|  |  |  |
| --- | --- | --- |
| **Saw Swee Hock School of Public Health**  **Data Management Unit** | **SOP:** | **SSHSPH-DM-B001-16-V3.1** |
| **Version:** | **V3.2** |
| **Access to SPHS Data and Samples** | **Implementation Date:** | **13/08/2018** |

**CONDUCTING RESEARCH USING DATA AND SAMPLES FROM THE SINGAPORE POPULATION HEALTH STUDIES (SPHS)**

The Singapore Population Health Studies (SPHS) is a population-based health research initiative of the Saw Swee Hock School of Public Health (SSHSPH) to facilitate high-quality scientific research on the Singapore population. Investigators can request datasets and biological samples from SPHS for their research projects with the appropriate IRB approvals. To ensure that results presented in publications can be independently replicated and to avoid repetition of analyses, investigators will have to adhere to the following procedures for use of these data and samples:

1. Complete the Data & Samples Request Form (Annex A) and submit it together with a Data Analysis Proposal (Annex B; if purpose is research) or a Data Use Proposal (Annex C; if purpose is operational/programmatic) to the Data Management Unit (DMU), SSHSPH, at [SSHSPHDataRequest@nus.edu.sg](mailto:SSHSPHDataRequest@nus.edu.sg).
2. Considering the request the DMU will generate a quotation of charges for the services described in the data request and inform the requestor for approval.
3. A reference number will be assigned to the proposal which will then be reviewed by the SPHS Scientific Committee. Members of this committee are:

* Associate Professor Jeannette Lee
* Associate Professor Rob van Dam
* Professor Tai E-Shyong
* Professor Teo Yik Ying (for genetics studies)

1. DMU will inform the requestor of the outcome of the review. Once the proposal has been approved, the requestor must execute the [Material Transfer Agreement / Data Access Agreement] before the requested data and/or samples will be provided to the requestor.
2. If additional data variables or subjects, or additional samples are required during the course of the research project, the requestor may specify these in Section D, “Additional data/samples request,” of the Data & Samples Request Form and submit it to the DMU. The DMU may request a new request form be submitted if required.
3. Upon approval of the proposal, the requestor may begin analysis of the provided data and/or samples and write a manuscript. During this period, the SPHS Scientific Committee, through the DMU, may request an update of the progress of the research.
4. Before submitting the manuscript to a journal, the requestor should:
   1. have a co-author (generally the second author) conduct a technical review of the manuscript (see Annex D);
   2. preferably, have an independent person check the code of your statistical program. At least, provide descriptions in your code that clarify the different steps of the analysis to an independent investigator; and finally,
   3. submit to the DMU for review by a member of the SPHS Scientific Committee:
      1. the Manuscript Submission Form (see Annex E);
      2. the manuscript that has been given a technical review; and
      3. the statistical program code.

