

Guidelines for Use of the Shanghai Cohort Study Data and/or Biorepository

External Collaborators

A. Submitting a Research Project Proposal

1. Any investigator wishing to develop a collaboration with the Shanghai Cohort Study (SCS) investigators to use SCS survey data and/or biospecimens should submit a detailed proposal via the SCS website. The format of the proposal is described in detail below.
2. The proposal's format is similar to an NIH grant (i.e., specific aims, background and significance, preliminary studies, and methods) but should be no longer than FOUR pages in length. The proposal should also include required survey data variables, and type and amount of biospecimens, if biospecimens are requested, and the justification for using the SCS resources.
3. Additional Considerations when biospecimens are requested

The reasons for proposing use of the SCS biospecimens, rather than another source, must be clearly described. Given a limited amount available, the biospecimens will be used only for analyses where other, less precious, biospecimen collections (e.g., retrospectively collected biospecimens) cannot provide adequate data. The assessment of markers of disease prognosis will generally not be considered as appropriate in the use of the SCS biorepository. In addition, proposals to evaluate highly speculative hypotheses are not considered appropriate and will not be approved. Finally, laboratory analyses which are either already funded or have been proposed by SCS investigators will not be considered for approval.

Before an approval is granted, the following issues must be addressed to confirm that a particular association can be reasonably evaluated using SCS biorepository samples.

- (a) Appropriateness of using serum samples and plain urine samples: The laboratory needs to confirm that these types of samples are routinely accepted for the analysis of interest, otherwise a pilot study will need to be conducted to establish the validity of assay performance.
- (b) Laboratory assay: All assays must be conducted using the best available technology to ensure the precision and accuracy of the measurement. The serum volume required is minimized, and the assay reproducibility is maximized. The definition of "acceptable" biospecimen volume will be determined on a study-by-study basis and will depend in large part on the importance/priority of the study hypothesis. In the proposal, the applicant should clearly describe the various assay methods currently available and their rationale for choosing the specific assay(s).
- (c) Reproducibility of the laboratory assay: The laboratory conducting the analyses must be able to conduct the assay with a high degree of precision (i.e., low coefficient of variation or high reliability coefficient). This information must be obtained through a blinded evaluation of the laboratory. Unfortunately, coefficients of variation provided by laboratory investigators are not

sufficient, as, in our experience, these data do not always reflect the true magnitude of laboratory error. The evaluation must be recent and, if at all possible, should have been performed by the same technician who will be conducting the assay on the proposed study samples.

(d) Range of the biomarker in the SCS population: For many biomarkers of interest, knowledge of a usual range in an adult population will be sufficient. In such a case, a usual range and how this range was determined (e.g., in what population and what method used) should be briefly described. However, for certain assays, where the range of proposed biomarkers may vary substantially across different populations, a population-specific distribution of biomarker levels should be provided. If necessary, a pilot study to determine biomarkers levels in the SCS may be conducted prior to final approval for the proposed project.

(e) Stability of the biomarker over time (i.e., how well does a single measure reflect long-term blood levels). In the SCS cohort overall, only one-time point blood and urine sample per participant have been collected at baseline. Thus, data must be available demonstrating that a single measurement at one point of time will provide a sufficiently integrated value of the biomarker reflecting a relatively long-term exposure (generally the exposure of interest with chronic diseases), in which an association between the biomarker and disease risk could reasonably be detected, if indeed one exists. If these data are not already available, applicants should consider conducting a pilot study to assess stability over a minimum of a 6-month period.

4. Study proposals will be reviewed within about four weeks of proposal submission. A decision of acceptance, pending for revision, and rejection will be emailed to the applicant. For either of the latter two outcomes, reasons for the decision will be provided. A “pending for revision” will be given if the proposal has considerable scientific merit, yet one or more issues need to be addressed before the project can proceed. The revised proposals will receive an expedite review. For proposals that will require the funding outside of the applicant's institution, the approval process must be factored into the timing of any grant application submission. The SCS Investigators cannot take responsibility for missed deadlines.

B. Conducting Studies Using the SCS Survey Data and/or Biorepository

1. The exact nature and scope of the project must be described in a written collaborative agreement and signed by the external collaborator, the primary SCS investigator, and a representative from each investigator's institution. Use of data from the SCS is limited to the defined, specific project for which the approval was granted. If further research or analytic activities develop from the originally proposed project, the external collaborator must obtain appropriate approval for such activities by submitting a proposal amendment. In signing the collaborative agreement, external collaborators also will be confirming that they have read these guidelines and both understand and agree to comply with them.

2. Since no funds have been allocated to manage the development of these outside collaborative projects, all costs must be borne by the outside applicant's institution. Unless the initial development and review of the proposal requires substantial data exploration to determine feasibility, it is not anticipated that this cost would exceed \$2,000/proposal. The actual cost will be based on the time required of an SCS investigator and programmer to

determine approximate numbers of cases and the distribution of relevant exposure variables in the targeted study population.

3. Outside collaborators must provide a draft of any grant proposal (e.g., NIH grant) to the primary SCS investigator at least one month prior to the application due date. This will allow the SCS investigator an opportunity to provide feedback and will allow time to obtain any additional data for the grant application. Failure to meet this deadline will result in delay of submission. The primary SCS investigator will provide a letter of support to the external investigator to be included in the grant application that indicates SCS interest in collaboration with the outside investigator.

