to be enrolled at Stanford-affiliated site(s); (ii)the total number of participants expected to enroll at all sites; (iii) the type of participants (i.e. students, patients with certain cancer, patients with certain cardiac condition) and the reasons for using such participants.

This is a data repository and the datasets vary in terms of their size and scope. For example, one dataset represents the entire population of Denmark for four generations. We will facilitate their use by PHS researchers.

Other marquee datasets include:
20% CMS sample
MarketScan: 200 million lives, 8 years, commercial claims data
Optum: 150 million lives, 12 years, commercial claims data

Data will come from a wide variety of sources including but not limited to:

1. Data which are purchased or obtained explicitly for PHS (ie, CMS, Optum).
2. Data arrangements where the Director of PHS or his representative reaches an agreement with the data owner to either allow access to the data for PHS members or for PHS to import the data. (Danish Biobank, Indian Railroad, Taiwanese health claims).
3. Data which are already on campus where an investigator wishes to make the data more broadly available. (H-CUP, CHIS, etc).

Data Core Vision and plan attached in Section 16.

Title : Secure Academic Data Commons - TEMPLATE for participating institutions.
Approval Period: **Draft**

b) State the age range, gender, and ethnic background of the participant population being recruited.

All ages, sexes, genders and, we hope, races and ethnicities.

c) State the number and rationale for involvement of potentially vulnerable subjects in the study (including children, pregnant women, economically and educationally disadvantaged, decisionally impaired, homeless people, employees and students). Specify the measures being taken to minimize the risks and the chance of harm to the potentially vulnerable subjects and the additional safeguards that have been included in the protocol to protect their rights and welfare.

Vulnerable subjects are either included incidentally as these are large claims datasets (eg, Truven) or data are collected on vulnerable populations explicitly (eg, March of Dimes) to address risks and needs specific to that population.

d) If women, minorities, or children are not included, a clear compelling rationale must be provided (e.g., disease does not occur in children, drug or device would interfere with normal growth and development, etc.).

Not applicable. No excluded populations.

e) State the number, if any, of participants who are laboratory personnel, employees, and/or students. They should render the same written informed consent. If payment is allowed, they should also receive it. Please see Stanford University policy.

Not applicable. Inclusion of employees in any datasets will be incidental.

f) State the number, if any, of participants who are healthy volunteers. Provide rationale for the inclusion of healthy volunteers in this study. Specify any risks to which participants may possibly be exposed. Specify the measures being taken to minimize the risks and the chance of harm to the volunteers and the additional safeguards that have been included in the protocol to protect their rights and welfare.

Health status varies widely. Many of these datasets are representative of the general population so have large percentage of data from healthy individuals included.

g) How will you identify participants for recruitment? (E.g., by: chart review; referral from treating physician; response to ad). Attach recruitment materials in Section #16 (Attachments). All Final or revised recruitment materials, flyers, etc. must be submitted to the IRB for review and approval before use. You may not contact potential participants prior to IRB approval. See Advertisements: Appropriate Language for Recruitment Material.

The majority of datasets are delivered to PHS already de-identified. However, in some cases, the de-identification is carried out by the data core team.

Analysts in the Stanford PHS, who regularly work with the type of data to be requested, will fulfill an "honest broker" role. They will enrich and improve the utility of datasets via data overlay and linkage (where permissible), risk scores (where applicable), harmonization, organization, cleaning, and, if necessary, de-identify the pertinent data.

The Stanford PHS data portal and data visualization tools allow investigators to explore available datasets, do cursory analysis and visualization and decide which data are most appropriate for their study question.

Access to limited datasets will be granted to researchers who have completed the procedures outlined in Section 1.3.

h) Inclusion and Exclusion Criteria.

Identify inclusion criteria.

The study population will be determined by the original owner or collector of the data. In some cases, the datasets include the entire population of a country so there are virtually no inclusion or exclusion criteria. In other cases, datasets have very specific criteria.

Identify exclusion criteria.

