Last updated: 22 March 2024
Compiled by the NAACCR Data Security & Confidentiality Workgroup

NAACCR Data Release Guidelines

Editors:

Lauren Maniscalco, Louisiana Tumor Registry

Bozena M. Morawski, Cancer Data Registry of Idaho

Carolyn Bancroft, Maine Cancer Registry

Jean-Michel Billette, Canadian Cancer Registry – Statistics Canada

David Chesnut, Information Management Services, Inc.

Castine Clerkin, North American Association of Central Cancer Registries

Steven Friedman, National Cancer Institute, National Institutes of Health

Susan Gershman, Massachusetts Cancer Registry

Tabassum Insaf, New York State Cancer Registry

Selina Khatun, Nunavut Cancer Registry

Robert McLaughlin, Cancer Registry of Greater California

Carmina Ng, Canadian Cancer Registry – Statistics Canada

Recinda Sherman, North American Association of Central Cancer Registries

Anshu Shrestha, Registry of Greater California

Qianru Wu, Nebraska Cancer Registry

Heather Zimmerman, Montana Central Tumor Registry

Publication Date: March 22, 2024[1]

[1]These guidelines shall be reviewed and updated every 12 months.

Table of Contents

Last updated: 22 March 2024
Compiled by the NAACCR Data Security & Confidentiality Workgroup

Key Terminology

Term	Definition	Reference(s)
Common Rule (and 2018 Common Rule)	Published in 1991, HHS regulations for the protection of human subjects in research are codified at 45 CFR 46. These regulations are informed by the ethical principles of respect for persons, beneficence, and justice as described in the 1978 Belmont Report and reflect the consensus and agreement by 15 federal departments to the regulatory structure previously established by HHS and the FDA under their respective authorities. Subpart A, or "the Common Rule," provides a set of protections for research subjects. In 2018, a revision to the Common Rule went into effect. These changes aimed to reduce unnecessary burden on researchers and enhance protections for human subjects.	https://www.hhs.gov/ohrp/regulations-and-policy/regulations/45-cfr-46/index.html https://www-nejm-org/doi/10.1056/NEJMp1700736
Covered Entity	Under the Heath Information Portability and Accountability Act (HIPAA), an individual, organization, or agency that transmits protected health information (PHI) electronically for certain transactions related to healthcare. Nota bene: most (but not all) central registries are not covered entities under HIPAA, but use HIPAA standards as part of their business operations.	
Aggregated data	Data that collapse differences between and/or combines information from multiple people or tumors into a summary table or set of fields. Summary measures such as counts or rates can be provided as a data set or displayed in a table. Aggregated data do not include line-level data, even if aggregation over multiple variables results in a single record or count for a particular category, e.g. age group.	
De-Identified Data	HIPAA regulations define health information as de-identified if: • the data do not directly identify an individual and; • the covered entity has no reasonable basis to believe the data can be used to identify an individual.	45 CFR 164.514(a) Standard: De-identification of protected health information; (b) Implementation specifications: Requirements for de-identification of protected health information. https://www.hhs.gov/hipaa/for-professionals/privacy/special-topics/de-identification/index.html

Term	Definition	Reference(s)
Limited Data Set	Per HIPAA Privacy Rule, a data set containing identifiable health-related information that a Covered Entities is permitted to share for authorized purposes (such as research purposes, public health activities, and healthcare operations) without obtaining the prior authorization of the patient to whom the data relate, provided certain conditions are met.	45 CFR 164.514(e)(1) Standard: Limited data set
Confidential Data	Information subject to restrictions of access and use, which are intended to protect the identity of and/or information about persons. Confidential data are typically inclusive of data that allow for the identification of an individual person, and extend to private matters such as health, finances, kinship, genetics, criminality, etc. Confidential data may directly or indirectly identify a person. Examples of the former include name, address, social security number (SSN), or health card number (HCN). Examples of the latter include place of treatment or physician or patient census tract of residence.	For cancer registry data, see state statutes.
Personally Identifiable Information (PII)	Information that can be used to identify an individual, e.g., their name, Medicare Number, social security number (SSN), address, alone or in combination with other personal or identifying information.	45 CFR 160.103 Definitions. See “individually identifiable health information”. https://www.cms.gov/privacy
Protected Health Information (PHI)	Information related to an individual’s health or a medical condition, provision of health care to the individual, or payment for the provision of health care to the individual, and that identifies the individual or for which there is a reasonable basis to believe that it can be used to identify the individual. PHI is created or received by a health care provider, health plan, employer, or health care clearinghouse and the regulatory definition does not extend to self-reported information as is commonly obtained from participants in epidemiologic survey research.	45 CFR 160.103 Definitions. https://www.hhs.gov/hipaa/for-professionals/privacy/special-topics/de-identification/index.html

Term	Definition	Reference(s)
Privacy	"An individual's desire to limit the disclosure of personal information."	National Research Council (US) Committee on Maintaining Privacy and Security in Health Care Applications of the National Information Infrastructure. For the Record Protecting Electronic Health Information. Washington (DC): National Academies Press (US); 1997. Available from: https://www.ncbi.nlm.nih.gov/books/NBK233429/ doi: 10.17226/5595
Confidentiality	"A condition in which information is shared or released in a controlled manner. Organizations develop confidentiality policies to codify their rules for controlling the release of personal information in an effort to protect patient privacy."	National Research Council (US) Committee on Maintaining Privacy and Security in Health Care Applications of the National Information Infrastructure. For the Record Protecting Electronic Health Information. Washington (DC): National Academies Press (US); 1997. Available from: https://www.ncbi.nlm.nih.gov/books/NBK233429/ doi: 10.17226/5595
Security and Data Security	"[A] number of measures that organizations implement to protect information and systems. It includes efforts not only to maintain the confidentiality of information, but also to ensure the integrity and availability of that information and the information systems used to access it." "Data security is the process of making sure data are available only to those who need to use it for a legitimate purpose. Controlling access to data helps ensure privacy and is required by various federal agency policies and regulations."	National Research Council (US) Committee on Maintaining Privacy and Security in Health Care Applications of the National Information Infrastructure. For the Record Protecting Electronic Health Information. Washington (DC): National Academies Press (US); 1997. Available from: https://www.ncbi.nlm.nih.gov/books/NBK233429/ doi: 10.17226/5595 https://www.cdc.gov/cancer/npcr/tools/security/index.htm
Registry Data	Data collected by the central cancer registry as part of routine cancer registry operations.	

Term	Definition	Reference(s)
Augmented Data or Dataset	Registry data combined with other data collected by a third party. For example, cancer registry data linked with national cohort data would be considered augmented data.	https://www.naaccr.org/wp-content/uploads/2023/05/VPR-DUA_V4-Final-Clean_4.4.23-2.docx https://www.naaccr.org/wp-content/uploads/2023/02/Cancer-Registry-Secondary-Data-Sharing-Fact-Sheet_Final_2.9.23.pdf
Central IRB and Central IRB Processes	A Central Institutional Review Board (IRB) is an oversight committee that reviews, approves, and provides regulatory oversight for multi-site research protocols involving human subjects. A central IRB is often referred to as a "single IRB" because it serves as the sole body of oversight across the sites in multi-site research. The relying sites retain ultimate responsibility for the research activities performed locally in connection with the regulatory oversite of the Central IRB. A central (or single) IRB review process involves an agreement under which relying study sites in a multicenter study agree to the review of an IRB ***other*** than the IRB affiliated with the research site. This process, which is required (with some exceptions) for any U.S. institution engaged in cooperative research, is intended to achieve a more streamlined, efficient review process than local review at each participating research site.	https://www.hhs.gov/ohrp/regulations-and-policy/requests-for-comments/draft-guidance-use-single-institutional-review-board-for-cooperative-research/index.html
Engagement in Research…		https://www.hhs.gov/ohrp/regulations-and-policy/guidance/guidance-on-engagement-of-institutions/index.html
When Institutions are Engaged in Research	In accordance with guidance of the federal Office of Human Research Protections (OHRP), institutions are considered engaged in human subjects research when the involvement of their employees/agents in that project includes any of the following: 1. Receiving an award through a grant, contract, or cooperative agreement directly from HHS for the non-exempt human subjects research (i.e. awardee institutions), even where all activities	

Term	Definition	Reference(s)
	involving human subjects are carried out by employees or agents of another institution. 2. Institutions whose employees or agents intervene for research purposes with any human subjects of the research by performing invasive or noninvasive procedures. 3. Institutions whose employees or agents intervene for research purposes with any human subject of the research by manipulating the environment. 4. Institutions whose employees or agents interact for research purposes with any human subject of the research. (Examples of interacting include engaging in protocol dictated communication or interpersonal contact; asking someone to provide a specimen by voiding or spitting into a specimen container; and conducting research interviews or administering questionnaires.) 5. Institutions whose employees or agents obtain the informed consent of human subjects for the research. 6. Institutions whose employees or agents **obtain** for research purposes identifiable private information or identifiable biological specimens **from any source** for the research.	
When Institutions are **not** Engaged in Research	OHRP guidance provides that institutions would be considered not engaged in a human subjects research project if the involvement of their employees or agents in that project is limited to the scenarios described by HHS: https://www.hhs.gov/ohrp/regulations-and-policy/guidance/guidance-on-engagement-of-institutions/index.html The scenarios listed in the above link describe the types of institutional involvement that would make an institution not engaged in human subjects research. Situations commonly relevant to cancer registries include:* 1. Institutions whose employees or agents perform commercial or other services for investigators provided that all of the following conditions also are met:	

Term	Definition	Reference(s)
	a. the services performed do not merit professional recognition or publication privileges; b. the services performed are typically performed by those institutions for non-research purposes; and c. the institution's employees or agents do not administer any study intervention being tested or evaluated under the protocol. 2. Institutions whose employees or agents release to investigators at another institution identifiable private information or identifiable biological specimens pertaining to the subjects of the research. **There may be additional such scenarios that are relevant to your registry.*	
Active Consent	In the context of research using cancer registry data, the recruitment of participants often necessitates prior access to identifiable information about potential participants. Depending on local requirements, recruitment communications can be initiated through either the permission of the registry, and/or the physician source of the potential participant's cancer case record. Active Consent-Physicians: The Active consent requirement means an affirmative response from physician during a specific time period (such as a recruitment period) will allow for inclusion of patient data in the potential cohort of study participants whose PHI/PII will be released to researchers. Active Consent-Patients: In contrast to recruitment processes that allow for direct, initial contact with potential research participants by a researcher, "Active Consent" means the registry requires an affirmative, prior response from an individual about whom data in the cancer registry may be of interest to researchers. The active consent authorizes the registry to release that individual's PHI/PII to a research team for further screening, informed consent/assent, and enrollment in a research study.	

Term	Definition	Reference(s)
Passive Consent	Passive Consent-Physicians: Reversing the communication requirement associated with Active Consent, Passive Consent involves, in the context of patient contact studies, a process in which a registry provides physicians with a defined amount of time to contact the registry and indicate that a patient **should not be contacted** as a potential participant in a research study. In the absence of a communication from the physician, the registry will include an individual's data in the PHI/PII released to researchers for screening, informed consent/assent, and enrollment in a research study, or may move on to contacting the patient to obtain their active or passive consent. Passive Consent-Patients: In the context of patient contact studies, Passive Consent means a process in which a registry notifies an individual and provides a specific timeframe for response in which the potential participant can indicate that they **do not want** to be contacted by a research team. in the absence of a communication from the potential participant, the registry will include an individual's data in the PHI/PII released to researchers for screening, informed consent/assent, and enrollment in a research study.	

Last updated: 22 March 2024
Compiled by the NAACCR Data Security & Confidentiality Workgroup

Introduction

> *"Cancer registries have been used to provide compelling data documenting variations in cancer incidence and cancer mortality within and among different populations. Combined data from the NPCR and SEER registries have provided sufficient numbers of incident cases to examine rare cancers, specific histologic types, cancers at specific ages, and regional variations. In addition to contributing critical knowledge regarding patterns in cancer occurrence and a resource for cancer researchers, cancer surveillance data provide essential data with which to guide cancer prevention and control activities at the national, state, and local levels."*
>
> -White *et al*. The history and use of cancer registry data by public health cancer control programs in the United States. Cancer. 2017;123 Suppl 24(Suppl 24):4969-76.

The purpose of this document is to describe different conditions under which cancer registry data are requested from a central registry **for research** and to outline the steps that a registry can responsibly take prior to, during, and after data release to ensure patient confidentiality while also supporting the utility of cancer data. Cancer registries, as stewards of data that pertain to individuals with cancer, must work to ensure that the data they collect and curate are used in accordance with high ethical standards and for the common good. The value of cancer registry work lies in the use of these data. The cancer registry community is committed to achieving the maximum public health benefits for which cancer registry data may be used to reduce the burden of cancer.

In the Summer 2022 edition of the NAACCR narrative[2], T. Patrick Hill describes how data sharing (i.e. release) is a point of ethical obligation for individual registries because of the role that cancer surveillance data play in research that serves the common good. "[T]here is justification for surveillance and research conducted in a manner that not only avoids harming persons but actually benefits them." Because there is the potential for harm to an individual in the conduct of cancer surveillance, research, and data sharing, it is critical to use available tools (e.g. electronic and physical safeguards against breach) that minimize the potential for harm while maximizing the use of data. These guidelines are intended to be responsive to the challenge Hill presents, supporting an ethical and productive balance of harms and benefits, consistent with the legal, regulatory, and technical contexts in which cancer registries operate.