**ANNEX A. DATA & SAMPLES REQUEST FORM**

|  |  |
| --- | --- |
| **For official use** | |
| Request received date: | Request number: |
| **Section A: Requestor details** | |
| Date of request: Click or tap to enter a date. | Full name: |
| Department / organisation: | Designation: |
| Phone number: | Email address: |
| **Section B: Details of request** | |
| B1) Purpose of this request:  Research, including preliminary studies. *Please attach a Data Analysis Proposal (Annex B)*  Operational / programmatic (go to B4). *Please attach a Data Use Proposal (Annex C)* | |
| B2) Title of research study: | |
| B3) IRB approval reference:  *Attach a copy of the IRB approval letter and all protocol amendment approval letter(s) that list the approved documents and the latest approved IRB application which describes the research methodology/protocol.* | |
| B4) Requested delivery date: Click or tap to enter a date.  *Kindly note that the date entered should have a minimum of 6 working weeks interval from the date of request, excluding weekends and public holidays.* | |
| B5) What is being requested:  Samples (please also complete Section C)  Individually identified data, i.e. data that contains information that identifies the research participant  Key-coded individual, record-level data, i.e. data containing surrogate identifiers in place of information that identifies the research participant  Non-identifiable individual, record-level data, i.e. data will not be linkable to the original dataset (go to B7 if this is the only type of data requested)  Aggregate data (go to B7 if this is the only type of data requested)  Others, please specify: | |
| B6) Select the subject consent conditions that are required for your request:  Not applicable  consent for samples to be used for genetic research  consent for commercial development of research on the samples  consent not to be notified of incidental findings  consent to contact subject for public health research  other, please specify: | |
| B7) Does this request relate to any previous data or samples that you have requested from DMU?  No, this is a first request/this has no connection with previous request(s).  Yes. *Please state request form reference number and specify how they are related:* | |
| B8) State the measures you will undertake to ensure the security of the data requested for. | |
| **Section C: Specifications of samples (omit this section if samples are not requested for)** | |
| C1)   |  |  |  |  |  |  | | --- | --- | --- | --- | --- | --- | | Type of samples | Fasting status (Yes, No or N/A) | Number of samples (subjects) | Volume (ml) per subject | Other specifications (e.g. DNA concentration) | Tests that will be performed on the samples | | DNA |  |  |  |  |  | | EDTA Plasma |  |  |  |  |  | | Citrate Plasma |  |  |  |  |  | | Serum |  |  |  |  |  | | RBC |  |  |  |  |  | | Whole blood |  |  |  |  |  | | PBMC |  |  |  |  |  | | Clot |  |  |  |  |  | | Urine (buffered) |  |  |  |  |  | | Urine (neat) |  |  |  |  |  |   C2) Describe how the samples obtained will be stored. | |
| C3) What will happen to the samples after your research and analysis have been completed? | |
| **Section D: Acknowledgment by requestor** | |
| I certify that all the information provided above is true and accurate.  \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_  Signature of requestor & date | |
| **Section E: Additional data / samples request** | |
| Indicate the request number that was assigned to the original proposal:  *State the reason for the additional data and/or samples here and attach your revised data analysis proposal with the revisions marked out/highlighted.*  \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_  Signature of requestor & date | |
| **For official use** | |
| Additional request received date: Click or tap to enter a date. | |

**ANNEX B. DATA ANALYSIS PROPOSAL**

To be submitted to the Data Management Unit, Saw Swee Hock School of Public Health, at [SSHSPHDataRequest@nus.edu.sg](mailto:SSHSPHDataRequest@nus.edu.sg). An example of an analysis proposal is provided below.

The analysis proposal should be about 2-3 page-long and consist of the following:

1. Title
2. Proposed authors (can be alphabetical) but should identify first author and senior author
3. Research questions
4. Brief scientific background (~200 words)
5. Study design
6. Study population
   1. Eligible population
   2. Exclusions
7. Outcome variables
8. Exposure variables
9. Covariables
   1. Potential confounders considered
   2. Potential mediators
   3. Potential effect modifiers
10. Statistical analyses
    1. Primary analyses
    2. Secondary analyses
11. Please list the data source and all the data variables (as according to the data dictionary) which you wish to request for this proposal here:

***EXAMPLE OF A DATA ANALYSIS PROPOSAL***

**Research title:** Can body fat distribution, adiponectin and C-reactive protein explain ethnic differences in insulin resistance?

**First author:** Gao He

**Proposed co-authors (alphabetical):** Jeannette Lee, Rob M. van Dam, Salome A. Rebello, Tai E Shyong

**Research question:** In an Asian context, we will use path analysis to evaluate to what extent body fatness, adiponectin levels, C-reactive protein and their interconnections mediate the relation between ethnicity and insulin resistance.

**Background:**

South Asians have long been identified in the comparison with Europeans to have a higher risk of insulin resistance [1]; however, so far there is no study that has examined the difference in insulin resistance between ethnic groups in Asia and potential mediators of the relation between ethnicity and insulin resistance. Body fatness and body fat distribution as well as adipokines and inflammation may contribute to ethnic differences in insulin resistance.

Adiponectin level has been shown to be inversely related to risk of type 2 diabetes across diverse populations [2] and it has a protective role against insulin resistance [3]. There is evidence that South Asians have lower adiponectin level than Europeans [4]. In addition, there is a positive association between state of inflammation and insulin resistance and South Asians have been shown to have significantly higher CRP levels than do Europeans [5,6].