Title : Secure Academic Data Commons - TEMPLATE for participating institutions.
Approval Period: **Draft**

As above. The study population will be determined by the original owner or collector of the data. In some cases, the datasets include the entire population of a country so there are virtually no inclusion or exclusion criteria. In other cases, datasets have very specific criteria.

i) Describe your screening procedures, including how qualifying laboratory values will be obtained. If you are collecting personal health information prior to enrollment (e.g., telephone screening), please request a waiver of authorization for recruitment (in section 15).

Analysts in the Stanford Center for Population Health Sciences, who regularly work with the type of data to be requested, will fulfill an "honest broker" role. They will overlay, link, organize, and, if necessary, de-identify the pertinent data. Data will not be shared outside the personnel designated for each project and will only be used for approved research projects. In the case where data are de-identified, or a researcher is conducting exploratory work, they will be covered under the Protocol 40974 Stanford PHS Data Protocol: Data Exploration, Cohort Identification Educational Activities and Research in Population Sciences.

PHS data are only shared with individuals who: a) have completed required Data Security, HIPAA and human subject protection (CITI) training(s) as applicable b) have had all electronic devices which may potentially be used to access the data or analytic datasets encrypted c) provided proof of this this encryption either via the institutional encryption tracking and verification system or a signed statement from the Data Security Office of the researcher's institution d) have institutional review board (IRB) approval for their study or are included on an IRB approved study, e) have a "need to know" status with regard to the data and cannot practically work on the project without it and; f) signed a data use agreement with Stanford PHS stating that they will only use the data for the stated and IRB approved research purposes and that they cannot share the data with any third party, download the data and will abide by all regulatory and security requirements for that dataset. In some cases it will be necessary to complete special trainings and separate approvals as required by the data owner. These will be tracked in the Stanford PHS data portal within the users account.

In addition to the conditions above, the signed Data Use Agreement will stipulate that only output results from analyses can be downloaded – that is, the researcher agrees never to download data or subsets of those/analysis files or publish a formula which could be used to derive a cell with fewer than 11 individuals and that all data outputs will conform to the data owners or Stanford PHS cell size policies (no cell smaller than 11), whichever is more restrictive. In rare instances, data download is permitted, however, individuals downloading analytic datasets must receive prior written approval from PHS and additional encryption and security verification.

In the event of an employee leaving Stanford of their own volition or being terminated their access to the data will be terminated.

DATA REPORTING AND PUBLICATION

Stanford will not disclose any non-aggregated information derived from analyses including listings, or information derived from the file(s) with or without direct identifiers, if such findings, listings, or information can, by themselves or in combination with other data, be used to deduce an individual's identity. Additionally, Stanford will not identify or report any identifiable pharmacy, provider or prescriber in any publication. All findings will be reported in aggregate. Tables or outputs with cell sizes less than 10 will not be published or made public.

Stanford agrees that any use of data in the creation of any document (manuscript, table, chart, study, report, etc.) will adhere to the distributor of the data's current cell size suppression policy or Stanford's cell suppression policy, whichever is more restrictive. No cell (e.g. admittances, discharges, patients, services) 10 or less may be displayed or used in any publication. Also, no use of percentages or other mathematical formulas will be used if they result in the display of a cell 10 or less.

COMPLETION OF RESEARCH TASKS AND DATA DESTRUCTION

The Stanford Principal Investigator or their designee must complete the Study Form on the PHS Data Portal (phsdata.stanford.edu). All trainings and compliance information will be tracked using the administrative panel.

The portal keeps detailed records on study personnel, which role they have, protocol or personnel changes (on an annual basis or as they occur, whichever is more frequent), and conflict of interest. In addition to listing relevant personnel on IRB protocols, the portal also confirms that required data security, human subjects and privacy trainings are completed and up to date. In the event research projects are completed, the Stanford PI records the completion of the project and the disposition of data access as terminated. The Principal Investigator also notes protocol closures and posts a link to resultant products of research in the data portal.

j) Describe how you will be cognizant of other protocols in which participants might be enrolled. Please explain if participants will be enrolled in more than one study.