The NAACCR Data Security and Confidentiality Workgroup produced these guidelines to inform standards, policies, and practices for data release, i.e., the provision of a data set to an outside party by a population-based cancer registry. These guidelines are not intended to be and may not be read as legal advice. Rather, they are intended to express best practices for data release within contemporary frameworks of authorized, permissive, and ethical data sharing. The provision of data by central registries is subject to laws , regulations, and institutional policies that may be more or less restrictive based on local context, legislative intentions, and cultural norms. Readers are encouraged to hold the laws and policies governing their registries at the forefront of their consideration of these guidelines when using the content presented herein.

Where state law or registry policy restricts or prohibits data sharing, registries may need to work with their legal teams and state health departments to get amendments to public health law to allow release of data for research. Although beyond the scope of this document, registries should also weigh the merits of privacy, the merits of the purposes of data

[2] "Data Sharing For The Common Good: An Ethical Obligation?" Found at: https://narrative.naaccr.org/wp-content/uploads/2022/08/Summer-2022-PDF-.pdf

release, and an assessment of the probability that data release may breach privacy when developing guidelines for data sharing within their programs. Although we refer to HIPAA legislation, most – but not all – central cancer registries are not considered covered entities under HIPAA. For HIPAA rules developed primarily with data collected during an individual's medical care, privacy is considered an absolute. In the context of data release and research, while indubitably critical, privacy is not absolute. As result, defining privacy in a manner that is appropriate to this context requires great care.

Data Request Considerations

Although the subsequent sections of this document describe discrete steps and other recommended processes by which to release data to outside parties, releasing data includes manifold considerations that don't easily conform to a process or algorithm. Some of these considerations are described below; readers may want to review the following when using these guidelines and prior to and during any data release.

Referring requestors to existing resources

Data for many requests can be found in existing resources from NAACCR, the National Cancer Institute's Surveillance, Epidemiology and End Results (SEER) program, and the Centers for Disease Control and Prevention's National Program of Cancer Registries (NPCR). Consider recommending that the requestor verify that the equivalent of their request can't be found on publicly available websites/data repositories before running statistics in-house. Suggested resources are listed below:

- **CiNA Explorer** is an on-line, publicly accessible, interactive data visualization tool for quick access to cancer incidence statistics for major cancer sites. It is accessible to any user here: https://apps.naaccr.org/explorer/
    - **Cancer data visualization tools from other organizations include:**
        - NCI/CDC's State Cancer Profiles: https://statecancerprofiles.cancer.gov/
        - SEER's Cancer Statistics Explorer Network (SEER*Explorer): https://seer.cancer.gov/statistics-network/explorer/application.html
        - CDC's United States Cancer Statistics: Data Visualizations: https://gis.cdc.gov/Cancer/USCS/#/AtAGlance/

  Note: Data users should be mindful of the differences in data sources and methods utilized across these platforms.

- **CiNA Maps** is an on-line, publicly accessible, interactive mapping tool for quick access to the most recent five years of cancer incidence statistics. It is accessible to any user here: https://www.cancer-rates.info/naaccr/
- **Top Five Most Commonly Diagnosed Cancers Infographics** are static resources describing counts and percentages of the most common incident cancers by country, sex, and race/ethnicity. It is accessible here: https://www.naaccr.org/top-5-cancers/
- **CiNA Public Use Data Set** is a comprehensive data set comprised of cancer case data submitted to NAACCR from member registries in the U.S. and Canada. It is a publicly accessible incidence dataset

requiring only a signed Data Assurance Agreement for access. More information about the **CiNA Public Use Data Set** may be found here: https://www.naaccr.org/cina-public-use-data-set/

- **CiNA Research Data Sets** are comprehensive data sets comprised of cancer case data submitted to NAACCR from member registries in the U.S. and Canada. These datasets are available to NAACCR members and researchers working with a NAACCR member for use in specific analytic projects that are approved by a NAACCR application review working group. More information about the **CiNA Research Data Sets** may be found here: https://www.naaccr.org/cina-research/
    - Both the **CiNA Public Use Data Set** and the **CiNA Research Data Sets** must be accessed via the NAACCR Data Request Tracking (DaRT) system, which may be found here: https://apps.naaccr.org/dart
- **Virtual Pooled Registry** is a resource for investigators seeking cancer or other follow-up information on a pre-specified cohort of individuals enrolled in a study. It is a "single location to facilitate timely access to and use of high-quality cancer surveillance data for minimal risk linkage studies." Registries working with researchers requesting data linkages from multiple registries may refer researchers to the VPR: https://www.naaccr.org/about-vpr-cls/.

Research Review Processes

Registries may constitute a review team to evaluate research requests. Such teams may be comprised of the registry director, epidemiologists, and registry data managers. Research requests should be reviewed based on scientific merit, strength of the research team, institutional review and oversight at the study site, feasibility of the project, data security protocols, and other criteria as identified by the review team. The review team may meet periodically to discuss project requests and make a timely determination based on these considerations.

Research Review Processes for Accessing Canadian Data

The following resources may be of use to researchers interested in Canadian cancer registry data.

National/federal: Statistics Canada's Research Data Centres (RDCs) page includes procedures for application, legal agreement requirements, review processes, ethical considerations, and security clearance requirements: https://www.statcan.gc.ca/en/microdata.

Data release is reviewed by Statistics Canada staff for conformity to confidentiality release guidelines that generally include random rounding of counts, meeting a minimum cell size threshold, minimum population size for geographic areas must be at least 5,000 etc.

Provincial and/or territorial: Regional data access and linkage requests are facilitated by the Health Data Research Network Canada: https://www.hdrn.ca/dash/.

Data User Qualifications

To responsibly provide access to or permit use of registry data, central registries confirm the qualifications of the researchers. State laws may even require that an individual be a bona fide researcher in order to be

eligible for access and use of registry data. For specialized data requests, central registries undertake review the research team's qualifications – including their ability to appropriately analyze cancer surveillance data – to ensure that any product or publication resulting from the provision of cancer registry data will be appropriately analyzed, faithfully interpreted, and accurately presented to the public. If during its review of an application, a central registry notes a deficiency in research team expertise relative to the methods and objectives of the proposed research, a central registry may consider requiring the addition of a team member with the appropriate expertise to accomplish the research goals as a condition of the provision of the data set. Registries should also review the institutional support available to the requestor, both for scientific adequacy, but also with respect to the maintenance of confidentiality and data security. For example, data release to students may require faculty mentorship/supervision to ensure that data will be maintained in a secure environment and there will be oversight on the research. Specific qualifications are discussed within each section of the data release types in this document.

Prioritization of Requests

Data requests should be evaluated as they are received, and data requestors should be given an estimated turnaround time. Requests for aggregated data will likely be handled more quickly than case-level requests. As a public health resource that provides data to all data users and is not owned by or indebted to any other entity, including academic centers, the registry should not be expected to prioritize requests from one user over another. Additionally, registries may not wish to engage in ranking or prioritizing research studies based on scientific merit, even when requests for data effectively compete for the same patients or limited resources, e.g. tissue specimens.

Data Request Fees

While central registries do not typically charge fees for data, registries may charge researchers or other requestors for services including the preparation of data for access or release, and/or to conduct linkages or for other types of data release activities. Fees may be determined on a combination of factors, e.g. size of the cohort, per tumor, or time spent on specific components of the data request, e.g. IRB applications, conducting linkages, manually reviewing linkage results, dataset creation. In the context of patient contact studies, fees may also be applied to cover costs related to contacting physicians, and patient or next of kin consent to share their data. Some registries are not able or allowed to request fees because of statute or other reasons.

Tracking Data Requests

Data requests and releases should be documented in a tracking database (e.g., Excel, Access, other database-specific software). Items to consider recording include:

1) requestor's name;
2) requestor's institution;
3) requestor's email address;
4) brief description of request;
5) request date;
6) release/completion date;
7) type of request (e.g., research, hospital/service provider, state's department of health or office of public health, special study);
8) data use category (e.g., de-identified, limited data set, patient contact, etc.)
9) length of time spent completing the request;
10) the external data source provided (e.g., NAACCR, SEER, CDC-NPCR) if an external source is provided instead of a specialized data set from the central registry;
11) IRB status or other human subjects determination information, as applicable: IRB approval, exemption justification, and non-human subjects research determination, etc.;
12) data security considerations; and
13) registry staff completing request (name/initials)

Data Use and Confidentiality Agreements

Depending on the type of data request, confidentiality agreements and/or data use agreements may be required. Additional detail describing when to use which of these types of agreements is in Sections II to VI of "Types of Data Release and Recommended Protocols" below. Examples of these documents are included as appendices to this document.

Data Use Agreement (DUA) or Data Sharing Agreement (DSA)

DUAs/DSAs establish who may receive and use a data set, and how these data may be used. Parameters of re-release of data are typically included in the DUA/DSA. Additional broad information about DSAs may be found here: https://www.hhs.gov/hipaa/for-professionals/special-topics/emergency-preparedness/data-use-agreement/index.html.

Data Assurances Agreement (DAA)

Similar purpose to DUA/DSA as described above. NAACCR uses a DAA format as part of the Call for Data documentation to outline terms of data use, confidentiality, and the rights and responsibilities of NAACCR and the sending registry. See example: https://www.naaccr.org/wp-content/uploads/2023/09/4-Data-Assurance-AgreementUpdatedfor2023Sep2023.pdf

Data Confidentiality Agreement (DCA)

Registries may require DCAs for individuals accessing registry data. DCAs generally outline the conduct of researchers with respect to handling cancer data (e.g., agrees to not identify patients, adhere to small cell publication requirements, etc.).

Business Associate Agreement (BAA)

A legal document between a Covered Entity (such as a healthcare provider) and its Business Associate (such as a contractor engaged in providing goods or services—including software, for example—on behalf of or in service to the Covered Entity). BAAs do not apply to the relationships between covered entities and the central cancer registries to which the covered entities report cancer case data. In their interactions with healthcare providers/facilities, a central registry is not a contractor for the healthcare entity, but rather independently serves a public health surveillance function.[3]

Data Transfer Agreement (DTA)

A legal document/contract used to transfer human subject data between institutions or organizations for use in research. (IRB University Utah: https://irb.utah.edu/guidelines/repository/transfer.php)

Materials Transfer Agreement (MTA)

A legal document/contract used to transfer tangible research materials between entities. Although uncommon in the context of central cancer registry work, an MTA may be appropriate when a registry is affiliated with a health center and is facilitating tissue studies or other studies using biological materials (IRB University Utah: https://irb.utah.edu/guidelines/repository/transfer.php), and these specimens may be coded using registry data. Some registries have experienced an increase in requests for MTAs specifically for the proper handling of data sent to researchers, independent of linkage with specimen samples.

Memorandum of Understanding (MOU)

An MOU typically indicates a voluntary agreement between parties to assist in a shared, aspirational goal such as the implementation plans of a grant funded collaborative project. The agreement is written between the lead agency/applicant and a partnering entity. Some entities use an MOU when sharing data with researchers within the same institution. The MOU will generally define the overall program goals and describe the collaborative nature and relationship between the identified project and MOU-referenced participant. (Reference: https://www.acf.hhs.gov/sites/default/files/documents/fysb/mou_508.pdf)

Mode of Data Release/Transfer

A release of data refers to data released outside of the registry. Aggregated data may be released via unencrypted email, although as a best practice, registries may consider only releasing these data via secure means. Case-level datasets should always be released via secure transfer methods, such as secure email or secure file transfer protocol (SFTP).

[3] Healthcare entities unfamiliar with the role of a cancer registry may initially request that the registry enter into a BAA with the healthcare entity, and in most cases, another type of agreement, e.g. a DUA/DSA will be more appropriate.

Responding to Subpoena or other Legal Filings

Registry data are typically covered by Certificates and Assurances of Confidentiality, which broadly means that participants'/patients' identifiable information may not be released except under limited circumstances. State laws may also protect cancer registry data from subpoena; registry data are not primary source records like medical records and have less authoritative value for litigation purposes.

Assurance of Confidentiality (AoC)

An AoC offers formal confidentiality protection authorized under Section 308(d) of the Public Health Service Act (https://www.cdc.gov/os/integrity/confidentiality/index.htm). Under an AoC, "no identifiable information may be used for any purpose other than the purpose for which it was supplied unless such institution or individual has consented to that disclosure."

Certificate of Confidentiality (CoC)

A CoC protects the privacy of participants enrolled in research. COCs protect information, documents, and/or biospecimens that contain identifiable, sensitive information related to a participant (https://grants.nih.gov/policy/humansubjects/coc.htm). "Identifiable, sensitive information covered by a CoC must not be disclosed or provided:

- In any Federal, State, or local civil, criminal, administrative, legislative, or other proceeding OR
- To any other person not connected with the research.

Limited circumstances when the investigator and institution may release participant's identifiable sensitive information include:

- If required by other Federal, State, or local laws, such as for public health reporting of communicable diseases or child or elder abuse reporting;
- If the participant consents;
- If necessary for the medical treatment of the participant and made with the consent of the participant; or for the purposes of scientific research that is compliant with human subjects' regulations."

