# University of Wisconsin-Oshkosh
# Institutional Review Board (IRB)

| STANDARD OPERATING PROCEDURES | |
|---|---|
| SOP Number: 5<br>Effective Date: November 14, 2013<br>Last Reviewed: April 5, 2016<br>Attachment: Page 1 of 5<br>Note: Previously SOP #2<br>Prepared By: Anca Miron & Kelly Schill | Title:<br><br>Policy on Analysis of Secondary Datasets |

This policy applies only to activities that involve the secondary analysis of existing data, such as public data sets, medical records, student records, data collected from previous studies including audio/video recordings, etc. that were initially collected for another purpose.

**A. What is research involving human subjects and when does analysis of secondary data sets require IRB review?**

**In order for research to constitute "human subjects" research under federal law (45 CFR 46.102):**

- The data used in the research must have been obtained by the investigator through intervention or interaction with the subject of the data, OR
- The data must include identifiable private information about the subject of the data. Private information is information which has been provided for specific purposes by an individual and which the individual can reasonably expect will not be made public (e.g., a medical or school record).
- For permission to analyze existing data that meets the above definition of human subjects' research, please complete an **IRB Existing Human Subjects Data Form**.

**When is the secondary use of existing data exempt under Category 4?**

- When the investigators initially have access to identifiable private information but extract the data needed for the research in such a way that the information can no longer be connected to the identity of the subjects, then the protocol is considered human subjects research and needs IRB review. This protocol involves research activities involving human subjects that may be exempt from the requirements of the Federal Policy for the Protection of Human Subjects (45 CFR 46) under Category 4 of the federal regulations. In order to qualify for an exempt determination, an **IRB Existing Human Subjects Data Form** must be submitted to IRB for determination. The following criteria must be met for data sets to qualify for exempt status:
    a. The extracted data set should not include direct identifiers (names, social security numbers, addresses, phone numbers, etc.) or indirect identifiers (codes or pseudonyms that are linked to the subject's identity) and
    b. It must not be possible to identify subjects by combining a number of characteristics (e.g., date of birth, gender, position, and place of employment).

  The following do not qualify for exemption: research involving prisoners, research involving protected health information (PHI) from HIPAA-covered entities, and FDA-regulated research.
- Examples of potentially exempt projects:
    a. A researcher conducts a study of treatment outcomes that involves the review of patient charts at a medical facility. The researcher records patient age, sex, diagnosis, and treatment outcome in such a way that the information cannot be linked back to the patient. The project could potentially qualify for an exemption but this determination will be made by the IRB.
    b. A student will be given access to data from her faculty advisor's survey research project. The data consists of coded survey responses, and the advisor will retain a key that would link the data to identifiers. The student will extract the information she needs for her project without including any identifying information and without retaining the code. The use of the data does constitute research with human subjects because the initial data set is identifiable

(through a coding system); depending on its level of risk, this protocol could qualify for exempt or non-exempt status.

**When is the secondary use of existing data non-exempt?**

- If secondary analysis of existing data does involve research with human subjects and does not qualify for exempt status as explained above, the project must be reviewed either through expedited procedures or by a full (convened) Board, and an **IRB Existing Human Subjects Data Form** must be submitted to the IRB for review.
- Researchers using data previously collected under another study should consider whether the currently proposed research is a "compatible use" with what subjects agreed to in the original consent form. For non-exempt projects, a consent process description or justification for a waiver must be included in the research protocol. The IRB may require that informed consent for secondary analysis be obtained from subjects whose data will be accessed. Alternatively, the IRB can consider a request for a waiver of one or more elements of informed consent under 45 CFR 46.116(d). In order to approve such waiver, the IRB must first be satisfied that the research:
    a. presents minimal risk (no risks of harm, considering probability and magnitude, greater than those ordinarily encountered in daily life or during the performance of routine examinations or tests); and
    b. the waiver or alteration will not adversely affect the rights and welfare of the subjects; and
    c. the research could not practicably be carried out without the waiver or alteration; and
    d. whenever appropriate, the subjects will be provided with additional pertinent information after participation.
- Restricted Use Data: Certain agencies and research organizations release files to researchers with specific restrictions regarding their use and storage. The records frequently contain identifiers or extensive variables that combined might enable identification, even though this is not the intent of the researcher. Research using these data sets most often requires non-exempt level review.
- Examples of non-exempt projects:
    a. A student is applying to the National Center for Health Statistics for use of data from the National Health and Nutrition Examination Survey that includes geographic identifiers and date of examination. The analysis of this restricted use data would require non-exempt review by IRB.
    b. A student will be given access to coded mental health assessments from his faculty advisor's research project. The student plans to analyze the data with a code attached to each record, and the advisor will retain a key to the code that would link the data to identifiers. The use of the data does constitute research with human subjects and does not qualify for exempt status since subjects can be identified. This student project would require an application to be submitted for non-exempt review by the IRB or the faculty advisor may wish to submit an amendment application to add the student and his/her work to the original protocol.