4. Study costs

- (a) External collaborators must provide funds to cover the cost of initial programming needed to identify cases and to produce the distributions of relevant exposure variables.
- (b) The cost of all pilot studies required to determine the feasibility and validity of the proposed project must be assumed by the potential external collaborator.
- (c) At least one SCS investigator may be included as a co-investigator (with appropriate time commitment) on any grant proposal where use of SCS data is proposed. The level of effort will vary according to the size and complexity of the project, but will usually be 5% to 10% FTE per year.
- (d) To ensure integrity and safety of SCS data, it is the policy of the SCS that no original data with patient's identification number leave the University of Pittsburgh. Limited de-identified dataset will be provided to outside collaborators who can perform the proposed study. Not all variables can be included in the limited dataset because of the prohibition of the confidentiality and patient privacy law. In the case of complex data analysis and required certain variables that are not allowed to be included in the analytic dataset, the statistical analyses for such projects will be conducted at the SCS Data Management and Statistical Core (DMSC). Analysis plans will be drawn up by the outside collaborator in conjunction with the primary SCS investigator; these plans will be given to the SCS DMSC statistician who will oversee all analyses.

The outside applicant should have funds to cover SCS personnel effort (and other costs) associated with preparing data sets and conducting statistical analyses. Upon approval of the Request, SCS staff will provide the applicant with an estimate of the hours and costs required to carry out the work. Routine requests requiring less than 10 hours to fulfill, such as the preparation of limited data sets, simple tables displaying tabulations of study data, or straightforward statistical computations may be handled as a free service. If 10 or more hours are required to complete the request, the applicant will be either requested to provide funds to support an appropriate amount of time for SCS personnel (such as statisticians or/and programmers, lab technician to identify and pull biospecimens from biorepository) or billed according to the following fee schedule:

Fee Schedule for Requests	
< 10 hours	No charge for routine requests
10-40 hours	Flat rate of \$4000

> 40 hours	\$100/hour
------------	------------

If desired, prior to submitting the Request, the Applicant may contact the primary SCS investigator to estimate the anticipated effort needed to complete a Request.

- (e) The arrangement for payments will be made through formal subcontracts with the University of Pittsburgh, in which full overhead as approved by NIH will be considered a direct cost to the proposing institution cost base.

5. For a proposed study using a nested case-control study design, the SCS statistician will generate case-control matched sets based on the SCS established matching criteria. If the specific matching criteria are required, the applicant should consult with the SCS primary investigator to be certain that sufficient number of eligible controls are available within the cohort. The final matched case-control sets should be reviewed and approved by the primary SCS investigator, Director of the SCS DMSC, the PI of the proposed project.

6. To the extent possible, all analyses will be conducted as a single batch with appropriate masked QC samples added to the batch. If, as is frequently the case, a large number of samples are being assayed in a study, the precision of the assay must be monitored on an ongoing basis using masked QC samples. Results from these QC samples must be reported on a batch-by-batch basis to Director of SCS DMSC who will be responsible for monitoring the performance of the assay.

7. A proposed timeline for completion of biospecimen aliquoting and preparation of the analytic dataset should be discussed prior to submission of any grant. All projects need to be completed within the constraints of the current SCS infrastructure. It is very difficult, if not impossible, for the SCS research team to increase staff member for a particular project with a short deadline without substantial increase of expenditure on resources. Thus at the beginning of a project, external collaborators should review with the SCS to establish a proposed schedule for project completion.

8. The external collaborator must agree to keep the SCS investigators updated on the progress of the study by providing either a written or verbal report at least every year. Failure to adhere to a reasonable progress schedule could lead to termination of the collaborative relationship with no further data tables or additional analyses provided.

9. Human Subjects Considerations

- (a) It is PI's obligation to comply with the human subject protection policies and to obtain appropriate approval from institute of the leading Investigator's Human Research Protection Program prior to implementation.
- (b) As analyses of genetic susceptibility to disease are associated with complex ethical considerations, a full discussion of the ethical implications of these analyses must be part of the initial proposal. Investigators should be aware that analyses that identify human subjects at very high risk of disease are particularly problematic in this regard.

C. Data Analysis and Publication Issues

1. The external collaborating investigator should forward all laboratory results to the SCS. All primary data sets of laboratory results will be maintained in the DMHS database.
2. Typically, all data analyses will be conducted at the SCS DMSC. The most efficient way for these analyses to be accomplished will be for the outside investigator and the collaborating SCS investigator to agree upon the analysis plan in advance. The external collaborating investigator will provide to the statistician a set of data analysis requests and an outline of result tables that indicate how the results are to be presented. The SCS DMSC will proceed to complete the analyses and return the completed tables to the collaborating investigator. The SCS investigator will work as needed with the statistician to execute the analysis plan.
3. In the case that data analyses are conducted at the external collaborating investigator's institution, the external investigator should forward all study results to the SCS DMSC where they will be carefully reviewed and signed off by the DMSC Director before they are included in grant application as preliminary results, in progress report, and in final manuscript for publication.
4. At least one member of the SCS investigative team may be a coauthor on any manuscript resulting from this collaboration and, as such, will need to sign off on any manuscript prior to its submission for publication. This will take the form of a brief note indicating review and approval of the final manuscript by the SCS investigator. External investigators should plan on the entire process taking at least 4 weeks (and longer if there are issues to be resolved concerning analysis or interpretation of the data). Any initial presentation of these data at meetings also must receive sign-off from the designated SCS investigator(s).
5. Any dispute regarding data interpretation may be brought to the Advisory Committee for consideration. Where appropriate, the Advisory Committee will seek additional consultation from independent experts. Since the Advisory Committee meets as a group only once per year, considerable delay for a resolution could occur. Final decisions rest with Dr. Yuan, the Principal Investigator of SCS, in consultation with the Advisory Committee.