Considerations for Patient Contact Studies

Central registries that regularly engage in patient contact studies may encounter instances in which patients whose cancer cases are recorded in the registry are potentially eligible for inclusion in multiple and/or concurrent patient contact studies. Registries may also find that patients are eligible for inclusion in multiple studies across time (but not at the same time) or that patients are eligible for inclusion (and therefore potentially contacted) around the same time for concurrent studies—with some researchers using cancer registries to source potential participants, and others relying on other sources such as medical records review at an institution of care.

There are numerous considerations that registries should consider when deciding which patient information to release to researchers for contact and potential inclusion in a research study. Registries may wish to provide patients with sufficient time (after receiving a recruitment letter, for example) to contact the registry and request not to be contacted again in the future for research studies. Registries may also wish to limit possible confusion among patients receiving recruitment letters for different research studies (from different or even the same institution) at or around the same time.

Given the importance of voluntary participation to the ethics of research, we recommend that all potentially eligible patients who have not indicated that they would prefer to be excluded from patient contact studies be given the opportunity to participate in research, regardless of prior or recent opportunities for engagement in research.[4] This approach may require additional work on the part of the registry and researchers to devise a plan that maximizes patient agency and minimizes patient confusion, inconvenience, and possible distress. Some mechanisms by which to reduce confusion and inconvenience among patients include combining communication for multiple studies into one mailing or phone call when possible.

Registries that engage in patient contact studies may find it beneficial to require investigators to provide updated information to the registry about participants that have been contacted. One mechanism by which to do this is a "recruitment report," in which researchers report back: the number of patients that the researcher has contacted; how many and which patients (or their representatives) have asked not to be contacted again; and how many patients have enrolled. Registries may also wish to provide researchers with contact information for potential participants on a rolling basis that allocates a number of potential research participants to researchers that comports with study capacity. Communications are then made over a given time period, e.g. the next three months. Additional release of patient contact information may be contingent on the provision of the recruitment report for the prior period. (See Appendix C for example.)

At minimum, researchers should notify the registry if a potential participant expresses a desire to not be contacted for any future studies, or to provide the registry with updated contact information or vital status. Registries should flag those individuals who have requested no further contact within the registry database, so that information about these individuals is not made available for future patient contact studies.

When registries are contacting patients to determine interest in study participation, patients should also be given the opportunity to opt out of being contacted for future research studies. Suggested language for patient contact materials is below:

> "I, _____, hereby request to not be contacted by the [registry name] or researchers for any future patient contact studies."
>
> "I, _____, would like to opt out of all future patient contact studies."

[4] The NAACCR data standard includes fields that can be used to identify patients who do not want to be contacted for patient contact studies. The "No Patient Contact Flag" (NAACCR Item #1854, NAACCR version 23) identifies when a patient, family member, or provider has notified the central registry that the patient does not want to be contacted for participation in research studies. This flag is applied at the patient level and associated with all tumors for that patient.

For populations that are commonly studied but may experience social or cultural marginalization or oppression, the registry should encourage researchers to engage a community liaison or community-based organization that can assist in the communication to potential participants of the study's importance and the implications of participation. These liaisons may, in turn, require that researchers return to communicate their findings to the population under study as a means of increasing trust in the research process and engagement by commonly studied populations that may otherwise experience marginalization.

Attribution Language

The registry should request that any person or organization using registry data include a statement of acknowledgement in any presentation, report, or other publication using registry data.[5] This attribution or acknowledgment language should be included in data use and/or confidentiality agreements between the researcher and the registry. Suggested language for otherwise unpublished data sets is as follows:

- For studies using data from a single registry: "The authors would like to acknowledge the contribution to this study from [cancer registry name] supported by [cancer registry's funding agencies]."
- For studies using data from multiple registries: "The authors would like to acknowledge the contribution to this study from central cancer registries supported through the Centers for Disease Control and Prevention's National Program of Cancer Registries (NPCR) and/or the National Cancer Institute's Surveillance, Epidemiology, and End Results (SEER) Program. Central registries may also be supported by state agencies, universities, and cancer centers. Participating central cancer registries contributing data to this study include the following: [list of participating registries]."

Secondary Data Sharing

> *The law relating to things even today remains linked to the law relating to persons […] What imposes obligation in the present received and exchanged, is the fact that the thing received is not inactive. Even when it has been abandoned by the giver, it still possesses something of him. Through it the giver has a hold over the beneficiary just as, being its owner, through it he has a hold over the thief."*
>
> -(Marcel Mauss, The Gift, 11-12).

Secondary data sharing refers to data shared by a registry with a researcher that is then re-released or otherwise made available to a third party. The data sharing is "secondary" in the sense that a registry has defined a direct or primary recipient—a bona fide researcher—with a need that the registry can satisfy through providing access to and/or use of its data. Secondary data sharing activities—and the needs for them—derive or extend from that direct or primary relationship and research need.

[5] For instances using existing reports/publications, researchers should use the attribution language provided in the report/publication or cite appropriately.

Secondary data sharing takes a variety of forms. For example, a researcher may be required under a grant for which data is received to deposit a complete set of the data in a federal or federally-designated repository. A researcher might independently combine cancer registry data with other data and create a dataset amenable to future analyses. Sharing data with a researcher interested in exploring a hypothesis using that dataset would be a form of secondary data sharing. Many data linkage projects and participation in data consortia involve secondary data sharing. From the VPR Secondary Data Sharing Fact Sheet, "[t]he NIH Policy for Data Management and Sharing (https://sharing.nih.gov/data-management-andsharing-policy/about-data-management-and-sharing-policies) reinforces NIH's longstanding commitment to making the results and outputs of NIH-funded research available for further research through effective and efficient data management and data sharing practices."

Some cancer registries restrict secondary data sharing categorically whereas others might specify specific purposes for which secondary data sharing is not permitted. The State of Montana, for example, does not permit secondary data release for market research purposes. Cancer registries with more restrictive policies tend to view each data release as a cul-du-sac. In these cases, any bona fide researcher, repository or consortium interested in the same data would need to make an independent data request of the source under the policies and procedures of these cancer registries. A major limitation of this approach, however, is the lost value of data being combined and processed through research activities that enhance the scientific use of the underlying cancer registry data. In practice, cancer registries with more restrictive policies are often driven by a mandate to identify, record, and have a direct relationship of authorization and/or permission with each end user of the cancer registry data. Through processes such as aggregation, de-identification, and/or combination with other data, the data may or may not be conceived as having severed connection to the source cancer registry. Registries that engage in secondary data sharing should – and may be required to – track:

1. How shared data will be used?
2. Who will have access to these data?
3. Will the data be shared beyond this secondary release? If so, by which mechanisms and with which entities?
4. In which format and where will data reside?

The topic of secondary data sharing has become one of elevated interest as revisions to the US federal Common Rule (2018 Common Rule) have lessened the level of review, ethical oversight, and human subjects' protection afforded to secondary data analyses in the interest of advancing scientific research for the common good using existing – and typically de-identified – data. These changes to the Common Rule also reduce the amount of resources required to review applications for use of de-identified data in secondary analyses. Current exemptions from the oversight of Institutional Review Boards now extend to "Secondary research uses of identifiable private information or identifiable biospecimens, if at least one of the following criteria is met: … (ii) Information, which may include information about biospecimens, is recorded by the investigator in such a manner that the identity of the human subjects cannot readily be ascertained directly or through identifiers linked to the subjects, the investigator does not contact the subjects, and the investigator will not re-identify subjects;…" (https://www.hhs.gov/ohrp/regulations-and-policy/regulations/45-cfr-46/revised-common-rule-regulatory-text/index.html#46.104)

The standard of de-identification embodied in the Common Rule exemption does not have universal application with respect to classification of permissible versus impermissible secondary data sharing of cancer registry data. As of this writing, more consistent standards, calibrated to balance privacy interests with potential scientific benefits, would be a significant contribution to the fields of cancer registration, surveillance, and research. NAACCR has developed a factsheet on Secondary Data Sharing that registries may find helpful when explaining the concept of secondary data sharing and its importance in "biomedical research discovery, enhance[ing] research rigor and reproducibility, provid[ing] accessibility to high-value datasets, and promot[ing] data reuse for future research studies": https://www.naaccr.org/wp-content/uploads/2023/02/Cancer-Registry-Secondary-Data-Sharing-Fact-Sheet_Final_2.9.23.pdf

Types of Data Release and Recommended Protocols

The following sections describe types of requested data and recommended protocols and documents required for the release of said data to researchers.

I. Aggregated Data

1) **Definition**: Aggregated data collapses differences between and/or combines information from multiple people or tumors into a summary table or set of fields. Summary measures such as counts or rates can be provided as a data set or displayed in a table. Aggregated data do not include line-level data, even if aggregation over multiple variables results in a single record or count for a particular category, e.g. age group.
2) **Recommended Requirements**:
   a) Release data to requestor following receipt of a complete and verified request.
   b) Ensure the data recipient agrees to acknowledge the data source in any reports and publications.
   c) Align with state statue or other applicable policies (see NAACCR Suppression Document for additional information) when releasing data stratified by geographical unit. Additional levels of review may be required when the geographical unit being requested is small, e.g. census tract.
   d) Follow state law, registry policy and contractual requirements, with respect to cells or categories with a small number of cases/counts will be suppressed for patient confidentiality, typically if < 6 cases (see NAACCR Suppression Document for additional information).
   e) Suppress rates for statistical stability per registry's judgment and/or guidelines. As a general rule, rates are suppressed if based on < 16 cases or deaths and/or the underlying population consists of fewer than 20,000 people (see NAACCR Suppression Document for additional information).
   f) Require that the investigator follow the guidelines in the de-identified data section (i.e. a protocol must be submitted and approved) if the requestor intends to engage in secondary analyses of the aggregated data or if the aggregated data includes sensitive data (see Section II).
3) **Re-release of data**: Aggregated data are typically amenable to being re-released for secondary analyses.

II. De-Identified Case-Level Data

1) **Definition**: Although registry data are generally considered to derive from public health surveillance activities in accordance with Health Insurance Portability and Accountability Act (HIPAA) regulations, the "Safe Harbor" method (45 CF 164.514(b)(2)(i)) of removing 18 specific identifiers is commonly used as a standard for de-identification. Under these provisions, a de-identified dataset as defined by the HIPAA Privacy Rule requires that data items be removed (Safe Harbor Method) or an expert determines that the dataset is not individually identifiable (Expert Determination Method) or the data set be hashed via one-way cryptography.
   a) Safe Harbor Method: The data items in Appendix A must be removed. A unique identifying number, characteristic, or code can be included provided the stipulations specified in Appendix A are followed.
   b) Expert Determination Method: A person with appropriate knowledge of and experience with generally accepted statistical and scientific principles and methods for rendering information not individually identifiable:
      i) When applying such principles and methods, determines that the risk that the information could be used, alone or in combination with other reasonably available information by the intended recipient to identify an individual who is a subject of the information is very small; and
      ii) Documents the methods and results of the analysis that justify such determination.
      iii) Expert determination is not widely used by registries.

   Registries may also leverage Privacy Protecting Record Linkages (PPRL) to link records with sensitive data without revealing identifying information. With PPRL, records from two different sources are linked by one-way cryptography, commonly known as hashing, [6] whereby each person's PHI/PII is rendered encrypted for the linkage and data transfer.[7] A third party can then compare the hashed values to see if a pair of records are from the same person, without revealing that person's identity. Additional information describing hashing may be found here: https://www.naaccr.org/wp-content/uploads/2023/02/Data_Encryption_and_Hashing_Primer_v2_FINAL.pdf.

2) **Requirements**: Researchers may receive case-level de-identified records by satisfying a set of criteria that includes:
   a) Providing a written description of the proposed research project.
   b) Providing CV/Resume and contact information of principal investigator.

[6] Guidance Regarding Methods for De-identification of Protected Health Information in Accordance with the Health Insurance Portability and Accountability Act (HIPAA) Privacy Rule: https://www.hhs.gov/hipaa/for-professionals/privacy/special-topics/de-identification/index.html
[7] CDC Foundation: Data Linkage and Identity Management – Privacy Protecting Linkage: https://www.cdcfoundation.org/CDCFoundationPPRLSummary.pdf?inline

c) Providing a signed confidentiality agreement or data assurance agreement from all those who might have contact with the data.
d) Obtaining approval by the Director of cancer registry or other designated person. Typically, the registry's research or data request review committee does not need to review de-identified data requests. Registries can, however, choose to be more restrictive and require committee review for de-identified data sets.
e) Agreeing, as a recipient of the data, to acknowledge the cancer registry for the data in any reports and publications (see Attribution Language in the section on Data Request Considerations) and provide manuscripts to the registry for review prior to submission, if requested.
f) Agreeing to provide a progress report or status update periodically (which would include a list of publications utilizing the registry's data), destroy the data upon completion of the project, and notify the cancer registry of its destruction.
g) Agreeing to the deletion, destruction, or other final disposition of data in accordance with registry policy. The registry may provide documentation to the data requesting party describing best practices in data destruction prior to the provision of data. Please see the data destruction primer for more details: https://www.naaccr.org/wp-content/uploads/2022/01/NAACCR-Data-Destruction-Primer-2021115.pdf.

