Guidelines for Use of Survey and Clinical Data Collected in the Shanghai Men's Health Study

A. Submitting a Research Project Proposal

1. Any investigator wishing to use data collected in the Shanghai Men's Health Study (SMHS) should first submit a detailed proposal via the SMHS 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 variables and justification for using the SMHS resources. The reasons for proposing use of SMHS data must be clearly described.

3. Study proposals will be reviewed within four weeks of proposal submission. A decision to accept, accept pending revisions, or reject a proposal will be emailed to the applicant. The applicant may submit a revised application if the original application is not accepted. Arrangements will be made to provide an expedited review of a revised proposal.

B. Conducting Studies Using the SMHS Resources

1. The exact nature and scope of the project must be described in the application. Use of data (or other covariate data) from the SMHS is limited to the defined, specific project for which the approval has been obtained. Research activities beyond the scope of the original project must obtain appropriate approval by submitting a proposal amendment before implementation. In submitting a research application, the applicant also confirms that he/she has read these guidelines ("Guidelines for use of the Shanghai Men's Health Study data") and both understood and agreed to comply with them.

2. Applicants may be asked to provide funds related to the development and review of a research application if the application requires significant data exploration to determine study feasibility and assess study design. The actual cost will be based on the time required of an SMHS investigator and programmer to determine approximate case numbers that might be considered appropriate for the proposed analyses and related exposure distributions.

3. Study Costs

(a) The cost of all pilot studies required to determine the feasibility and validity of the proposed project must be assumed by the applicant.

(b) To ensure integrity of SMHS data, in general, no original data will be sent to outside investigators. Secondly, because of the complexity of the database and the SMHS investigators' knowledge of the strengths and limitations of these data, substantial input is required by SMHS investigators to ensure both valid and maximal use of the available data. For these reasons, the SMHS Data Management and Statistical Core (DMSC) will perform the analyses for all outside investigators. Analysis plans will be drawn up by the outside investigators in conjunction with a SMHS investigator; these plans will be given to a DMSC statistician who will oversee all analyses. With reasonable justification, exceptions can be made to provide a subset of SMHS data, typically on a list of well-defined exposure variables for pooling projects, such as consortium studies.

The applicant should have funds to cover SMHS personnel effort (and other costs) associated with preparing data sets and conducting data analyses. Upon approval of the request, SMHS staff will provide the applicant with an estimate of the hours and costs required to carry out the work. The applicant will either be required to provide funds to support an appropriate amount of time for SMHS personnel (such as statisticians or/and programmers) or be billed at $125 per hour.

If desired, prior to submitting the request, the applicant may contact the SMHS team to estimate the anticipated effort needed to complete a request.

4. Human Subjects Considerations

(a) It is the applicant's obligation to comply with the human subject protection policies and to obtain appropriate approval from the applicant institute'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 men at very high risk of disease are particularly problematic in this regard.

5. The programs used for data analysis must be carefully reviewed and signed off on by an SMHS epidemiologist and statistician in addition to the study programmer and the external applicant. Importantly, the sign off must be by an SMHS investigator who understands how the cases and population for analysis are being defined, is familiar with SMHS variable definitions, and can understand the code generated by the programmer.

6. The applicant must agree to keep the SMHS 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 approved study.

C. Data Analysis and Publication Issues

1. The applicant should forward all analysis results to the SMHS.

2. All data analyses will be conducted at Vanderbilt by DMSC of the SMHS (see section B.3.b above). The most efficient way for these analyses to be accomplished is for the outside investigator and the SMHS team to agree on the analysis plan in advance (to whatever extent possible). The external investigator will provide to the DMSC statistician a set of data analysis requests and an outline template of result tables that indicates how the results will be presented. The DMSC of the SMHS will proceed to complete the analyses and return the completed tables to the investigator. In completing the analysis plan, the SMHS investigator also will work as needed with the statistician in supervising the SMHS programmer assigned to the project.

3. Any manuscript prepared using SMHS data should be reviewed by the SMHS team prior to its submission for publication. External investigators should plan on the entire process taking at least four 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 SMHS team.

