

California Cancer Registry Data Disclosure Guidelines for Secondary Data Sharing

This document describes guidelines to fulfill the California Cancer Registry¹ (CCR) requirements for obtaining CCR data for studies participating in secondary data sharing. When applied, these guidelines allow the Principal Investigator (PI) or the equivalent of the research study to use CCR data to promote scientific research while complying with Health and Safety Code Sections 103875 – 103885.

1. Definition of Terms

Disclosure: Disclosure of CCR data is defined as the granting of the right to examine the data and to create or retain a copy for the use of the institution. Only data necessary for the stated purpose of the request may be disclosed. The data may be used only for the approved purpose.

Secondary Data Sharing: Further sharing of data with a third party, including external investigators, pooling projects, and/or repositories that allow further access.

Controlled Access Repository: At a minimum, a controlled-access repository should meet the following requirements: (1) Relies on appropriate standards to deidentify data (pursuant to HIPAA 45 CFR 164.514(b)(2)(i)), including using random, unique Subject IDs. (2) Reviews data access requests. (3) Controls/audits dataset access and download. (4) Requires a data use agreement with repository recipients that prohibits further data sharing and includes requirements for maintaining confidentiality and destroying data when the project is completed.

De-identified Data: Per the Health Insurance Portability and Accountability Act's (HIPAA) Safe Harbor provision (45 CF 164.514(b)(2)(i)), de-identified data **excludes** 18 identifiers ([Methods for De-identification of PHI | HHS.gov](#)).

1. Names
2. All geographic subdivisions smaller than a state, including street address, city, county, precinct, ZIP code, and their equivalent geocodes, except for the initial three digits of the ZIP code if, according to the current publicly available data from the Bureau of the Census:
 - (a) The geographic unit formed by combining all ZIP codes with the same three initial digits contains more than 20,000 people; and
 - (b) The initial three digits of a ZIP code for all such geographic units containing 20,000 or fewer people are changed to 000
3. All elements of dates (except year) for dates that are directly related to an individual, including birth date, admission date, discharge date, death date, and all

¹ CCR is California's statewide, population-based cancer surveillance system and is a statutorily mandated program per Health and Safety Code Sections 103875 – 103885. On October 13, 2023, Governor Newsom signed Senate Bill (SB) 344 into law (SB 344 Chaptered), authored by the California Legislature. This legislation makes amendments to the Health and Safety Code § 103885 and went into effect on January 1, 2024.

ages over 89 and all elements of dates (including year) indicative of such age, except that such ages and elements may be aggregated into a single category of age 90 or older.

4. Telephone numbers
5. Vehicle identifiers and serial numbers, including license plate numbers
6. Fax numbers
7. Device identifiers and serial numbers
8. Email addresses
9. Web Universal Resource Locators (URLs)
10. Social security numbers
11. Internet Protocol (IP) addresses
12. Medical record numbers
13. Biometric identifiers, including finger and voice prints
14. Health plan beneficiary numbers
15. Full-face photographs and any comparable images
16. Account numbers
17. Any other unique identifying number, characteristic, or code
18. Certificate/license numbers

2. Guidelines Applicable to Secondary Sharing of CCR Data

- A. Any disclosure of CCR Data shall include only the information necessary for the stated purpose of the requested disclosure, used for the approved purpose, and not be further disclosed, unless the researcher complies with CCR's secondary data sharing guidelines.
- B. Cases of secondary data sharing are to occur under the direction and control of the originally approved data recipient, who remains the researcher responsible for data security and integrity.
- C. Secondary sharing of CCR data is only allowed if the data are de-identified.
- D. If a research study is required to share data per the terms of their funding agreement, and/or if the investigator intends to make the data shareable in the future, the data sharing plan should be detailed in the CPHS and CCR applications, and the two must be consistent. Further, researchers should consider the sensitivity of each variable being re-disclosed and determine whether data transformations are required to ensure the data cannot be used to identify or reidentify the data with an individual person². The minimally necessary rule should always be considered.

² Please refer to [California Health and Human Services De-identification Guidelines \[PDF\]](#) for further information on ensuring that any data shared is sufficiently de-identified and aggregated to prevent the identification of individual patients.

- E.** Data may be shared via controlled access repositories or other investigator-initiated ways of sharing de-identified data. Investigator-initiated ways of sharing de-identified data require, at a minimum, a data use agreement that prohibits further data sharing and includes requirements for maintaining confidentiality and destroying data when the project is completed.
- F.** When CCR data are linked with external datasets (for example, hospital discharge files, Medicare or Medi-Cal claims, or survey data), the entire linked dataset must be de-identified in accordance with applicable de-identification standards before any secondary sharing occurs.
- G.** Secondary sharing of CCR data will not be tracked by CCR.
- H.** Please note that vital statistics data cannot be redisclosed without prior approval from the Research and Analytics Branch, Center for Health Statistics and Informatics, California Department of Public Health ([Research and Analytics Branch \(RAB\)](#)). For questions related to vital statistics data, contact HIRS@cdph.ca.gov.