3) **Release of registry data**: Researchers shall NOT release **registry** data or use the data for any projects other than the project specified in the proposal. Note that the release of augmented data may be accommodated separately, subject to specifics of executed agreements, that may allow for re-release based on the idea that the augmented data are a unique resource of separate and distinct value than the cancer registry data that may have informed or contributed to them. Please review the Secondary Data Sharing section or this fact sheet for more information regarding secondary data sharing: https://www.naaccr.org/wp-content/uploads/2023/02/Cancer-Registry-Secondary-Data-Sharing-Fact-Sheet_Final_2.9.23.pdf

III. Limited Data Set

1) **Definition**: A limited data set excludes direct identifiers in order to reduce the burden of data security and confidentiality requirements associated with access and use data, but nevertheless includes sufficient information that the data may potentially be used to identify individual persons. Details regarding limited data sets released for linkage studies or patient contact studies can be found in Sections IV-VI. Central registries typically follow the HIPAA definition of a "limited data set" in preparation of data for research purposes as a useful standard for ensuring the ethical and effective management of cancer data. The following identifiers are removed in order to compose a limited data set:[8]
    a) Names
    b) Postal address information (other than town or city, state, and ZIP code)
        a. While this is not included in the HIPAA Administrative Simplification Regulation Text, census tract and block group should also be removed. ZIP code should be truncated to the first three (3) digits.

[8] Items with an asterisk above are not standard/uncommonly encountered data items for cancer registries. Reference, pg. 98: https://www.hhs.gov/sites/default/files/ocr/privacy/hipaa/administrative/combined/hipaa-simplification-201303.pdf

c) Phone and fax numbers
d) Electronic mail addresses*
e) Social security number
f) Medical record numbers
g) Health plan beneficiary numbers*
h) Account numbers*
i) Certificate/license numbers*
j) Vehicle identifiers and serial numbers, including license plate numbers*
k) Device identifiers and serial numbers*
l) Web Universal Resource Locators (URLs)*
m) Internet Protocol (IP) address numbers*
n) Biometric identifiers, including finger and voice prints*
o) Full face photographic images and any comparable images*

Additionally, while not mentioned in the HIPAA Administrative Simplification Regulation Text, identities of physicians, hospitals, or other healthcare providers and facilities should also be removed.

2) **Requirements**: Researchers may receive individual case records in a limited data set by providing the following items to the registry for review:
   a) A written research proposal/protocol (Appendix B)
   b) A CV/Resume and contact information of principal investigator.
   c) A signed data use agreement between institutions and/or data user and confidentiality agreements by all individuals who may access the registry's case-level data. The suggested template for the data confidentiality agreement is the Data Confidentiality Agreement for NAACCR Researchers (Appendix D2) and the Data User Agreement for Access to the NAACCR CiNA Dataset (Appendix D3). The suggested template for the data use agreement is the Virtual Pooled Registry Templated DUA (Appendix D4).
   d) The IRB Approval of the study from the following:
      1) Researcher's Institution: Researcher should submit their IRB application (or protocol) and approval documents to the registry.
      2) Registry's Institution: According to the Revised Common Rule, IRB approval from the registry's institution is ***only required if*** the registry is engaged in the research (e.g. registry is part of the research team, obtains data about biospecimens from participants through intervention/interaction, obtains identifiable private information about the participants, or obtains informed consent from study participants).[9] If the registry is engaged in the research, the registry may elect or be required to use the researcher's IRB of record as a single IRB for the study (i.e. use a central IRB).
   e) An agreement in which the Data recipient(s) commits to acknowledge the registry for the data in any reports and publications (see Attribution Language in the section on Data Request Considerations).

[9] NAACCR Webinar on the Revised Common Rule: https://education.naaccr.org/products/revised-common-rule

f) Written confirmation that the researchers will provide a progress report or status update periodically (which would include a list of publications utilizing the registry's data), destroy the data upon completion of the project, and notify the cancer registry of its destruction.
g) Written confirmation that the researchers will ensure the deletion, destruction, or other final disposition of data in accordance with registry policy. The registry may provide documentation to the data requesting party describing best practices in data destruction prior to the provision of data. Please see the data destruction primer for more details: https://www.naaccr.org/wp-content/uploads/2022/01/NAACCR-Data-Destruction-Primer-2021115.pdf.

4) **Release of data**: Researchers shall NOT release registry data or use the data for any projects other than the approved one. Registry data may not be disclosed for any civil, criminal, administrative or other legal proceeding. Note that the release of augmented data may be accommodated separately, subject to specifics of executed agreements, that may allow for re-release based on the idea that the augmented data are a unique resource of separate and distinct value than the cancer registry data that may have informed or contributed to them. Please see this fact sheet for more information regarding secondary data sharing: https://www.naaccr.org/wp-content/uploads/2023/02/Cancer-Registry-Secondary-Data-Sharing-Fact-Sheet_Final_2.9.23.pdf

3) **Approval of data request**: The registry's review team will review the proposal for compliance with accepted confidentiality procedures and appropriateness of research design, including the ability of the researchers to answer their stated question with the available data.

IV. Data Linkages

The data release policies under this category apply only to researchers who have names and identifiers of patients and request linkages with the registry's database to obtain additional diagnostic, treatment, or follow-up information.

Registry personnel perform linkages between research data sets with PII/PHI and the registry data with PII/PHI to identify potential matches. The registry should discuss and establish a matching protocol with researchers, in particular if researchers will be linking with multiple population-based cancer registries. Matched cases will be returned to the researchers with the study number as the only patient-specific identifier. Although investigators will receive PHI from the registry for their cohort, name, date of birth, SSN, and other direct identifiers should be removed from the data set being returned to the researchers. Registry staff may additionally advise that the researchers contact the Virtual Pooled Registry to request linkage with registries, in particular if they are requesting data from multiple registries.

1) **Requirements**: Provide the following items to the registry for review:
   a) A written research proposal/protocol (Appendix B)
   b) A CV/Resume and contact information of principal investigator(s).

c) A sample patient consent form to determine if study participants were/are/will be notified of medical record or cancer registry review. ***If they were not***, then an approved waiver of consent should be provided by the researcher's institutional IRB.
d) IRB Approval
   i. Researcher's Institution: IRB application and approval documents. *Note: This requirement could also be covered by a central IRB.*
   ii. Registry's Institution: IRB approval ***only if*** registry's investigators will participate in the research. *Note: This requirement could also be covered by a centralized IRB.*
e) An agreement that covers data use and confidentiality as applicable. An agreement is typically required for each person who will access the data. Please review section "Data Use and Confidentiality Agreements" for more information on specific agreement types.
f) Written confirmation that the Data recipient agrees to acknowledge the registry for the data in any reports or publications (see Attribution Language in the Data Request Considerations section).

2) **Re-release of data**: Researchers shall NOT release registry data or use the data for any projects other than the approved one. All U.S. registry data are covered by either an Assurance of Confidentiality, Certificate of Confidentiality, or both; as such, registry data may not be disclosed for any civil, criminal, administrative or other legal proceeding. Note that the release of augmented data may be accommodated separately, subject to specifics of executed agreements, that may allow for re-release based on the idea that the augmented data are a unique resource of separate and distinct value than the cancer registry data that may have informed or contributed to them. Please see this fact sheet for more information regarding secondary data sharing: https://www.naaccr.org/wp-content/uploads/2023/02/Cancer-Registry-Secondary-Data-Sharing-Fact-Sheet_Final_2.9.23.pdf

3) **Approval of data request**: The registry's research committee and/or IRB will review the proposal for compliance with accepted confidentiality procedures and appropriateness of research design.

V. Requests for Data with Identifiers – No Patient Contact

Case-level data that include any identifier listed in Appendix A are considered confidential. Researchers request the registry to provide a dataset that includes cases meeting the study's eligibility criteria and data is considered identifiable, but researchers do not plan to contact cohort members.

1) Requirements: Provide the following items to the registry for review:
a) A Written proposal/protocol (Appendix B)
b) A CV/Resume and contact information of principal investigator(s).
c) A signed data use agreement or confidentiality agreements appropriate to the data request. See Appendix D for examples of templated agreements.
d) IRB Approval
   1) IRB application and approval documents from researcher's institution. *Note: This requirement could also be covered by a centralized IRB.*

2) IRB approval from registry's institution ***only if*** registry's investigators will participate in the research. *Note: This requirement could also be covered by a centralized IRB.*

e) Written confirmation that the data recipient agrees to acknowledge the registry for the data in any reports and publications (see Attribution Language in the Data Request Considerations section).

2) Re-release of data: Researchers shall NOT re-release Registry Data or use the Registry Data for any other projects other than the approved one. Registry data may not be disclosed for any civil, criminal, administrative or other legal proceeding. For more details about sharing Augmented Data, see Secondary Data Sharing.

3) Approval of data request: The registry's research committee will review the proposal for compliance with accepted confidentiality procedures and appropriateness of research design. Reviewers will consider the minimum necessary standard when approving identifiers for release.

VI. Requests for Data with Identifiers – Patient Contact Studies

Case-level data that include any identifier listed in Appendix A are considered confidential. These data are most commonly requested when researchers request to use registry data to identify potentially qualifying subjects for a study cohort, with the intent to contact, screen, and consent patients for study participation. The registry will follow agreed upon protocols for releasing patient contact information to researchers (see Table below).

Please note: Patient contact studies are handled differently within each registry. Given that registry time and resources are often required for patient contact studies, a budget based on the protocol and scope of work, which is dependent upon the scenarios outlined below, is typically developed and further collaboration is dependent upon the approval of the budget.

1) **Potential Study Participant Contact:** How a registry chooses to contact potential participants for patient contact research studies is dependent on registry policy and staff time. A patient's physician may be the first point of contact, however registries should follow policies and procedures that maximize patient agency and ability to participate in research. Notifying a patient's physician that their patient is to be contacted for a research study (typically referred to as Physician Notification) may be required by registry guidelines, which may be based on state statute. In Table 1 below are mechanisms by which a registry may contact potential participants for participation in research studies.

Last updated: 22 March 2024
Compiled by the NAACCR Data Security & Confidentiality Workgroup

Table 1: Mechanisms by which central cancer registries may choose to engage with physicians and potential participants in the context of patient contact studies.				
Type of Contact	**Physician Notification**	**Patient Notification**	**Party Initiating Patient Contact**	**Contact information provided for:**
Patient Consent Only	None	Active/passive consent to share contact information with researchers.	Registry	Any patient providing active consent or passive consent.
Physician Notification and Patient Consent	Registry notifies physician that their patient might be contacted for research. No opportunity for consent/refusal provided.	Active/passive consent to share contact information with researchers.	Registry	Any patient providing active consent or passive consent.
Physician Consent and Active Patient Consent	Registry obtains active/passive physician consent to contact patient for research.	Active consent to share contact information with researchers.	Registry	Any patient data with active or passive consent from physician and active consent from patient.
Physician Consent and Passive Patient Consent	Registry obtains active/passive physician consent to contact patient for research.	Passive consent to share contact information with researchers.	Registry	Any patient data with consent from physician and active or passive consent from patient.
Physician Consent and Patients Contact Researchers	Registry obtains active/passive physician consent to contact patient for research. ** Physician contact could be bypassed within this model.**	Registry notifies potential participants of eligibility and provides researcher contact information.	Registry	None. Patients contact researchers if they are interested in participating.
Active/passive physician consent and direct contact by researcher to potential participants	Registry obtains active/passive physician consent to contact patient for research.	None.	Researchers	Any patient data with consent from physician.
No physician contact and direct contact with patient by researchers	None.	None.	Researchers	Researchers receive contact information for any potentially eligible participant in registry data.

For any of the above physician or potential participant contact work, a registry may work with a trusted third party to either contact physicians, patients, or next of kin for consent to be contacted. Contact information for patients who have previously refused to participate in research or those who have opted out of the current research study shall not be shared with researchers.

2) Requirements

Provide the following items to the registry for review:

i) A written research proposal/protocol including description of technical components (analytic plan, power calculations, methods by which patients will be identified and contacted) and ethical considerations (Appendix B).
ii) A CV/Resume and contact information of principal investigator(s)
iii) A signed data use agreement or confidentiality agreements appropriate to the request. A DUA and confidentiality agreements are required in all scenarios **except** in cases where the patient contacts that researcher. Please review section "Data Use and Confidentiality Agreements" for more information on specific agreement types.
iv) Written confirmation that the data recipient agrees to acknowledge the registry for the data in any reports and publications (see Attribution Language in the Data Request Considerations section).
v) IRB Approval
   (1) Researcher's Institution: IRB application and approval documents. *Note: This requirement could also be covered by a centralized IRB.*
   (2) Registry's Institution: IRB approval ***only if*** the registry's investigators will participate in the research. *Note: This requirement could also be covered by a centralized IRB.*

3) Re-release of data: Researchers shall NOT re-release data or use the data for any other projects other than the approved one. Registry data may not be disclosed for any civil, criminal, administrative or other legal proceeding.

4) **Approval of data request**: The registry's research committee will review the proposal for compliance with accepted confidentiality procedures and appropriateness of research design. The registry's research committee will also consider the registry's capacity to accomplish the scope of work and the research team's ability to support the registry's work. In other words, there must be agreement between the research and registry teams on the registry's scope of work and compensation for this work prior to any data transfer or approval by the registry to participate in the study. The registry may enter into a subcontract or have another mechanism with the investigator's institution to be able to participate in the project and receive compensation for the work. Registry researchers may also be included as collaborators or co-investigators on the research protocol depending on the required effort and involvement of registry personnel.

Restricted Data

I. Cases from the Department of Veterans Affairs (VA) or Department of Defense (DOD)

The VA and DOD, noting that DoD healthcare facilities are managed by the Defense Health Agency (DHA), typically prohibit the re-release of data for research purposes, regardless of de-identification. However, some registries may be able to re-release de-identified data depending on the agreement in place with the institution. Always review your state's agreement prior to release to understand the applicable restrictions.