**Study Design:** Cross-sectional study

**Study population:** Eligible population: Singapore prospective study program (SP2)

Exclusion criteria:

* people with diabetes
* people with cardiovascular diseases
* people of other ethnicity than Chinese, Malay and Indian
* people with missing data for the key variables involved in the analysis

**Outcome variables:** Insulin resistance by HOMA-IR

**Exposure variables:** Ethnicity (Chinese, Malay, Indian)

**Covariables:**

Potential confounders

* Age (years)
* Sex (male=1, female=0)
* Smoking status (non-, ex-, light smokers, heavy smokers)
* Physical activity (kcal/day)
* Alcohol consumption (non-drinker, light drinker, heavy drinker)

Potential intermediates:

* [Body fatness measurements]
  + BMI (kg/m2) derived from height and weight
  + Waist circumference (cm)
  + Hip circumference (cm)
  + Waist-hip ratio
* [Adipokines]
  + Total adiponectin level (ug/ml)
  + High molecular weight adiponectin level (ug/ml)
  + HMW adiponectin/Total adiponectin ratio
* [Inflammation marker]
  + C-reactive protein level (mg/l)

**Statistical analyses:**

* 1. Examine the distribution of each continuous variable by histogram, log-transform those with skewed distribution
  2. Tabulation of baseline characteristics across 3 ethnic groups using proportions for categorical variables, means/geometric means and standard deviation for continuous variables
  3. ANOVA and/or pair-wise ethnic comparison of the characteristics between ethnic groups with necessary adjustment for multiple comparison
  4. As a preliminary step, validate literature-suggested bivariate relationships among variables by correlations (Pearson’s r, Spearman’s r), pay close attention to those highly correlated; and then use multiple linear regression analysis to see whether the effect of a proximal variable on a distal variable is mediated by a mediator
  5. Integrate previous findings and theoretical rationales into a tentative model and conduct path analysis using Pathreg package in STATA for each path to get path coefficients (standardized regression coefficients/beta weights)
  6. Refine the model by dropping paths with non-significant path coefficients and compare it with our hypothetic model
  7. Run path analysis on the refined model and examine the path coefficients to conclude the mediating effects and potential causal relationship

**References**

1. McKeigue PM, Shah B, Marmot MG. Relation of central obesity and insulin resistance with high diabetes prevalence and cardiovascular risk in South Asians. *Lancet*. 1991;337:382-386.

2. Li S, Shin HJ, Ding EL, van Dam RM. Adiponectin levels and risk of type 2 diabetes: a systematic review and meta-analysis. *JAMA*. 2009;302:179-188.

3. Garaulet M, Hernández-Morante JJ, de Heredia FP, Tébar FJ. Adiponectin, the controversial hormone. *Public Health Nutr*. 2007;10:1145-1150.

4. Martin M, Palaniappan LP, Kwan AC, Reaven GM, Reaven PD. Ethnic differences in the relationship between adiponectin and insulin sensitivity in South Asian and Caucasian women. *Diabetes Care*. 2008;31:798-801.

5. Misra A. C-reactive protein in young individuals: problems and implications for Asian Indians. *Nutrition*. 2004;20:478-481.

6. Chambers JC, Kooner JS. Diabetes, insulin resistance and vascular disease among Indian Asians and Europeans. *Semin Vasc Med*. 2002;2:199-214.

**ANNEX C. DATA USE PROPOSAL**

To be submitted to the Data Management Unit, Saw Swee Hock School of Public Health, at [SSHSPHDataRequest@nus.edu.sg](mailto:SSHSPHDataRequest@nus.edu.sg).

1. Please describe the use of the data you wish to request:
2. Please list the data source and all the data variables (as according to the data dictionary) which you wish to request:

**ANNEX D. TECHNICAL REVIEW**

For quality control, we request the conduct of a technical review.

The purpose of the technical review is to have a second person look at all of the data presented in the manuscript tables and text, checking for consistency and plausibility.

Examples include:

* Do the numbers in the tables and for the exclusions add up?
* Are the ranges of variables consistent with other publications using the SP2 cohort data?
* Are the numbers cited in the text the same as those presented in the Table and the Abstract?
* Do the table / figure numbers in the text refer to the right table/ figure and do the reference numbers in the text refer to the right citation.
* Was the output from the statistical programs accurately copied into the tables and text.

The technical reviewer must be a co-author other than the first author or person who did the computer analyses. Usually, the second author is assigned the responsibility of doing the technical review.

**ANNEX E. MANUSCRIPT SUBMISSION FORM**

To be submitted to the Data Management Unit, Saw Swee Hock School of Public Health, at [SSHSPHDataRequest@nus.edu.sg](mailto:SSHSPHDataRequest@nus.edu.sg).

1. Manuscript title:
2. Authors:
3. Corresponding author:
4. Name of person who performed the technical review:
5. Reference number (provided when analysis plan is approved):
6. Version and date of SPHS dataset used:
7. Journal to be submitted to:
8. Attach the following documents/files:
   1. Manuscript to be submitted to the journal
   2. Statistical program code