

## Guidance for Using Secondary/Archival Data

Although projects that only involve secondary data analysis (i.e., data that has been collected previously) do not involve interactions or interventions with humans, they may meet the federal definition of human subjects research (HSR). Thus, the IRB recommends researchers submit their application even when they will use existing data or contact the Office of Research Integrity at [orihelp@bsu.edu](mailto:orihelp@bsu.edu) for additional assistance.

### 1. When do researchers need to submit the IRB application (or contact the IRB)?

The IRB requests that researchers who conduct studies using secondary data including de-identified, publicly available data not listed below in item #3 and coded private information seek a determination from the IRB as to whether or not the study constitutes human subjects research.

The IRB will determine 1) that the project does not constitute human subjects research (not HSR) or 2) that the project needs the IRB review (exempt, expedited, or full board).

PIs will be asked to provide copies of data use or security agreements required by data holder, a description of data security and access procedures, or a permission letter from the data holder, if applicable.

### 2. What IRB needs to know?

In order for the IRB to assess the risks to the participants through the use of existing data sources and make recommendations for ethical use of the data, the following information should be provided. The list is not exclusive.

- How did you obtain access to the data?
- Who provides the data sets?
- What do the data consist of (i.e., provide a list of the variables in the data set) ?
- How many records will you access?
- Will the data be combined with other data sources?
- Can the participants be linked to their data?
- If de-identified data or coded data will be used, who prepares the data?

### 3. When does secondary use of existing data NOT require IRB review?

The analysis of de-identified, publicly available data does not constitute human subjects research as defined at 45 CFR 46.102, and thus, it may not require IRB review. The data sets listed below do not require BSU IRB review if the datasets are publicly available and de-identified.

- [Inter-University Consortium for Political and Social Research \(ICPSR\)](#)
- [National Center for Health Statistics](#)
- [National Center for Education Statistics](#)
- [National Election Studies](#)
- [U.S. Bureau of the Census](#)

**Note:** Not all datasets from the sites above do require the IRB review. If they are restricted data which the data provider requires IRB review and if the data sets are merged with other data, they should be reviewed by the IRB.

### 4. When is the secondary use of existing data exempt?

Research involving secondary data analysis can be exempted under Category 4 of the federal regulations if the sources of such data are publicly available or the information is recorded by the investigator in such a manner that the dataset contains no information that can identify participants directly or through identifiers linked to the participants. Data must be de-identified before any analysis is conducted in order to qualify under exempt category 4.

The following do not qualify for exemption.

- Research involving prisoners
- Research involving protected health information from HIPAA-covered entities

- FDA-regulated research.

**5. When does the secondary use of existing data require expedited or full board review?**

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. Generally, research involving the analysis of non-publicly available data that contains private identifiable information about living individuals is considered to constitute human subjects research that is not exempt from 45 CFR 46.

**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.

**Consent:** 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.