Data repository protocol. Not applicable.

k) Payment/reimbursement. Explain the amount and schedule of payment or reimbursement, if any, that will be paid for participation in the study. Substantiate that proposed payments are reasonable and commensurate with the expected contributions of participants and that they do not constitute undue pressure on participants to volunteer for the research study. Include provisions for prorating payment. See payment considerations

Data repository protocol. Not applicable.

l) Costs. Please explain any costs that will be charged to the participant.

Not costs to subjects. Data repository protocol.

m) Estimate the probable duration of the entire study. Also estimate the total time per participant for: (i) screening of participant; (ii) active participation in study; (iii) analysis of participant data.

This is a data repository. i) we anticipate data updates at least annually for most datasets. ii) We anticipate that the repository will remain at Stanford for a minimum of five years. iii) Protocols for analyses will continue for a minimum of one year and will be submitted separately to the IRB.

9. Risks

a) For the following categories include a scientific estimate of the frequency, severity, and reversibility of potential risks. Wherever possible, include statistical incidence of complications and the mortality rate of proposed procedures. Where there has been insufficient time to accumulate significant data on risk, a statement to this effect should be included. (In describing these risks in the consent form to the participant it is helpful to use comparisons which are meaningful to persons unfamiliar with medical terminology.)

The risks of the Investigational devices.

Data repository protocol. Not applicable.

The risks of the Investigational drugs. Information about risks can often be found in the Investigator's brochure.

Data repository protocol. Not applicable.

The risks of the Commercially available drugs, reagents or chemicals. Information about risks can often be found in the package insert.

Data repository protocol. Not applicable.

The risks of the Procedures to be performed. Include all investigational, non-investigational and non-invasive procedures (e.g., surgery, blood draws, treadmill tests).

Analysts in the Stanford Center for Population Health Sciences, who regularly work with the type of data to

be requested, will fulfill an "honest broker" role. They will link, organize, and, if necessary, de-identify the pertinent data. Data will not be shared outside the personnel designated for each project and will only be used for approved research projects. All personnel approved as Stanford-PHS personnel or approved to use analytic data sets will have a) completed required HIPAA and human subject protection training; and b) have "need to know" status with regard to the data and cannot practically work on the project without it and c) if applicable, have been approved to access analytic data sets by the original owner of the data and the Stanford IRB.

In the event of an employee leaving Stanford of their own volition or being terminated their access to the data will be terminated.

DATA REPORTING AND PUBLICATION

Stanford does not disclose any non-aggregated information derived from analyses including listings, or information derived from the file(s) with or without direct identifiers, if such findings, listings, or information can, by themselves or in combination with other data, be used to deduce an individual's identity. Additionally, Stanford will not identify or report any identifiable pharmacy, provider or prescriber in any publication. All findings will be reported in aggregate. Tables or outputs with cell sizes less than 10 will not be published or made public.

Stanford agrees that any use of data in the creation of any document (manuscript, table, chart, study, report, etc.) will adhere to the distributor of the data's current cell size suppression policy or Stanford's cell suppression policy, whichever is more restrictive. No cell (e.g. admittances, discharges, patients, services) 10 or less may be displayed or used in any publication. Also, no use of percentages or other mathematical formulas will be used if they result in the display of a cell 10 or less.

COMPLETION OF RESEARCH TASKS AND DATA DESTRUCTION

The Stanford Principal Investigator must be included in the Study Form on the PHS Data portal for projects which wish to use data.

The PHS Data Portal (phsdata.stanford.edu) keeps records on study personnel, which role they have, protocol or personnel changes (on an annual basis or as they occur, whichever is more frequent), and conflict of interest. In addition to listing relevant personnel on IRB protocols, the portal Study Form and administrative panel tracks and verifies that all required trainings are complete and up to date. In the event research projects are completed, the Stanford PI records the completion of the project, products of research and the disposition of data access.

