PUBLIC DATA SETS

Unit: Human Research Protections Program (HRPP), Office of Research

Applies to: Faculty and Staff involved in human research

PURPOSE:
Research using certain publicly available data sets may not constitute human subjects research because the data contained within these specific data sets do not meet the federal definition of "human subject."

DEFINITIONS:
Public Data Sets: Collections of data that have been made publicly available.
- x   Includes files consisting of medical and/or research data prepared with the intent of public use.
- x   Includes data files that have been reviewed under the jurisdiction of an IRB with the intent of making them available for public use. The federal government has responsibility for the review of federal statistical data collections.
- x   Does not include data shared informally among colleagues.

Publicly Available: Material is widely available, even when:
- x   Identifiers are included, so long as individuals whose data is represented would not reasonably expect daeun Tf imEOn made publis.
- xgred forobntaiking ths dats.
  - xe data islimitred tg resed
    - cmadmlic/or resea(chtafr)61(bilactionha.)5.7s a.

  Retrdidedt; Nedt conwidared publicly elailab43.8(e.)]TJ /TT7 1 Tf -0 .5743 -1.2186 TD 0 Tc 0 Tw <0078>Tj /TT
            aer tdoseaeetrdictions.
        d byl govereontal genclie,r resta(chorganizdat65(i-5.5(ne,r)]TJ T* [(andothier.)-1( )]TJ -3.2787 -1
          s AesuApplirf myt have theaibiliys oivihe iffearent

            spetsosearer d.5(D Tht) [(spets d)5.5eprohesuAv

If a data set is not featured on this list, Investigators should submit an IRB application via eBridge. If Investigators wishes to propose that a data set be added to this list, see instructions below under "Adding to the list of data sets considered public."

HRPP Staff Review
1. Investigators should submit an IRB application in eBridge if any of the following conditions apply:
    a. The data set is not listed on the IRB website as a public data set.
    b. The data set is listed on the IRB website as a public data set but a third party requires documentation from the IRB confirming that the use of this data does not constitute human subjects research.
    c. Data from multiple public data sets are merged, as this may alter the nature of one or both sets (e.g. two de-identified data sets could be merged to create an identifiable data set).
    d. Investigator seeks to enhance a public data set with identifiable or potentially identifiable data.
    e. Data sets which are not publicly available are used alongside publicly available data sets.
    f. Activities interacting or intervening with human subjects are proposed in addition to use of publicly available data sets.
    g. Using controlled-access data sets such as NIH dbGaP data repository.
    h. The investigator plans to create a public data set by preparing files with the intent of making them available for public use.

2. The data supplier may require the researcher or researcher's institution to sign a Data Use Agreement that explicitly requires IRB approval or a determination of exempt status. If IRB review is required:
    a. Complete and submit an IRB application in eBridge.
    b. Upload a copy of the Data Use Agreement to the application. An unexecuted copy is acceptable.

Adding to the list of data sets considered public
1. A data set may be nominated for inclusion on the MCW List of Data Sets Considered Public by completing and submitting the IRB Form: Public Data Set Nomination Checklist.
        a. The IRB does not include data sets that directly or indirectly identify subjects on this list.
2. The nominated data set and documentation provided to the MCW IRB will be reviewed by the HRPP Office.
3. Written notification of the review outcome will be provided to the individual who submitted the nomination request.
4. If the data set is found eligible, the data set will be added to the MCW List of Data Sets Considered Public.
5. If a data set is found ineligible, individuals will receive written notification along with instructions to submit an IRB application for review.

REFERENCES:
N/A

SUPPORTING DOCUMENTS:
MCW HRPP Website> Public Data Sets
IRB Form: Public Data Set Nomination Checklist

Effective Date: 07/01/2023
Version number: 3.0
Previous Version/date: 2.0, 06/15/2018
Responsible Office: HRPP Office
Approval Date: 05/29/2023

Approved By
HRPP Authorized Official: Ryan Spellecy, PhD, Director, HRPP
                            Human Research Protections Program (HRPP)
                            Office of Research
                            Medical College of Wisconsin