4. Any dispute regarding data interpretation may be brought to the External Advisory Board for consideration. Where appropriate, the External Advisory Board will seek additional consultation from

independent experts. Since the External Advisory Board meets as a group only once per year, considerable delay in coming to a resolution could occur. Therefore, it behooves all investigators to work closely with the SMHS team in resolving any dispute. Final decisions rest with Dr. Shu, the SMHS Principal Investigator, in consultation with the External Advisory Board.

Guidelines for Use of Biological Samples Collected in the Shanghai Men's Health Study

A. Submitting a Proposal

1. Any investigator wishing to use data collected in the Shanghai Men's Health Study (SMHS) should first submit a detailed research proposal. 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 variables, biospecimen type and amount, assay methods, laboratory name, justification for using the SMHS resources, and timeline. The reasons for proposing use of the SMHS biospecimens, rather than another source, must be clearly described. The SMHS biorepository is a unique and finite resource. The assessment of markers of disease prognosis will generally not be considered an appropriate use of the SMHS biorepository. In addition, proposals to evaluate highly speculative hypotheses are not considered appropriate and will not be approved.

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

(a) Appropriateness of using blood samples collected in EDTA and urine samples spiked with vitamin C. SMHS blood samples were collected using EDTA as the anticoagulant, and urine samples were spiked with vitamin C. The laboratory needs to confirm that these samples are appropriate for the analysis of interest; otherwise, a pilot study will need to be conducted to establish that EDTA or vitamin C will not interfere with assay performance.

(b) Laboratory assay to be used. All assays must be conducted using the best available technology to ensure that the appropriate parameter is assayed, the volume of sample required is minimized, and the assay reproducibility is maximized. The definition of "acceptable" sample 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 be clear in describing the various assay methods currently available and their rationale for using the specific assay being proposed.

(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 possible, should have been performed by the same technicians who will be conducting the study analyses.

(d) Range of the biomarker in the SMHS cohort. For many biomarkers of interest, knowledge of a usual range in an adult population will be sufficient (e.g., plasma antioxidant levels); in this instance, the usual range and how this range was determined (i.e., in what population) should be briefly described. However, for other assays, where the range may vary substantially by population (e.g., plasma levels of DDE/PCBs), a population-specific distribution of biomarker levels should be provided. If necessary, a pilot study to determine levels observed in the SMHS may need to be conducted prior to receiving final approval for conducting a project.

(e) Stability of the biomarker over time (i.e., how well a single measure reflects long-term exposure levels). In the SMHS cohort overall, only one blood and urine sample per participant was collected. Thus, data must be available to indicate that a single measurement provides a sufficiently integrated measure of longer-term exposure (generally the exposure of interest with chronic diseases) so that an association between the biomarker and disease could reasonably be detected, if it indeed exists. If these data are not already available, applicants should consider conducting a pilot study to assess the stability.

4. It is anticipated that the decision to accept, accept pending revisions, or reject a proposal will be made within four weeks of proposal submission. The applicant may submit a revised application if the original application is not accepted. Arrangements will be made to provide an expedited review of a revised proposal.

B. Conducting Studies Using the SMHS Biorepository

1. The exact nature and scope of the project must be described in the application. Use of biomarker data (or other covariate data) from the SMHS is limited to the defined, specific project for which the SMHS approval has been obtained. Research activities beyond the scope of the original project must obtain appropriate approval before implementation. In submitting a research application, the applicant also confirms that he/she has read these guidelines ("Guidelines for Use of Biological Samples") and both understood and agreed to comply with them.

2. Applicants may be asked to provide funds related to the development and review of a research application if the application requires data exploration to determine study feasibility and assess study design. The actual cost will be based on the time required of an SMHS investigator and programmer to determine approximate case and control numbers that might be considered appropriate for the proposed analyses and related exposure distributions.

3. Every effort will be made to preserve biological samples collected in the SMHS so that these precious samples can be used for future studies. Multiplex assay technologies will be used when appropriate to preserve samples.

4. Study Costs

(a) Applicants must provide funds to cover the cost of retrieving, aliquoting and shipping of specimens, data entry of results, and for additional freezer space (necessitated by the aliquoting of samples). Funds also must be provided for the initial programming needed to identify case and control samples, prepare data sets, and perform statistical analyses (if appropriate) (see fee schedules in 4(d) below).