The risks of the Radioisotopes/radiation-producing machines (e.g., X-rays, CT scans, fluoroscopy) and associated risks.

Data repository protocol. Not applicable.

The risks of the Physical well-being.

We do not anticipate that this study will impact the physical well-being of participants.

The risks of the Psychological well-being.

We do not anticipate that this study will impact the psychological well-being of participants.

The risks of the Economic well-being.

We do not anticipate that this study will impact the economic well-being of participants.

The risks of the Social well-being.

We do not anticipate that this study will impact the social well-being of participants.

Overall evaluation of Risk.

Low - innocuous procedures such as phlebotomy, urine or stool collection, no therapeutic agent, or safe therapeutic agent such as the use of an FDA approved drug or device.

Title : Secure Academic Data Commons - TEMPLATE for participating institutions.
Approval Period: **Draft**

b) If you are conducting international research, describe the qualifications/preparations that enable you to both estimate and minimize risks to participants. Also complete the International Research Form and attach it in the Attachments section. If not applicable, enter N/A.

In the case where we have international data partners, we will adhere to all data governance laws and policies of the host country and institution.

c) Describe the planned procedures for protecting against and minimizing all potential risks. Include the means for monitoring to detect hazards to the participant (and/or to a potential fetus if applicable). Include steps to minimize risks to the confidentiality of identifiable information.

Please see the detailed description provided in the data management plan attached in Section 16. In brief:

1. For the most part, data will be de-identified before being transmitted to Stanford. 2. Storage and distribution of data will adhere to Stanford's restricted data policies: http://adminguide.stanford.edu/62.pdf and information security policy http://adminguide.stanford.edu/63.pdf. as described in our Data Management Plan attached in Section 16.

We will also work with data scientists to advance innovations around data security and risk detection.

d) Explain the point at which the experiment will terminate. If appropriate, include the standards for the termination of the participation of the individual participant Also discuss plans for ensuring necessary medical or professional intervention in the event of adverse effects to the participants.

Data repository protocol. Not applicable.

e) Special Participant Populations

Children's Findings OHRP. As children are involved in your research, please select one or more regulatory categories (46.404 through 46.407) below that your research falls under and provide the necessary rationale for each determination. See full regulation citation.

Y 46.404 Research not involving greater than minimal risk. The research must present no greater than minimal risk to children and adequate provisions must be made for soliciting the assent of the children and the permission of their parents or guardians. Please provide rationale for the above statement.

Rationale for category selected above

We intend to collect and store datasets which may contain information about children. We will make limited datasets available to researchers who have an IRB approval to do the research and have completed all required encryption and trainings as outlined in the Data Management Plan in section 1.3.

10. Benefits

a) Describe the potential benefit(s) to be gained by the participants or by the acquisition of important knowledge which may benefit future participants, etc.

No direct benefits will accrue to the study subjects. This repository will enable analyses which may potentially lead to a greater understanding of health and well-being.

11. Privacy and Confidentiality

Privacy Protections

a) Describe how the conditions under which interactions will occur are adequate to protect the privacy interests of participants (e.g., privacy of physical setting for interviews or data collection, protections

Title : Secure Academic Data Commons - TEMPLATE for participating institutions.
Approval Period: **Draft**

for follow-up interactions such as telephone, email and mail communications).

Data repository protocol. Not applicable.

Confidentiality Protections

b) Specify PHI (Protected Health Information). PHI is health information linked to HIPAA identifiers (see above). List BOTH health information AND HIPAA identifiers. If you are using STARR, use the Data Privacy Attestation to ensure that your request will match your IRB-approved protocol. Be consistent with information entered in section 15a.