"VA-only Cases" or "DOD-only Cases" shall mean a case in the registry database where the only information on the cancer was received from the VA or the DOD. There is no other report of the case submitted by any other institution.

a. Surveillance activities: Cases that are only reported by the VA/DOD may be released to national agencies for routine surveillance.
b. Research studies: VA/DOD-only cases are typically excluded from research data sets because of existing agreements established by the VA/DOD with central cancer registries. If a case is reported by the VA/DOD ***and*** another institution, the case can be included in research datasets provided other restrictions on release are not present.
    i. Starting in NAACCR version 23 there is a new field characterizing facility reporting restrictions for re-release for research.
    ii. NAACCR version 23 provides a "Reporting Facility Restriction Flag" (NAACCR Item # 1856). Please see the NAACCR Data Dictionary for more information regarding this field: https://apps.naaccr.org/data-dictionary/
c. Patient Contact Studies: Unless the registry's agreement with the VA/DOD explicitly states otherwise, the registry shall exclude VHA-only and DOD-only cases from all patient contact studies.
d. Linkage Studies: VHA- and DOD-only cases must be removed from all datasets used for research (even if patient consent was obtained) unless the registry's agreement with the VA/DOD explicitly states otherwise. However, there may be instances where VA cases may be included in de-identified datasets. Please review your current DUA/DTA with the VHA.

II. Cases received from other states through the Inter-Registry Data Exchange

Cases reported to the cancer registry **solely** via inter-registry data exchange agreements are handled as follows according to the NAACCR National Interstate Data Exchange Agreement from 2021[10]. See Table X for a summary.

a. "Interstate-exchange-only cases" means a case for which the only information on the cancer case available to the receiving registry was obtained through interstate exchange from the

[10] https://www.naaccr.org/wp-content/uploads/2021/08/NAACCR-interstate-exchange-agreement_for-signature_7.23.21.docx

sending registry.[11] In this context, there is no other report of the case submitted by any other institution within the catchment area of the receiving registry. (Death certificates and other non-NAACCR abstract data sources are typically not considered sufficient additional sources of cancer-related data, such that the out-of-state case is rendered releasable.)

b. If the sending registry has signed the NAACCR July 2021 interstate data exchange agreement without addendums, including the sending registry's interstate-exchange-only cases in de-identified research data sets is permitted without permission from the sending registry. The sending registry's written permission for receiving registry's release of data is only required when release includes ***identifiable information*** on interstate-exchange-only cases. Written permission from the Sending Registry is not required for the following:
    i. Release of identifiable information where the receiving registry has also received cancer information from another reporting facility without restrictions on including data in research data sets.
    ii. Release of a limited data set to researchers conducting studies facilitated by the Virtual Pooled Registry Cancer Linkage System.
    iii. Release of de-identified data.
c. **Please note**: For states that have not signed the NAACCR Agreement or have not signed the updated agreement from 2021, the specific signed agreements must be reviewed for details on handling interstate-exchange-only cases.
    i. https://www.naaccr.org/national-interstate-data-exchange-agreement/

[11] "Receiving Registry" shall mean the Central Cancer Registry that receives resident cancer information from the Central Cancer Registry where the cancer or non-malignant brain tumor was diagnosed or treated.
"Sending Registry" shall mean the Central Cancer Registry that shares resident cancer information with the Central Cancer Registry where the patient resides.

Last updated: 22 March 2024
Compiled by the NAACCR Data Security & Confidentiality Workgroup

Table 2: Summary of Research Use for Restricted Data Sources			
Where are data coming from?	**Can be released for patient contact studies?**	**Can be released for de-identified data sets for researcher?**	**Virtual Pooled Registry, Other Linkage Studies, and Beyond (i.e. genomics)**
No restrictions on release based on reporting facility.	Yes (providing there are no physician or patient objections)	Yes	Yes
Out Of State Cases: Tumor records received only from Out of State (OOS) data exchange with another central registry	In instances where data being requested are not de-identified, i.e., patient contact studies, re-release is typically not permitted without explicit permission from the sending OOS registry.	In instances where the data being released is deidentified, OOS tumors may be released for research when the sending OOS registry has signed the NAACCR Interstate Exchange Agreement (revised July 2021) and did not otherwise specify a restriction in an addendum.	In instances where the data being released is deidentified, OOS tumors may be released for research when the sending OOS registry has signed the NAACCR Interstate Exchange Agreement (revised July 2021) and did not otherwise specify a restriction in an addendum.
VHA or DoD: Tumor records received only from Veterans Health Administration (VHA) and/or Department of Defense	No, unless your registry's agreement specifically stipulates otherwise.	No, unless your registry's agreement specifically stipulates otherwise. For SEER registries, the NCI-VHA 2023 MOU supersedes local VA agreements.	No, unless your registry's agreement specifically stipulates otherwise.
VHA/DoD and OOS *or* DOD, VHA, and OOS: Tumor records received only from Veterans Health Administration (VHA) and/or Department of Defense AND out-of-state data exchange	No, unless your registry's agreement specifically stipulates otherwise.	In instances where data being released are deidentified, these tumors may be released for research when the sending OOS registry has signed the NAACCR Interstate Exchange Agreement (revised July 2021) and did not otherwise specify a restriction in an addendum.	In instances where data being released are deidentified, these tumors may be released for research when the sending OOS registry has signed the NAACCR Interstate Exchange Agreement (revised July 2021) and did not otherwise specify a restriction in an addendum.
Other situations			

Death Certificate/Autopsy Only cases	N/A	There may be value in providing DCO cases for linkage and use in de-identified research. Deceased persons are not considered research subjects/covered under HIPAA, but your ability to release depends on agreement with vital statistics in your state.	There may be value in providing DCO cases for linkage and use in de-identified research. Deceased persons are not considered research subjects/covered under HIPAA, but your ability to release depends on agreement with vital statistics in your state.

III. Release of Death Certificate Data

Release of cause of death depends on registry policy, the source of the information, and agreements with the state's bureau of vital statistics. Registry staff will review the source of the information before releasing data.

a. **Cause of Death from State's Bureau of Vital Statistics**: Release of COD from the state is subject to state policy and data agreements between the registry and the vital statistics bureau; however, this information is generally available for release without an additional request to the state's bureau of vital records. In instances where COD is not able to be released, SEER's cause-specific death classification may meet researcher needs (see "c" below).
b. **Cause of Death from the National Death Index (NDI):** The registry will conduct a formal review of a request for cause of death information (which may or may not require IRB approval) before releasing cause of death obtained from the NDI. Cause of death obtained from NDI may be released for research and surveillance purposes. This information may not be re-released by the researcher. For more information, please review NAACCR's NDI Fact Sheet: https://www.naaccr.org/wp-content/uploads/2020/04/NDI-Factsheet-for-NAACCR_Final.pdf.
    i. The registry relates to the NDI as an independent data repository with confidentiality and data security requirements separately applicable to its data.
    ii. The registry must provide the NDI with information regarding the release of such data (i.e. researcher name, organization, study title, date) annually.
c. **SEER Cause-specific Death Classification**: Registries may be able to release SEER cause-specific death (https://seer.cancer.gov/causespecific/) in lieu of actual cause of death information without additional permission from the state's vital statistics office.
d. **Vital Status**: Vital status is typically not restricted.
e. **Date of Last Contact (Date of Death):** Restrictions on date of death typically follow full date restriction rules. If the full date of death/follow-up is included in the data set to researchers, the data set is considered a limited data set. Registries may opt to calculate days from diagnosis (or other timepoint) for researchers in lieu of releasing full date or death/follow-up.

Appendices

Appendix A: Safe Harbor Method

The following data items must be removed in the Safe Harbor Method of de-identification.

(1) Name
(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 is 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 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) Fax numbers
(6) Email addresses
(7) Social Security Numbers
(8) Medical record numbers
(9) Health plan beneficiary numbers
(10) Account numbers
(11) Certificate/license numbers
(12) Vehicle identifiers and serial numbers, including license plate numbers
(13) Device identifiers and serial numbers
(14) Web Universal Resource Locators (URLs)
(15) Internet Protocol (IP) addresses
(16) Biometric identifiers, including finger and voice prints
(17) Full-face photographs and any comparable images
(18) Any other unique identifying number, characteristic, or code, except as permitted with respect to a re-identification code as defined below:
    "(c) *Implementation specifications: re-identification.* A covered entity may assign a code or other means of record identification to allow information de-identified under this section to be re-identified by the covered entity, provided that:
        (i) *Derivation.* The code or other means of record identification is not derived from or related to information about the individual and is not otherwise capable of being translated so as to identify the individual; and
        (ii) *Security.* The covered entity does not use or disclose the code or other means of record identification for any other purpose, and does not disclose the mechanism for re-identification."

Source: https://www.hhs.gov/hipaa/for-professionals/privacy/special-topics/de-identification/index.html

Appendix B: Suggested Contents of Research Proposal

1) Title of the Project
2) Name of the funding agency and the funding period
    a) If unfunded, please document.
3) Grant/contract number, if applicable
4) IRB Status/Documentation
    a) Provide Consent Documents for Review
5) Description of the study
    a) Include years of diagnosis requested
    b) Include primary cancer sites or other tumor characteristics
6) Background and Objectives
7) Methodology
    a) Research Design
        1. Eligibility Criteria
    b) Data Analysis Plan
        1. Including comparison group as applicable
        2. Description of how the registry's data will be used
        3. List of Requested Data Items (with NAACCR or registry-specific identifiers)
            a. Including justification for need of sensitive or potentially identifying data items.
8) Project Timeline
9) Description of physical and policy safeguards to protect data
10) Description of data destruction protocol/plan
11) Study personnel: description of the study team's staffing and expertise to complete project

Appendix C: Example Recruitment Report

Recruitment Report n of N **IRB # (enter here)** **Date: (enter here)**	*Previous 3 months* *(From DD-MM-YYYY to DD-MM-YYYY)*	*Entire Recruitment Period* *(From DD-MM-YYYY to DD-MM-YYYY)*
How many patients were provided to you by the Registry?		
Out of these patients, how many have you contacted?		
How many patients -or those speaking on behalf of the patients- have contacted you, asking not to be contacted again?*		
How many patients have enrolled?		

*Please also send the list of Registry patient IDs in this category to XXXXXXXX@registryemail.com every three months.

Appendix D: Examples of Data Assurance, Confidentiality, and Use Agreements used for NAACCR Programs

NAACCR, Inc. CALL FOR DATA ASSURANCES AGREEMENT

Agreement executed this _____day of _____, 20 ___, by and between the **NORTH AMERICAN ASSOCIATION OF CENTRAL CANCER REGISTRIES, INC.** ("NAACCR, Inc."), a California corporation, and _____ ("REGISTRY") of _____, _____.

(Name) *(City)* *(State/Province)*

NAACCR, Inc. is engaged in an annual Call for Data to conduct data evaluation, aggregation, analysis and publication of cancer incidence, specifically described in Attachment A:

NAACCR, Inc. uses and analyzes certain cancer incidence data (the "Data"). NAACCR, Inc. agrees and acknowledges that patient confidentiality is of the utmost importance in the use of the Data and in the manner in which all research results are presented and/or published. Accordingly, in consideration of NAACCR's receipt of the Data from the Registry, NAACCR, Inc. assures REGISTRY as follows:

1. NAACCR, Inc. agrees to treat the Data received from Registry as private, non-public health information. The Data will be used solely for the specified analyses and research described in Attachment A and not for any other purpose. The Data will never be used as a basis for legal, administrative or other adverse actions that can directly affect any individual about whom personal and/or medical information is included in the Data.

2. NAACCR, Inc. understands and agrees that any and all Data which may lead to the identity of any patient, research subject, physician, other person, or reporting facility is strictly privileged and confidential and agrees to keep all Data strictly confidential at all times.

3. If, in the course of evaluation, analysis, and research, NAACCR, Inc. believes it necessary to provide access to the Data to any NAACCR, Inc. researcher, NAACCR, Inc. will not do so unless and until such individual has properly executed a Data Confidentiality Agreement for NAACCR Researchers which has been accepted, in writing, by NAACCR, Inc. NAACCR, Inc. agrees to notify Registry in writing within forty-eight (48) hours of becoming aware of any violation of this Assurances Agreement or any Assurances Agreement executed by any other individual, including full details of the violation and corrective actions to be taken by NAACCR, Inc.

4. NAACCR, Inc. further agrees that all data provided under the provisions of this Assurances Agreement may only be used for the purposes described in Attachment A. Requests for ad hoc uses will only be provided after obtaining consent from each registry for each use.

5. NAACCR, Inc. agrees that (i) any and all reports or analyses of the Data prepared by NAACCR, Inc. shall contain only aggregate data. NAACCR, Inc. further agrees that (ii) at no time will any individual names or other personally identifying information or information which could lead to the identification of any Data subject ever be published, (iii) no report of the Data containing statistical cells with less than six (6) subjects shall be released without the prior written authorization of REGISTRY, (iv) aggregate data that identify individual REGISTRY will not be published without approval from REGISTRY.

6. NAACCR, Inc. further agrees that all data provided under the provisions of this Assurances Agreement shall remain the sole property of REGISTRY and may not be copied or reproduced in any form or manner without REGISTRY's prior written consent. Notwithstanding the foregoing, a NAACCR researcher may copy and maintain the Data on personal computer as long as such computer is secure and accessible only to the NAACCR, Inc. researcher.