**B. When does the secondary use of existing data NOT require IRB review?**

**In general, the secondary analysis of existing data only requires IRB review when it falls within the regulatory definition of research involving human subjects.** The use of data from the following **list of IRB approved public data sets** is not considered human subject research *as long as the following two criteria are met*:

- Research will NOT involve merging any of the data sets in such a way that individuals might be identified
- The researcher will NOT enhance the public data set with identifiable, or potentially identifiable data

Example of data sets that do not require IRB review:

a. Public use data sets (such as portions of U.S. Census data, etc.) are data sets prepared with the intent of making them available for the public. The data available to the public are not individually identifiable and therefore their analysis would not involve human subjects. Information that contains identifiers and can be accessed freely by the public (without special permission or application) is not "private" and the research therefore does not involve human subjects. For example, a study involving only analysis of the

published salaries and benefits of public university presidents would not need IRB review since this information is not private.

**Before beginning research, investigators who wish to have a specific data set or data archive considered for inclusion on the IRB approved list should complete and submit the Public Data Set Nomination Form to irb@uwosh.edu, in order to have that specific data set or data archive approved.**

UW Oshkosh, October 2013/Policy based on http://cphs.berkeley.edu/secondarydata.pdf

**List of Approved De-identified Publicly Available Datasets**

Below is a list of publicly available datasets that meet the requirements of the UW Oshkosh's Existing Datasets Policy:

- Autism Brain Imaging Data Exchange (ABIDE)
- Better Access to Data for Global Interdisciplinary Research (BADGIR)
- Inter-University Consortium for Political and Social Research (ICPSR)
- National Center for Health Statistics
- National Center for Education Statistics
- National Election Studies
- Roper Center for Public Opinion Research
- University of Wisconsin-Madison Data and Information Services Center (DISC)
- U.S. Bureau of the Census

☐ U.S. Bureau of Labor Statistics

- Luxembourg Income Study (LIS)
- Integrated Public Use Microdata Series (IPUMS-International)
- Integrated Public Use Microdata Series (IPUMS-USA)
- Integrated Public Use Microdata Series (IPUMS-CPS)
- Medical Expenditure Panel Survey (MEPS- Household Component)

☐ American Hospital Association Annual Survey

- American College of Surgeons National Trauma Data Bank (NTDB)
- American College of Surgeons National Surgical Quality Improvement Program( ACS-NSQIP) Participant Use Data File

☐ Panel Study of Income Dynamics (PSID)

☐ Comprehensive Hospital Abstract Reporting System (CHARS; public data only)

☐ Demographic Health Survey (DHS), Standard and Interim Surveys

☐ Fatality Analysis Reporting System (FARS)

- ☐ Hospital Compare
- ☐ Medicare Healthcare Cost Report Information System (HCRIS)
- ☐ National Automotive Sampling System (NASS) General Estimates System (GES)
    - Behavioral Risk Factor Surveillance System (BRFSS; public data only)
    - British Household Panel Survey
    - Childhood Cancer Survivor Study (CCSS)
    - German Socio-Economic Panel Survey
    - Healthcare Cost and Utilization Project (H-CUP) healthcare databases
    - HIV Prevention Trials Network D01: Vaccine Preparedness Study/Uninfected Protocol Cohort – 4 files
    - Medicare Physician Supplier Procedure Summary Master File
    - National Child Development Study
    - National Collegiate Athletics Association (NCAA) Injury Surveillance Program (ISP)
    - National Epidemiologic Survey on Alcohol and Related Conditions (NESARC)- Wave 1 & Wave 2
    - National Hospital Ambulatory Medical Care Survey (NHAMCS)
    - National Longitudinal Survey (NLSY)
    - National Survey of Children's Health (NSCH)
    - National Survey of Children with Special Health Care Needs (NS- CSHCN)
    - Survey of Consumer Finances (SCF)

*Note:* If your research merges more than one of these datasets in such a way that individuals may be identified, the research is not covered by this policy and requires prior IRB approval.

## INSTITUTIONAL REVIEW BOARD (IRB)

### PUBLIC DATA SET NOMINATION FORM

A list of IRB approved public data sets that are not considered human subject research is available at the following link: www.uwosh.edu/grants/forms/

Use of data from the IRB approved public data sets list is not considered human subject research as long as the following criteria are met:

- Research will NOT involve merging any of the data sets in such a way that individuals might be identified
- Researcher will not enhance the public data set with identifiable, or potentially identifiable information

**Instructions for completing this form:**

Please complete this form if you wish to have a specific data set considered by the IRB for inclusion in the public data sets list. Please submit to irb@uwosh.edu. You will receive a confirmation from the IRB regarding the status of the nomination form.

**Name:** _____

**Department/Division:** _

**Data Set Information:**

| | |
|---|---|
| **1. Name of data set:** | |
| **2. Brief description of the host archive or institution:** | |
| **3. List the URL for the data set:** | |
| **4. URL documenting public availability of the data set:** | |
| **5. Description of how the data set may be obtained by researchers. Include any conditions, stipulations, data use agreements, etc. :** | |
| **6. Description of how the data were collected, from/about whom, who collected the data, and when it was collected:** | |
| **7. Description of any conditions or restrictions about the use of the data:** | |

**IRB Use Only Below This Line**

**Date:** _____ **Status:** ☐ Approved for Public Data Set ☐ Not Approved for Public Data Set

**Added to Public Data Set List (name and date):**