Datasets will vary widely in the PHI fields they make available. PHI fields include but are not limited to:

- Names
- Social Security numbers
- Telephone numbers
- Geographic subdivisions smaller than a State, including street address, city, county, precinct, zip code, and their equivalent geocodes, census tracts and similar, geographic units
- Dates directly related to an individual, including birth date, admission date, discharge date, date of death; ages over 89 other elements of dates (including year) indicative of such age
- Fax numbers (though this will be exceedingly rare)
- Electronic mail addresses
- Medical record numbers
- Social Security numbers
- Health plan beneficiary numbers
- Account numbers
- Certificate/license numbers
- Vehicle identifiers and serial numbers, including license plate numbers
- Device identifiers and serial numbers
- Web Universal Resource Locations (URLs).
- Internet Protocol (IP) address numbers
- Biometric identifiers, including finger and voice prints
- Full face photographic images and any comparable images; and
- Other unique identifying number, characteristics or codes

Demographic PHI data listed above. Where available we will also include or incorporate sociodemographic variables and neighborhood characteristics.

Health encounter data including but not limited to: ICD-9 codes, dates of service (month and year), location of service, specialty of physician, pharmacy data). Claims data include dental, inpatient and outpatient encounters.

Electronic Health records including but not limited to: lab values, procedures, vital signs and flow sheet variables, assessments, clinic and inpatient notes, medications including dose and timing, symptoms, radiology images and reports etc.

Where available, we will also collect information on practice patterns including referral patterns and cost.

Where available, we will also incorporate behavioral data.

c) You are required to comply with University Policy that states that ALL electronic devices: computers (laptops and desktops; OFFICE or HOME); smart phones; tablets; external hard disks, USB drives, etc. that may hold identifiable participant data will be password protected, backed up, and encrypted. See http://med.stanford.edu/datasecurity/ for more information on the Data Security Policy and links to encrypt your devices.

Provide any additional information on ALL data security measures you are taking. You must use secure databases such as RedCap https://med.stanford.edu/researchit/infrastructure/redcap.html

https://med.stanford.edu/researchit/infrastructure/redcap.html. If you are unsure of the security of the system, check with your Department IT representative. Please see http://med.stanford.edu/irt/security/ for more information on IRT Information Security Services and http://www.stanford.edu/group/security/securecomputing/mobile_devices.html for more information for securing mobile computing devices. Additionally, any PHI data on paper must be secured in an locked environment.

By checking this box, You affirm the aforementioned. Y

For a detailed description, please see the Data Management Plan attached in Section 16.

Sensitive and restricted data sets are received by the Stanford Principal Investigator (PI) via encrypted disk, the mode preferred by the distributor of the data, and are saved to a secure server or GCP bucket that abides by Stanford's computer and network usage policy (http://adminguide.stanford.edu/62.pdf) and information security policy (http://adminguide.stanford.edu/63.pdf). On prem servers are behind a firewall with access restricted to authorized investigators only. Two step authentication is required at each sign on to access the data. Intrusion monitoring software is handled through the central university network office and locally as a backup. Data are encrypted and backed up nightly onsite for both onsite and offsite data storage. The server is housed in a secure location with very limited access. This location is monitored 24/7 via camera and has limited keycard access. In case of power failure both the server and network switches are tied into the natural gas generator to eliminate any down time and/or disruption in service.

For analytic data sets, the PHS data portal tracks: Project or study PI, project title, IRB approval date, date DUA was signed and received, variables included in the analytic data set and date data were made available. Data access is routinely audited to ensure that there is no inappropriate access.

d) Describe how data or specimens will be labeled (e.g. name, medical record number, study number, linked coding system) or de-identified. If you are de-identifying data or specimens, who will be responsible for the de-identification? If x-rays or other digital images are used, explain how and by whom the images will be de-identified.

Most data will come to PHS de-deidentified by the data vendor. In the case where PHS receives data with identifiers, where practical, PHS will construct deidentified analytic datasets for distribution.

This work will either be conducted manually or using state of the art tools designed to automate both data de-identification and assessment of re-identification risk.

e) Indicate who will have access to the data or specimens (e.g., research team, sponsors, consultants) and describe levels of access control (e.g., restricted access for certain persons or groups, access to linked data or specimens).