7. NAACCR, Inc. will not take any action that will provide any Data furnished by REGISTRY to any unauthorized individual or agency or any other third party without the prior written consent of REGISTRY.

8. NAACCR, Inc. will not disclose in any manner, to any unauthorized person, information that would lead to identification of individuals described in the Data furnished by REGISTRY. Also, NAACCR, Inc.

will not provide any computer password or file access codes which protect the Data to any unauthorized person.

9. Should NAACCR, Inc. become aware of any unauthorized access or disclosure of the Data to other persons, NAACCR, Inc. will report it immediately to REGISTRY.

10. In the event that any attempt is made to obtain from NAACCR, Inc. any or all of the Data provided to NAACCR, Inc. by subpoena or other legal means, NAACCR, Inc. will notify REGISTRY immediately. NAACCR, Inc. agrees that REGISTRY may employ attorneys of its own selection to appear and defend the claim or action on behalf of REGISTRY. REGISTRY, at its option, shall have the sole authority for the direction of the defense and shall be the sole judge of the acceptability of any compromise or settlement of any claims or action against REGISTRY.

11. NAACCR's obligations hereunder shall remain in full force and effect and survive the completion of NAACCR's Call for Data projects described in Attachment A.

12. The terms of this Assurances Agreement shall be binding upon NAACCR, Inc. his/her agents, assistants, and employees.

13. Notwithstanding any contrary language in this Assurances Agreement, NAACCR, Inc. acknowledges and agrees that NAACCR's access to the Data shall at all times be in the sole discretion of REGISTRY.

14. REGISTRY reserves the right to review any and all of NAACCR's reports prior to dissemination or NAACCR's manuscripts before submission for publication to ensure that confidentiality is not violated and the Data are used appropriately.

15. This Assurances Agreement shall be governed by and interpreted under the laws of the State of Illinois.

Dated this 30th day of August 2022.

North American Association of Central Cancer Registries, Inc.

By:

Its: Executive Director

Print Name:

North American Association of Central Cancer Registries, Inc.

REGISTRY

By: _____

Its: _____

Print Name: _____

Address: _____

_____

_____

E-mail address: _____

Contact Recinda Sherman with questions at: rsherman@naaccr.org

Uses of Registry Data Submitted to the NAACCR Call for Data

1 Summary of Primary and Secondary Data Uses

Central registry data submitted to NAACCR, Inc. in the Call for Data include data from 1995-2020, as of December 6, 2022, and 2021 as late as January 31, 2023. Primary data use activities do not require additional consent from registries. Secondary data use activities do require additional, project-specific consent.

1.1 Cancer in North America Primary Use (registry data included in these activities for all registries based on signed, general DUA except where specific consent is noted)

1.1.1. Produce Cancer in North America (CiNA), 2016-2020 monographs (Vol 1: Combined Incidence; Vol 2: Registry-Specific Incidence, Vol 3: Registry-Specific Mortality, Vol 4: Survival, Vol 5: Prevalence, Vol 6: Population Attributable Risk Factors, and associated appendices).

1.1.2. NAACCR evaluations of 1995-2021 data to determine fitness for use in cancer surveillance and research projects and make assessments available on the NAACCR website and to the NAACCR community, e.g. 12-month data assessment, completeness of variables.

1.1.3. Produce the Annual Report to the Nation on the Status of Cancer (high quality U.S. registries only).

1.1.4. Create surveillance information and respond to data requests using the NAACCR CiNA research datasets for requests of aggregate data of high-quality U.S., Canadian, or North American data; state/provincial/territorial-specific data; or cancer site-specific data using a suppression rule for fewer than six cases for any requested rates and counts by state, province or territory.

1.1.5. Create aggregated measures by central registry for use in the NAACCR web-based public query systems available here https://www.naaccr.org/interactive-data-online/.

1.1.6. Development of datasets to create project-specific datasets for proposals approved by the NAACCR Research Application Review Workgroup (RApR) (consent from registries required prior to release of any data).

1.1.7. Utilize data internally within NAACCR and by NAACCR committees to assess quality, fitness for use, and appropriate methodologic approaches.

1.1.8. Use of CiNA Research datasets to support the ranking by state and province based on cancer-related health indicators to support comprehensive cancer control.

1.1.9. Produce a non-confidential, Public Use dataset with limited variables to be available upon request after signing a data assurances agreement. Inclusion in this dataset will require registry consent as well as meeting data quality criteria. Consent for this specific project is included in this document.

1.1.10. Produce an historical dataset annually to create delay factors based on multiple submission datasets. Consent for this project is included in this document.

1.1.11. Produce a new CiNA Geographic dataset, for US data only, which includes census tract, for the purposes of data quality and methodology evaluation by NAACCR only. These data will never be released without specific consent from the registries.

1.2 Cancer in North America Secondary Uses (consent requested for each specific project)

The NAACCR website describes how data releases are approved, with registry consent, and summarizes the variables available for each use; the data release procedure, and steps to ensure patient confidentiality: https://www.naaccr.org/cina-data-products-overview/. The NAACCR Data Request Tracking (DaRT) System tracks all data requests, data release, and associated processes: https://apps.naaccr.org/dart.

1.2.1 Create project-specific datasets from the 1995-2020 CiNA (e.g. CiNA Research, CiNA Survival/Prevalence) and special CiNA Research datasets for researchers--consent for inclusion will be sent to registries as projects are approved.
1.2.2 Produce CiNA Research dataset for calculation of incidence projections by the American Cancer Society (ACS) for their annual *Cancer Facts and Figures* publications (Active Consent attached)
1.2.3 Provide aggregated data for medullary thyroid cancer verification (Active Consent attached)
1.2.4 Produce CiNA Public Use datasets in SEER*Stat for Public Use. Registry inclusion is dependent upon standard data quality criteria as well as individual registry consent (Passive Consent attached)
1.2.5 Produce CiNA Research dataset for American Lung Association Annual Report (Passive Consent attached)
1.2.6 Produce an historical dataset annually for NCI collaborators to conduct delay-adjusted incidence rates (Passive Consent attached)
1.2.7 Produce CiNA Research dataset for evaluation of stage at diagnosis and impact of the Affordable Care Act by the American Cancer Society (Passive Consent attached)

Registry Certification Program

Diagnosis year 2020 will be used for Registry Certification. This involves an evaluation of a registry's data to determine whether they meet NAACCR's high-quality standards for use in computing incidence statistics. The Certification Committee reviews results annually. NAACCR Executive Office conducts the evaluation.

Physical and Electronic Data Security

Certificate of Confidentiality

The use of CiNA data is covered by a Certificate of Confidentiality (Certificate) that protects the privacy of research participants enrolled in research. The Certificate prohibits disclosure in response to legal demands, such as a subpoena. Effective October 1, 2017, NIH no longer provides documentation that specific NIH-funded studies are covered by a Certificate but NIH funded research activities are automatically issued a certificate under the NIH Policy on Certificates of Confidentiality. This has been confirmed with a Human

Subjects Protections Consultant from the NIH Office of Extramural Research. More information is available here: https://humansubjects.nih.gov/coc/faqs#definitions. If you have any questions, please contact NAACCR Program Manager of Data Use & Research (Recinda Sherman at rsherman@naaccr.org).

File Submissions

All files are submitted to the NAACCR Statistical Analytic Unit, Information Management Services, Inc. (IMS) through secure electronic channels. Annually, IMS provides an assessment of their Data Security processes using the NAACCR document, *Inventory of Best Practices Assurance of Confidentiality and Security*. The data submissions are accessible only by IMS staff under contract with NAACCR, Inc. to process the files and produce the primary analyses of data.

Datasets used for CiNA Products and other primary and secondary uses of data will be deleted after the publication of research or confirmation of end of study. However, historic data will be maintained to support the on-going production of reporting delay-factors.

Standard File Submission Workflow

Registry personnel log in via the MyNAACCR Login service and can only upload files on the NAACCR CFD Portal. The CFD Portal does not offer the ability to download any files by the registry or NAACCR staff.

All data in transmission from a registry to IMS are encrypted using industry standard TLS 1.2 technologies.

Each file received by IMS is automatically encrypted by AES 256 encryption with a unique, randomly generated key per submission year.

When IMS wants to download the files from the NAACCR CFD Portal, a single approved IMS staff member (Rick Firth) runs a script that connects to the web server. The script that the person uses can only be run from within the secure IMS network.

The files are downloaded to a secure, privileged-access directory on an internal IMS server.

During the download process, each encrypted registry file is retained and also unencrypted.

Each file is run through Edits and then IMS' SEER*Recode program and additional fields added on to the data record. An output file in CSV format is created.

Country-specific (US or Canada) SAS programs are run on the registry CSV files that check for invalid records and produce two CSV files:

One containing all the registry exclusion records

One containing all the records to be used

Once all the data are processed, the unencrypted files on the NAACCR CFD Portal are deleted. The encrypted version of each registry data file is retained in case an error is located later in the processing of the data and requires correction. The encrypted registry data files are kept for 5 years and then deleted. The intermediate registry- specific CSV files are deleted once all the data are processed and production on CiNA products has begun. All backups are deleted 12 months later.

Special Handling of Census Tract for CiNA Geographic Dataset (US Only)

All data in transmission from a registry to IMS are encrypted using industry standard TLS 1.2 technologies. Submission files can only be uploaded to the NAACCR Call for Data Portal and are not accessible for download. Each file received by IMS is automatically encrypted upon upload with a unique key per submission year.

When the data are ready to be processed by IMS, a single, approved IMS staff member (Rick Firth) downloads the encrypted registry files to a secure privileged-access directory on an internal IMS server. Each encrypted registry file is retained and then unencrypted, and the data processed.

A limited dataset containing census tract is created for SEER*Stat, separate from the standard CiNA dataset. Once the SEER*Stat file is created, all unencrypted, processing files are deleted. The encrypted, registry submission files that are stored on the secure IMS network are kept for one year and then deleted. All file backups will be deleted 12 months later.

A single, approved NAACCR Staff member (Recinda Sherman) and a single, approved IMS Staff member (Rick Firth) will have access to census tract information. However, the IMS SEER*Stat administrators (Dave Campbell, Don Green, Aaron Hall, Steve Scoppa, Gretchen Flynn) have the capacity to access any NAACCR dataset. However, no IMS personnel will access the CiNA census tract data without authorization from NAACCR via DaRT.

CiNA Production Database Processing Files Workflow

This workflow starts by using the CSV file containing all the records for the CFD submission years (from 2.3.2.d.ii).

Three SAS programs are used to process the CSV files and create resultant CSV files that are acceptable to SEER*Prep. The three SAS programs are for the three different types of databases to be created:

Survival & Prevalence DBs – using the output of the survival SAS program CSV files, a production SEER*Stat database for survival and prevalence is created.

Production DBs – using the output from the incidence SAS program, the CiNA, CiNA In situ, and Certification SEER*Stat databases are produced.

12-month DB – using the output from the incidence SAS program and the 12- month SAS program, a 12-month SEER*Stat database is created with the CFD submission years plus the 12-month data.

During this entire process, all original patient IDs are replaced with a new random patient ID. Other adjustments include adding in custom recodes and adjusting certain fields are performed.

The CSV files used by SEER*Prep for each database type are kept for 5 years. The SEER*Stat databases are kept for 5 years, except for the files required to support the delay adjusted.

SEER*Stat Research Databases Processing Workflow

This workflow starts by using the CSV files created above from step 2.5.2.b.

An additional set of CSV files is created limited by years of high-quality data for each registry.

These CSV files are used by SEER*Prep to create the expanded and NHIA research databases.

The CSV files are used by various SAS programs to create unique new CSV files for custom research SEER*Stat database requests (i.e., add quartiles for a field, breast subtype, non-bridged, etc.).

Public-Use Database: A SAS program is run on the additional set of CSV files to add in recode fields unique to the public use database. The registry-specific CSV files output from this program are used by SEER*Prep to create the public use SEER*Stat database.

All CSV files used by SEER*Prep to create any of these databases are kept for 5 years. The Public-Use SEER*Stat database is kept for 3 years. All research SEER*Stat databases are kept until the research has published or the study closed.

Delay-adjusted Database processing Workflow

This workflow starts by using the CSV files created above from step 2.6.2.

A new set of CSV files is created by a SAS program that adds the delay factors.

The CSV files are used by SEER*Prep to create the production delay database.

The CSV files are used by a SAS program to create an additional set of CSV files for registry-specific requests for a delay database.

The CSV files used by SEER*Prep to create any database in this section are kept for 5 years. All associated SEER*Stat databases are kept only as long as required to support calculation of delay-adjustment factors.

Secondary Datasets

The NAACCR Executive Office provides permission for client server access to the CiNA Research, CiNA Survival/Prevalence, or CiNA special datasets to approved NAACCR researchers. Each dataset is the property of NAACCR, Inc. Data may be exported to other statistical software, such as SAS, for analysis. General information and variables lists for all dataset types are available on the NAACCR Website: https://www.naaccr.org/cina-data- products-overview/.