(b) In addition to the cost of aliquots of the study samples, funds must be available to prepare quality control specimens to be analyzed along with the study samples (in approximately a 1:10 ratio).

(c) The cost of all pilot studies required to determine the feasibility and validity of the proposed project may be assumed by the applicant.

(d) To ensure integrity of SMHS data, it is the general policy of the SMHS that no data leave Vanderbilt University Medical Center. Secondly, because of the complexity of the database and the SMHS investigators' knowledge of the strengths and limitations of these data, substantial input is required by SMHS investigators to ensure both valid and maximal use of the available data. For these reasons, all data analyses will be carried out by the Data Management and Statistical Core (DMSC) of the SMHS. The outside investigators in conjunction with a SMHS investigator will draw up analysis plans; these

plans will be given to a DMSC statistician who will oversee all analyses. With reasonable justification, exceptions can be made to provide a subset of SMHS data, typically on a list of well-defined exposure variables for pooling projects, such as consortium studies.

The applicant should have funds to cover SMHS personnel effort (and other costs) associated with preparing datasets and biological samples, as well as conducting data analyses. Upon approval of the request, SMHS staff will provide the applicant with an estimate of the hours and costs required to carry out the work. The applicant will either be required to provide funds to support an appropriate amount of time for SMHS personnel (such as statisticians, programmers, and/or lab personnel) or billed at $125 per hour. Costs for biological sample retrieval, aliquot and preparation will be determined using the fee schedule approved by the VUMC.

If desired, prior to submitting the request, the applicant may contact the SMHS team to estimate the anticipated effort needed to complete a request.

(e) The arrangement for payments will be made through the Vanderbilt University Medical Center in which an appropriate overhead rate will be used to estimate the total cost for the study.

5. Human Subject Considerations

(a) It is the applicant's obligation to comply with human subject protection policies and to obtain appropriate approval from the applicant institute'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 men at very high risk of disease are particularly problematic in this regard.

6. Before aliquoting of samples begins, the programs used to generate cases and controls must be carefully reviewed and signed off on by an SMHS epidemiologist and DMSC statistician in addition to the study programmer and the external applicant investigator. Importantly, the sign-off must be by an SMHS investigator who understands how the cases and controls are being defined, is familiar with SMHS variable definitions, and can understand the code generated by the programmer. Laboratory assay of the wrong cases or controls because of errors in their initial identification can be very expensive and would waste a precious resource.

7. 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 the SMHS investigator who will be responsible for monitoring reproducibility.

8. The applicant must agree to keep the SMHS investigators updated on the progress of the study by providing either a written or verbal report at least annually. Failure to adhere to a reasonable progress schedule could lead to termination of the approved study.

9. Any biological sample remaining after the completion of the approved laboratory assays must be returned promptly to the SMHS sample archive.

C. Data Analysis and Publication Issues

1. The applicant should forward all laboratory results to the SMHS. All primary data sets of laboratory results will be maintained in the SMHS database to be used by other investigators in future studies.

2. All data analyses will be conducted at Vanderbilt by DMSC of the SMHS (see section B.4.b above). The most efficient way for these analyses to be accomplished will be for the outside investigator and the SMHS team to agree on the analysis plan in advance (to whatever extent possible). Once the laboratory assays are complete and results sent to the SMHS, the external investigator will provide to the DMSC statistician a set of data analysis requests and a template of result tables that indicates how the results are to be presented. The SMHS DMSC will complete the analyses and return the completed tables to the outside investigator. In completing the analysis plan, the SMHS investigator will also work as needed with the DMSC statistician in supervising the SMHS programmer assigned to the project.

3. Any manuscript prepared using SMHS data should be reviewed by the SMHS team prior to its submission for publication. External investigators should plan on the entire process taking at least four 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 must also receive a sign-off from the SMHS team.

4. Any dispute regarding data interpretation may be brought to the External Advisory Board for consideration. Where appropriate, the External Advisory Board will seek additional consultation from independent experts. Since the External Advisory Board meets as a group only once per year, considerable delay in coming to a resolution could occur. Therefore, it behooves all investigators to work closely with the SMHS team in resolving any dispute. Final decisions rest with Dr. Shu, the SMHS Principal Investigator, in consultation with the External Advisory Board.