Only PHS Data Core, SRCC or other Data Admins and approved research teams will have access to the data.

f) If data or specimens will be coded, describe the method in which they will be coded so that study participants' identities cannot be readily ascertained from the code.

The data will be deidentified either by the data vendor or by PHS staff in whatever manner is appropriate for the dataset at hand. Data de-identification procedure described in Section 5.1 but will vary according to dataset type, research needs, available technologies and other factors.

g) If data or specimens will be coded, indicate who will maintain the key to the code and describe how it will be protected against unauthorized access.

In the event we have access to personal identifiers such as name or MRN and data are coded, the data manager(s) will hold the codes and the keys. No other individuals will have access to the codes.

For the most part, the data owners (eg, CMS) will hold the identities and data will arrive deidentified. The data owners will retain the code.

h) If you will be sharing data with others, describe how data will be transferred (e.g., courier, mail) or

transmitted (e.g., file transfer software, file sharing, email). If transmitted via electronic networks, describe how you will secure the data while in transit.See http://www.stanford.edu/group/security/securecomputing/. http://www.stanford.edu/group/security/securecomputing/. Additionally, if you will be using or sharing PHI see https://uit.stanford.edu/security/hipaa https://uit.stanford.edu/security/hipaa.

Please see the detailed description in the Data Management Plan attached in Section 16. In brief:

No person will be permitted to download data onto a computer, laptop or workstation unless there are special circumstances and permission has been granted in writing. All analytic data sets will be kept on the PHS server, the PHS Portal or Nero (either on prem or in Google). Investigators will access the data and analytic environment via a secure VPN or browser which requires an active authorized account and two-step authentication.

i) How will you educate research staff to ensure they take appropriate measures to protect the privacy of participants and the confidentiality of data or specimens collected (e.g. conscious of oral and written communications, conducting insurance billing, and maintaining paper and electronic data)?

All researchers are required to complete the PHS Data Security and other trainings as appropriate in order to access the data. Additionally, in order to obtain data sets, researchers will be required to sign a DUA and sign a statement indicating they understand they are not permitted to download data except in rare circumstances and with special permission.

12. Potential Conflict of Interest

Investigators are required to disclose any financial interests that reasonably appear to be related to this protocol.

Financial Interest Tasks

Investigators	Role	Email	Has Financial Interest?	Date Financial Interest Answered	Date OPACS Disclosure Submitted	Date OPACS Review Completed
Lesley Park	PD	lesley.park@stanford.edu	N	05/06/2019		

13. Consent Background

13. 1 Waiver of Consent Waiver of Consent

1) Y **The research involves no more than minimal risk to the participants.**

This data repository will collect and store datasets which have been collected either for other purposes (eg administrative or claims data) or for other research projects (eg HRS, Danish Biobank). We have gone to great lengths to protect patient confidentiality as outlined in our DMP in section 16.

2) Y **The waiver or alteration will not adversely affect the rights and welfare of the participants.**

Administrative or medical records may be used without patient their authorization if the use is approved by the IRB. Because study procedures are in place to protect confidentially (outlined in the DMP) information learned during the study will not affect the treatment of the participants and thus will not adversely affect their welfare.

Title : Secure Academic Data Commons - TEMPLATE for participating institutions.
Approval Period: **Draft**

3a) Y **The research could not practicably be carried out with out the requested waiver or alteration.**

3b) Y **For research using identifiable private information or identifiable biospecimens, the research could not practicably be carried out without using such information or biospecimens in an identifiable format.**

If the IRB required informed consent of participants, this research would be impracticable to do because it would require contacting hundreds of millions of patients and other individuals, many of whom are long dead. Additionally, in most cases Stanford PHS will not have access to identifiers, so this would be impossible. In the case where it was possible, identification and contact of patients would represent a greater violation of privacy and risks than using their de-identified data. 3b. In the case where data are delivered to the Data Core with identifiers, we remove them prior to making them available for research unless the investigator provides a compelling reason why a particular identifier is necessary.