There are 2 types of CiNA datasets:

The first dataset type is made available only to approved researchers who are NAACCR members, have a NAACCR approved protocol, and have signed a data confidentiality agreement. These include the standard CiNA Research dataset and CiNA Survival/Prevalence and NAACCR is investigating the development of a CiNA Geographic dataset. Datasets with county identifiers, single-years of age, special variables, new methodology, or sensitive research (e.g. HIV) require ACTIVE CONSENT (see section 5) from individual registries to be included in the researcher's specific CiNA dataset. Datasets with the standard 19 age-groups and no county identifiers with well understood methodology and no special variables or sensitive research require PASSIVE CONSENT from individual registries to be included in the researcher's dataset. PASSIVE CONSENT

may also be used for some special variables such as “reside in Appalachian County (Y/N)” or recodes for breast cancer subtypes.

The second dataset is the limited variable, non-confidential, Public Use dataset. The dataset is available upon request after signing a data assurance agreement. The dataset automatically suppresses cells less than 16 and variables are recoded to enable all users to use standard analysis. This dataset does not allow data to be exported outside of SEER*Stat. The Data Assurance Agreement and a list of included variables and recodes are on the NAACCR website: https://www.naaccr.org/cina-public-use-data-set/.

All CiNA datasets created for research will be deleted 1 year after the project is documented as closed or 5 years after known publication date.

Further Explanation of Secondary Uses of Data

A bibliography of peer reviewed publications based on CiNA datasets is available on our website.

Individual CiNA Research Projects

Access to CiNA Research datasets requires a submission of a proposal for review and approval by the NAACCR Research Application Review Work Group (RApR). RApR reviews the applications for scientific merit and appropriateness of using CiNA data. Due to changes in the Common Rule, NAACCR IRB is no longer required. If RApR approves a project, the NAACCR IRB is notified and consent is requested from all registries eligible for participation in each study. A dataset is created for the Researcher that only includes data from registries that consent to include their data through the consent process.

After all approvals are in place, but before receiving access to the dataset, all recipients must sign a Data Confidentiality Agreement for NAACCR Researchers (See Attachment B). The NAACCR IRB monitors all projects annually. Copies of the NAACCR IRB procedures, forms, and meeting minutes are located on the NAACCR Website https://www.naaccr.org/irb-information- for-cina/. CiNA datasets are password protected and may not be accessed by anyone other than approved researchers. All manuscripts for publication resulting from the individual CiNA Research Projects are requested to be reviewed and approved by the NAACCR Scientific Editorial Board before

release. They are also reviewed by the IRB to ensure the researcher publishes the data in accordance with data agreement and approved proposal. When the studies are completed, researcher access to the SEER*Stat dataset is terminated except for Public Use Datasets. Access to the Public Use Datasets is automatically terminated annually after the release of the latest CiNA Public Use dataset—users must sign a new data assurances agreement to gain access to the latest dataset. The NAACCR Data Request Tracking (DaRT) System will track all data requests, data release, and associated processes: https://apps.naaccr.org/dart.

Assurances of Proper Use of CiNA data by Researchers

NAACCR members and their collaborators that approved for access to CiNA Research datasets must sign the NAACCR Data Confidentiality Agreement, which specifies the proper protection and limitations on use of the data. Recipients of the Public Use Dataset must initial and sign a Data Assurance Agreement.

Registry Consenting

NAACCR employs two approaches to obtain registry consent for *ad hoc* CiNA projects, summarized below. Both approaches are now conducted through the NAACCR DaRT system with a consent request sent the individual designated as the CiNA Approver by the Registry, as well as any alternate registry designates.

Passive Consent: Registries have 14 days to respond. If no response is received, approval is assumed. Projects qualifying for Passive Consent do not request single-years of age, do not request County at Dx, are not unique applications of surveillance data, and do not request special variables that increase the potential for identification of individual patients.

Active Consent: Registries have 14 days to respond. If no response is received, data from the non-responding registry will be excluded from the project. Projects requiring Active Consent may contain single-years of age or County at Dx, or both. Active Consent is also used for projects that are either unique applications of cancer registry data or that request special variables that have the potential to identify specific patients or residential locations of patients. Over time, as registries become familiar with projects and there is wide-spread participation, some Active Consents move to Passive Consents. Two examples

are the Delay Adjustment Project (now qualifies for Passive Consent) and the CiNA Survival Project (which no longer requires consent because it is now included in CiNA Primary Uses).

	Passive Consent Process	**Active Consent Process**
Variable list	Standard Dataset; Standard +area-based socioeconomic variables--county or tract-based poverty or urban/rural data ***without*** County (requires specific researcher justification)	Customized Request (e.g. Single Year of Age, County Identifier, county or tract-based socioeconomic variables, such as poverty or urban/rural status, released along with a County)
Geographic Presentation	United States, Canada, North America, regional, state-level analysis.	County
Linked Special Geographic Variables	County-level collapsed data, i.e. Appalachian Region Y/N, CHSDA region Y/N), coded economic or other SES data that does not uniquely identify a county or tract; data appended at the state-level is allowed	Any area-level data linking to continuous variables or coded data that could uniquely identify county

Rescinding Consent

For all primary uses of NAACCR submissions, a registry director has the opportunity to rescind consent up to the time that the files go into production to produce the various products. Thus, it is important that every registry be familiar with their data before it is submitted.

With regard to special studies (secondary data uses), once the dataset has been produced and released to a researcher, consent **may still be rescinded**. However, a researcher

may have already conducted analysis on the original file or presented/published data. If a registry rescinds consent, we immediately remove the registry's data from the SEER*Stat dataset, instruct the researcher to destroy any exported data for that registry, and instruct the researchers to remove the registry's data from any pending or future presentations/publications.

D2. Data Confidentiality Agreement for NAACCR Researchers

DATA CONFIDENTIALITY AGREEMENT FOR NAACCR RESEARCHERS

Agreement executed this____day of_____, 20___, by and between _____ *(Name)* ("Researcher") of _____ *(City)*, _____ *(State/Province)* and **NORTH AMERICAN CENTRAL CANCER REGISTRIES, INC.** ("NAACCR"), a California corporation. Researcher is engaged in research into the causes, control, or prevention of cancer, specifically described as follows:

Project Title: _____

NAACCR collects and maintains certain research data (the "Data") that will or may assist Researcher in this regard. Researcher agrees and acknowledges that patient confidentiality is of the utmost importance in the use of the Data and in the manner in which all research results are presented and/or published. Accordingly, in consideration of his/her receipt of the Data from NAACCR, Researcher agrees as follows:

1. Researcher agrees to treat the Data received from NAACCR as private, non- public health information. The Data will be used solely for the specified research described hereinabove and not for any other purpose. The Data will never be used as a basis for legal, administrative or other adverse actions that can directly affect any individual about whom personal and/or medical information is included in the Data.

2. Researcher understands and agrees that any and all Data which may lead to the identity of any patient, research subject, physician, other person, or reporting facility is strictly privileged and confidential and agrees to keep all Data strictly confidential at all times.

3. If, in the course of his/her research, Researcher believes it necessary to provide access to the Data to any other individual, Researcher will **NOT** do so unless and until such individual has properly executed a Data Confidentiality Agreement that has been accepted, in writing, by NAACCR. And, Researcher agrees to notify NAACCR in writing within forty-eight (48) hours of his/her becoming aware of any violation of this Confidentiality Agreement or any Confidentiality Agreement executed by any other individual, including full details of the violation and corrective actions to be taken by Researcher.

4. Researcher further agrees that all data provided under the provisions of this Data Confidentiality Agreement may only be used for the purposes described hereinabove, and that any other or additional use of the data may result in immediate termination of this Confidentiality Agreement by NAACCR.

5. Researcher agrees that (i) any and all reports or analyses of the Data prepared by Researcher shall contain only aggregate data. Researcher further agrees that (ii) at no time will he/she ever publish any individual names or other personally identifying information or information which could lead to the identification of any Data subject, and (iii) no report of the Data containing statistical cells with less than six (6) subjects shall be released without the prior written

authorization of NAACCR's Executive Director, who has received written authorization from contributing registries.

6. Researcher agrees that linkage to another database is not permitted for the purpose of identifying an individual on the file, but may be permitted if appropriate linkage is described in the proposal and this linkage is approved by the NAACCR IRB.

7. Researcher further agrees that all data provided under the provisions of this Confidentiality Agreement shall remain the sole property of NAACCR and may not be copied or reproduced in any form or manner without NAACCR's prior written consent.

8. Researcher shall indemnify NAACCR from any and all liability, loss, or damage (including attorneys' fees) suffered as a result of claims, demands, costs or judgments arising out of the failure of Researcher or those acting in connection with Researcher to conform to and obey the provisions of this Data Confidentiality Agreement. In the event a claim should be brought or an action filed against NAACCR in connection with any such failure, Researcher agrees that NAACCR may employ attorneys of its own selection to appear and defend the claim or action on behalf of NAACCR, at the expense of Researcher. NAACCR, at its option, shall have the sole authority for the direction of the defense and shall be the sole judge of the acceptability of any compromise or settlement of any claims or action against NAACCR.

9. Researcher will not take any action that will provide any Data furnished by NAACCR to any unauthorized individual or agency without the prior written consent of NAACCR.

10. Researcher will not discuss in any manner, with any unauthorized person, information that would lead to identification of individuals described in the Data furnished by NAACCR. Also, Researcher will not provide any computer password or file access codes that protect the Data to any unauthorized person.

11. Should Researcher become aware of any unauthorized access or disclosure of the Data to other persons, Researcher will report it immediately to NAACCR's Executive Director. Researcher understands that failure to report violations of confidentiality by others shall be considered as Researcher's own violation and may result in civil or criminal penalties and termination of current and future access to confidential data.

12. In the event that any attempt is made to obtain from Researcher any or all of the Data provided to Researcher by NAACCR by subpoena or other legal means, Researcher will notify NAACCR immediately. Researcher agrees that NAACCR may employ attorneys of its own selection to appear and defend the claim or action on behalf of NAACCR. NAACCR, at its option, shall have the sole authority for the direction of the defense and shall be the sole judge of the acceptability of any compromise or settlement of any claims or action against NAACCR.

13. Researcher's obligations hereunder shall remain in full force and effect and survive the completion of Researcher's research project described hereinabove.

14. The terms of this Confidentiality Agreement shall be binding upon Researcher, his/her agents, assistants and employees.

15. Notwithstanding any contrary language in this Confidentiality Agreement, Researcher acknowledges and agrees that Researcher's access to the Data maintained by NAACCR shall at all times be in the sole discretion of NAACCR.

16. NAACCR reserves the right to review any and all of Researcher's reports prior to dissemination or Researcher's manuscripts before submission for publication to ensure that confidentiality is not violated and the Data are used appropriately.

17. Researcher understands that access to the Data will be terminated when the report is submitted to the NAACCR Scientific Editorial Board or on May 1, the release date of an updated NAACCR analytic file, whichever is sooner. However, the researcher may request in writing an extension to access the Data.

18. If Researcher is required by any other party or parties, including the state or a state agency, to execute any additional confidentiality agreement(s) as a condition of access to the Data, in the event of a conflict between the provisions of such agreement and this Agreement, Researcher agrees that the most restrictive agreement shall prevail.

19. This Confidentiality Agreement shall be governed by and interpreted under the laws of the State of Illinois.

Dated this____day of_____, 20___.

Researcher _____ ("Researcher" Signature)

_____ (Print Name)

Address _____

_____

E-mail address_____

Phone: (_____)_____ ext. ___

Received and accepted this___day of_____, 20___.

North American Association of Central Cancer Registries, Inc.

By: _____

Its: _____

Data User Agreement for Access to the NAACCR CiNA (Cancer in North America) Dataset

Current Year: 2023 **Data Years Requested:** 1995-2019

These data are provided for the sole purpose of statistical reporting and analysis only. By using these data, you signify your agreement to comply with the following:

---

1. ☑ I will inform the Manager of Data Use and Research if I identify any issue with the data.
2. ☑ I will acknowledge the NAACCR CiNA Dataset file in any publications or presentations resulting from this study.
3. ☑ I will submit a copy of any journal articles to the NAACCR Scientific Editorial Board via the Manager of Data Use and Research prior to submission. I will ensure to fully describe any limitations or issues with the data or interpretation that were identified as part of the application review process.
4. ☑ I will send a courtesy copy of all publications using CINA data to NAACCR the Manager of Data Use and Research.
5. ☑I will adequately protect the rights and welfare of the data subjects.
6. ☑ I clearly and fully understand the risks to subjects, and the risks are outweighed by the importance of the knowledge to be gained.
7. ☑ Any proposed changes in analysis plan will be reported to NAACCR. I will not initiate these changes without NAACCR review and approval except if necessary to eliminate apparent immediate harm to the subjects.
8. ☑ I have reviewed and agree to comply with all federal, state, and local laws, rules, regulations, policies, and procedures related to the protection of human subjects.

My signature indicates that I agree to comply with the above-stated provisions.

