Public Data Sets

Projects that only involve secondary data analysis of data sets with identifiable information but do not involve interactions or interventions with humans, still require IRB review, because the definition of “human subjects” at 45 CFR 46.102(f) includes living individuals about whom an investigator obtains identifiable private information for research purposes.

Projects that use de-identified data sets only, and do not involve interactions or interventions with humans, would not require IRB review; however, only the IRB may make that determination.

Secondary analysis of restricted data (e.g. data that require a data use agreement, obtain permission to access, etc.) requires IRB review.

If a research study involves secondary analysis of publicly available data (de-identified) and utilizes additional procedures/methods that do meet the criteria of human research, the entire project must be submitted for IRB review.

The use of public use data (de-identified) will require IRB review if:

- Research involves merging of any of the data sets in such a way that individuals might be identified, or
- The researcher will expand the public data sets with identifiable, or potentially identifiable data.