4) Y **Whenever appropriate, the participants or legally authorized representatives will be provided with additional pertinent information after participation.**

It is not expected that the information learned from this these research studies will directly impact participants' treatment. Thus, it is not anticipated that there will be pertinent information for study participants. The study may lead to new knowledge that may affect the treatment of future patients. Additionally, in almost all cases, PHS will not have access to patient contact information, so contacting individuals would be impossible, unethical or both. That said, in the event of an important and actionable discovery, PHS will inform the data owner of the finding so that they may decide how to act upon the information. As PHS did not collect the data and, in most cases, does not have access to direct identifiers, direct contact would be inappropriate or take place under a completely separate protocol and agreement.

14. Assent Background (less than 18 years of age)

14. 1 Waiver of Assent Waiver of Assent

Address the following four regulatory criteria for a waiver of assent and provide a protocol-specific reasons for each:

Y **The research involves no more than minimal risk to the participants.**

This data repository will collect and store datasets which have been collected either for other purposes (eg administrative or claims data) or for other research projects (eg HRS, Danish Biobank). We have gone to great lengths to protect patient confidentiality as outlined in our DMP in section 16.

Y **The waiver will not adversely affect the rights and welfare of the participants.**

Administrative or medical records may be used without patient their authorization if the use is approved by the IRB. Because study procedures are in place to protect confidentially (outlined in the DMP) information learned during the study will not affect the treatment of the participants and thus will not adversely affect their welfare.

Y **The research could not practicably be carried out without the requested waiver.**

Y **For research using identifiable private information or identifiable biospecimens, the research could not practicably be carried out without using such information or biospecimens in an identifiable format**

If the IRB required informed consent of participants, this research would be impracticable to do because it would require contacting hundreds of millions of patients and other individuals, many of whom are long dead. Additionally, in most cases Stanford PHS will not have access to identifiers, so this would be impossible. In the case where it was possible, identification and contact of patients would represent a greater violation of privacy and risks than using their de-identified data.

In the case where data are delivered to the Data Core with identifiers, we remove them prior to making them available for research unless the investigator provides a compelling reason why a particular identifier is necessary.

Y **Whenever appropriate, the participants or legally authorized representatives will be provided with additional pertinent information after participation.**

Title : Secure Academic Data Commons - TEMPLATE for participating institutions.
Approval Period: **Draft**

It is not expected that the information learned from this these research studies will directly impact participants' treatment. Thus, it is not anticipated that there will be pertinent information for study participants. The study may lead to new knowledge that may affect the treatment of future patients. Additionally, in almost all cases, PHS will not have access to patient contact information, so contacting individuals would be impossible, unethical or both.

That said, in the event of an important and actionable discovery, PHS will inform the data owner of the finding so that they may decide how to act upon the information. As PHS did not collect the data and, in most cases, does not have access to direct identifiers, direct contact would be inappropriate or take place under a completely separate protocol and agreement.

15. HIPAA Background

15. 1 Waiver of Authorization waiver of hipaa authorization

a) **Describe the Protected Health Information (PHI) needed to conduct the research. PHI is health information linked to HIPAA identifiers. List BOTH health information AND HIPAA identifiers. If you are using STRIDE, use the Data Privacy Attestation to ensure that your request will match your IRB-approved protocol.**

As this is a data repository identifiable, or potentially identifiable, fields will vary widely by dataset. In many cases data will arrive at Stanford de-identified We expect databases may contain PHI and PII as described in Section 11b. In the event PHS will carry out de-identification, we have attached a detailed description of procedures to protect and de-identify data to prepare it for research in Section 5.2 of the DMP attached in Section 16. Identifiers are often necessary in order to conduct linkage and data enrichment. We will use tools such as Choice Maker or Datavant so that it will not be necessary to vend explicit identifiers to researchers in most cases. These tools append a scrambled code to records that allow linkage without inclusion of sensitive fields. In the event an investigator needs to use identifiers (eg, dates of service, zip code or census tract etc) they must complete a separate IRB and justify the need for these fields.

b) Please Answer:

Y **Do you certify that the use or disclosure of protected health information involves no more than a minimal risk to the privacy of individuals?**

Y **Do you certify that the research could not practically be conducted with out the waiver?**

Y **Do you certify that you have adequate written assurances that the protected health information will not be reused or disclosed to any other person or entity, except as required by law, for authorized oversight of the research project, or for other research for which the use or disclosure of protected health information would be permitted?**

Y **Do you certify that the research could not practically be conducted with out access to and use of the protected health information?**

c) **Please describe an adequate plan to protect any identifiers from improper use and disclosure.**

We have attached a detailed description of our security and de-identification strategy in the DMP attached in Section 16.

d) **Please describe an adequate plan to destroy the identifiers at the earliest opportunity consistent with conduct of the research, unless there is a health or research justification for retaining the identifiers or such retention is otherwise required by law.**

Plans to destroy the data vary by dataset and agreement with the data owners or generators. As datasets often represents a considerable investment, and destruction of identifiers makes further linkage impossible, for the most part, we will elect to strongly protect and limit access to identifiers rather than destroy them outright. These procedures are detailed in the DMP attached in Section 16.

Title : Secure Academic Data Commons - TEMPLATE for participating institutions.
Approval Period: **Draft**

16. Attachments

Attachment Name	Attached Date	Attached By	Submitted Date
5_Stanford_DMP_General_01Mar19	05/06/2019	itaylor	
7_DUA_Individual_15Apr19	05/06/2019	itaylor	

Obligations

The Protocol Director agrees to:

- Adhere to principles of sound scientific research designed to yield valid results
- Conduct the study according to the protocol approved by the IRB
- Be appropriately qualified to conduct the research and be trained in Human Research protection, ethical principles, regulations, policies and procedures
- Ensure all Stanford research personnel are adequately trained and supervised
- Ensure that the rights and welfare of participants are protected including privacy and confidentiality of data
- Ensure that, when de-identified materials are obtained for research purposes, no attempt will be made to re-identify them.
- Disclose to the appropriate entities any potential conflict of interest
- Report promptly any new information, modification, or unanticipated problems that raise risks to participants or others
- Apply relevant professional standards.

Any change in the research protocol must be submitted to the IRB for review prior to the implementation of such change. Any complications in participants or evidence of increase in the original estimate of risk should be reported at once to the IRB before continuing with the project. Inasmuch as the Institutional Review Board (IRB) includes faculty, staff, legal counsel, public members, and students, protocols should be written in language that can be understood by all Panel members. The investigators must inform the participants of any significant new knowledge obtained during the course of the research.

IRB approval of any project is for a maximum period of one year. For continuing projects and activities, it is the responsibility of the investigator(s) to resubmit the project to the IRB for review and re-approval prior to the end of the approval period. A Notice to Renew Protocol is sent to the Protocol Director 7 weeks prior to the expiration date of the protocol.

Department Chair must approve faculty and staff research that is not part of a sponsored project. VA applicants must have Division Chief or Ward Supervisor approval. E-mail the Department Chair approval to IRBCoordinator@lists.stanford.edu.

All data including signed consent form documents must be retained for a minimum of three years past the completion of the research. Additional requirements may be imposed by your funding agency, your department, or other entities. (Policy on Retention of and Access to Research Data, Research Policy Handbook, http://doresearch.stanford.edu/policies/research-policy-handbook/conduct-research/retention-and-access-research-data)

PLEASE NOTE: List all items (verbatim) that you want to be reflected in your approval letter (e.g.,

Title : Secure Academic Data Commons - TEMPLATE for participating institutions.
Approval Period: **Draft**

Amendment, Investigator's Brochure, consent form(s), advertisement, etc.) in the box below. Include number and date when appropriate.

Protocol, DMP, DUA, attachements.

Y By checking this box, I verify that I, as the Protocol Director (PD) responsible for this research protocol, have read and agree to abide by the above obligations, or that I have been delegated authority by the PD to certify that the PD has read and agrees to abide by the above obligations.