First Name	Last Name	Organization

Last updated: 22 March 2024
Compiled by the NAACCR Data Security & Confidentiality Workgroup

Phone	Email
Date	Signature

Agreement ID:

D4. Virtual Pooled Registry Templated Data Use Agreement

NAACCR Virtual Pooled Registry Data Use Agreement ("Agreement")

Provider Registry ("Provider"):	Recipient Institution ("Recipient"): Recipient Institution FWA#:
Provider Representative Name: Email:	Recipient Scientist Name: Email:
Agreement Term: Start Date: Date of last signature below End Date (select one): ☐ _____ (amount of time) after the Start Date OR ☐ At end of Project, as defined by: _____ OR ☐ No pre-defined end date	Project Title: Researcher Institution IRB Review Determination: ☐ Human subjects research, non-exempt IRB-Approved Protocol # _____ ☐ Human subjects research, exempt (per 45 CFR part 46) ☐ Not human subjects research

Preamble

The parties acknowledge that, prior to the execution of this Agreement, the Recipient disclosed to the Provider certain data about individuals in a dataset the Recipient intends to use for the Project. These preliminary data may have included various identifiers, including, but not limited to, full name, Social Security Number, date of birth, and medical record number. It is understood that the Provider is charged with handling such identifiers in its regular course of business and will protect the Recipient's data according to the same standards it protects its own data. The disclosure of this preliminary data was made to link Recipient's dataset with data from Provider and provide matched case counts. The parties now agree that the Provider has significant information to assist Recipient with performance of the Project and this Agreement covers the Recipient's use of individual-level data released by the Provider.

Terms and Conditions

1) Provider shall provide the data set described in Attachment 1 (the "Data") to Recipient for the research purpose set forth in Attachment 1 (the "Project"). Provider shall retain ownership of the Data, and Recipient does not obtain any rights in the Data other than as set forth herein.

2) Provider represents that it has full authority to share the Data with the Recipient.

3) By signing this Agreement, Recipient provides assurance that relevant institutional policies and applicable federal, state, or local laws and regulations (if any) have been followed, including the completion of, and compliance with, any IRB review or approval that may be required prior to Recipient's use of the Data. Upon Provider's written request to the Recipient's Contact for Formal Notices identified in the signature block, Recipient shall provide the IRB (or equivalent body) determination letter and protocol that is the basis for the IRB status determination (approval, exemption, not human subjects) herein referred to as "IRB-Approved Protocol". Recipient also provides assurance that the Data requested represents the minimum information necessary for the Recipient to complete the Project.

4) Recipient shall not use the Data except as authorized under this Agreement. Recipient shall use the Data only for

statistical, scientific, medical research, and public health purposes that are described in the IRB-Approved Protocol (and revisions thereof) and any application(s) requesting Provider Data. Usage of the Data outside the original IRB-Approved Protocol, and any associated data request application(s), will require review by Provider for amendment of this Agreement as appropriate.

5) Recipient agrees to use the Data in compliance with all applicable laws, rules, and regulations, including but not limited to 45 CFR part 46, as well as all professional standards applicable to such research.

6) Recipient will not use the Data, or permit others to use the Data, either alone or in concert with any other information, to contact individuals who are the subjects of the Data without written permission from the Provider. If Recipient contacts the subjects of the Data without having used the Data to identify these subjects, for example as part of routine activities, the Data may not be used to change what is communicated to the subjects unless written permission is obtained from the Provider.

7) Recipient is encouraged to make publicly available the results of the Project. The Recipient will provide a list of all publications that used the Data. All reports or analyses of the Data prepared by Recipient shall contain only aggregate data. At no time will Recipient publish any individual names or other personally identifying information or information which could lead to the identification of any Data subject. In presentations, manuscripts, or other public disclosures where registry- or state-specific data are presented, the following requirements apply: 1.) statistical cells containing less than six (6) subjects (or as otherwise specified by the Provider) must be suppressed; and 2.) the Recipient shall furnish Provider with a copy of the proposed presentation or manuscript prior to publication or use. Provider shall have thirty (30) days from receipt of a manuscript and fourteen (14) days from receipt of a presentation to disapprove such proposed presentation or submission for publication if Provider believes there is identifiable information that needs protection. In this event, Recipient may proceed with the manuscript or presentation only after addressing Provider's concerns regarding identifiable information.

8) Recipient agrees to acknowledge the contribution of the Provider as the source of the Data in all written, visual, or oral public disclosures concerning Recipient's research using the Data, as appropriate in accordance with scholarly standards and any specific format that has been indicated in Attachment 1.

9) The Data will be used solely by Authorized Persons that have a need to use, or provide a service in respect of, the Data in connection with the Project. Authorized Persons include:

   a) Recipient Personnel: Recipient Scientist and Recipient's faculty, employees, fellows, students, agents, and/or contractors.

   b) Collaborator Personnel: Faculty, employees, fellows, students, agents and/or contractors of an institution involved in the conduct of the Project, as listed in the IRB-Approved Protocol, that have executed an agreement with Recipient that is substantially similar to this Agreement.

10) Except as authorized under this Agreement or otherwise required by law, Recipient agrees to retain control over the Data and shall not disclose, release, sell, rent, lease, loan, or otherwise grant access to the Data to any third party, except Authorized Persons. Recipient agrees to keep the Data in a secured environment with appropriate administrative, technical, and physical safeguards to prevent unauthorized use of or access to the Data and to maintain appropriate control over the Data at all times.

11) The Data is subject to the Federal Privacy Act of 1974, as amended, at 5 U.S.C. § 552a and the Data is covered under a Certificate of Confidentiality, which must be asserted against compulsory legal demands, such as court orders and subpoenas for identifying information or characteristics of a research participant. See https://grants.nih.gov/grants/guide/notice-files/NOT-OD-17-109.html for further information.

12) Recipient acknowledges that Data is governed by Provider's state law which likely prohibits release of Data in response to subpoena in addition to the restrictions present in Term 11. If Recipient believes it is required by law

or legal process to disclose the Data, it will promptly notify Provider, prior to such use or disclosure and will disclose the least possible amount of Data necessary to fulfill its legal obligations.

13) In the event Recipient becomes aware of any use or disclosure of the Data not provided for by this Agreement, Recipient shall take any appropriate steps to minimize the impact of such unauthorized use or disclosure as soon as practicable and shall notify Provider of such use or disclosure within 24 hours of discovery of the unauthorized use or disclosure. Recipient shall cooperate with Provider to investigate, correct, and/or mitigate such unauthorized use or disclosure. Recipient acknowledges that Provider may have an obligation to make further notifications under applicable laws or regulations and shall cooperate with the Provider to the extent necessary to enable Provider to meet all such obligations.

14) Unless stipulated otherwise herein, Recipient is permitted to share a dataset including study-acquired information augmented with elements of the Data ("Augmented Dataset") with parties other than the Authorized Persons described in Term 9 above (the "Secondary Recipient(s)"), provided conditions a-d below are all met. The ability to share an Augmented Dataset with Secondary Recipients aligns with the NIH Data Management and Sharing Policy requirements (https://sharing.nih.gov/data-management-and-sharing-policy/about-data-management-and-sharing-policies/data-management-and-sharing-policy-overview.

a) The Augmented Dataset has been properly "de-identified" pursuant to 45 CFR 164.514(b)(2)(i);

b) The identity of the subjects of the Data cannot be readily ascertained from the Augmented Dataset and the Secondary Recipient will not attempt to learn the identity of the subjects of the Augmented Dataset;

c) The Secondary Recipient:

(i) is a controlled access repository (e.g., dbGAP) which reviews data access requests, controls and audits dataset access, and requires a data use agreement that prohibits repository recipients from further data sharing; or

(ii) has executed a flow-through data use agreement with Recipient that requires Secondary Recipient to abide by the terms and conditions of this Agreement in the same manner as Recipient, but prohibits Secondary Recipient from further data sharing, other than with a controlled access repository which reviews data access requests, controls and audits dataset access, and requires a data use agreement that prohibits repository recipients from further data sharing.

e) Recipient will submit an annual report to the Provider of any new or updated data sharing with Secondary Recipients. The report will include the name of the Recipient Personnel who authorized sharing, the Secondary Recipient name, title, address and associated organization or repository to which data were shared, the date the data use agreement was executed, the date of data sharing, the specific purpose for which the data will be used, and a list of the Data elements obtained from the Provider that were included in the Augmented Dataset. The following Provider-specific requirements, if any, must be applied to the sharing of elements of the Data:

Instructions to the drafter (completed by Provider); delete after completion of this section.

This section should include Provider-specific requirements for sharing elements of the Data with Secondary Recipients (e.g., removing data elements that Provider does not allow to be released, removing a unique patient ID that can be linked to identifiable data, cell suppression requirements, limitations on what type

15) Recipient agrees to securely destroy or return the Data, as directed by the Provider in Attachment 1, at the earliest

time at which destruction or return can be accomplished, consistent with the purpose of the Project. In all cases, Recipient shall destroy or return the Data upon expiration or termination of this Agreement. Notwithstanding the above, Recipient may securely retain one (1) copy of the Data in accordance with Term 16.

16) Unless terminated earlier in accordance with this section or extended via a modification in accordance with Section 21, this Agreement shall expire as of the End Date set forth above. Either party may terminate this Agreement without cause with thirty (30) days written notice to the other party's Authorized Official as set forth below. Provider may terminate this Agreement immediately upon breach by Recipient of any material provision of this Agreement. Upon expiration or early termination of this Agreement, Recipient shall follow the disposition instructions provided in Attachment 1, provided, however, that Recipient may retain one (1) archival copy of the Data in a secure manner as necessary to comply with the records retention requirements under any law, and for the purposes of research integrity and verification. Data retained in this manner shall be subject to the data security and confidentiality terms of this agreement so long as Recipient is in possession and/or control of the Data.

17) Except as provided below or prohibited by law, any Data delivered pursuant to this Agreement is understood to be provided "AS IS." PROVIDER MAKES NO REPRESENTATIONS AND EXTENDS NO WARRANTIES OF ANY KIND, EITHER EXPRESSED OR IMPLIED. THERE ARE NO EXPRESS OR IMPLIED WARRANTIES OF MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE, OR THAT THE USE OF THE DATA WILL NOT INFRINGE ANY PATENT, COPYRIGHT, TRADEMARK, OR OTHER PROPRIETARY RIGHTS.

18) Except to the extent prohibited by law, the Recipient assumes all liability for damages that may arise from its use, storage, disclosure, or disposal of the Data. The Provider will not be liable to the Recipient for any loss, claim, or demand made by the Recipient, or made against the Recipient by any other party, due to or arising from the use of the Data by the Recipient, except to the extent permitted by law when caused by the gross negligence or willful misconduct of the Provider. No indemnification for any loss, claim, damage, or liability is intended or provided by either party under this Agreement.

19) Neither party shall use the other party's name, trademarks, or other logos in any publicity, advertising, or news release without the prior written approval of an authorized representative of that party. The parties agree that each party may disclose factual information regarding the existence and purpose of the relationship that is the subject of this Agreement for other purposes without written permission from the other party provided that any such statement shall accurately and appropriately describe the relationship of the parties and shall not in any manner imply endorsement by the other party whose name is being used.

20) This Agreement and Attachment 1 (Project Specific Information) embody the entire understanding between Provider and Recipient regarding the transfer of the Data to Recipient for the Project.

21) No modification or waiver of this Agreement shall be valid unless in writing and executed by duly-authorized representatives of both parties.

22) The undersigned Authorized Officials of Provider and Recipient expressly represent and affirm that the contents of any statements made herein are truthful and accurate and that they are duly authorized to sign this Agreement on behalf of their institution.

Agreement ID:

By an Authorized Official of Provider:	By an Authorized Official of Recipient:
_____ _____ Date Name: Title: Email: <u>Contact Information for Formal Notices:</u> Name: Address: Email: Phone:	_____ _____ Date Name: Title: Email: <u>Contact Information for Formal Notices:</u> Name: Address: Email: Phone: This data use agreement has been read and understood by Recipient Scientist, who is responsible for ensuring compliance with these terms and conditions, and for communicating these obligations to Recipient Personnel who will handle the Data. _____ _____ Date Name: Title: Email

Attachment 1

Data Use Agreement

1. Summary Description of Project (completed by Recipient)

Instructions to the drafter; delete after completion of this section.

This section of this attachment should provide sufficient information such that each party understands the Project that the Recipient will perform using the Data. Content of this section will be very similar to the Statement of Work used in other types of Agreements. Examples of information that should be provided include:

* *Objective or purpose of the Recipient's work*
* *A general description of the actions to be performed by the Recipient using the Data and possibly the anticipated results*

2. Summary Description of Data Released by Provider (completed by Recipient, reviewed by Provider)

Instructions to the drafter; delete after completion of this section.

This section of the attachment should provide sufficient information such that each party understands the information requested by Recipient and released by the Provider under this Agreement. Examples of information that should be provided include:

* *Description of the data (e.g., diagnosis years, type of cancer, etc.) released by the Provider.*
* *General overview of the type of data elements (e.g., tumor characteristics, treatment characteristics, etc.) included in the data to be released by the Provider.*

3. Expected frequency for recurring linkages, if applicable (completed by Recipient)

Instructions to the drafter; delete after completion of this section:

If the Project involves linkage of a study file with the Provider Data, describe how frequently the Recipient will provide a study file for recurring linkage. If linkages do not recur, write "None".

Agreement ID:

4. Data Transmission and Provider Support (completed by Recipient)

Provider shall electronically transmit Data to Recipient via a secure mechanism. Upon execution of this Agreement, Provider will contact Recipient individual below to coordinate Data transfer.

Name:	
Address:	
Email:	
Phone:	

5. Technical Requirements for Data Transmitted to Recipient (completed by Provider)

Instructions to the drafter; delete after completion of this section.

This section of this attachment should provide sufficient information such that each party understands the technical requirements of the data to be supplied to the Recipient. Examples of information that may be appropriate to include in this section are:

* *Format of Data*
* *Provision of Data dictionary*
* *Availability of Provider to assist Recipient in understanding the Data structure (e.g. variables, code lists, etc.)*
* *If/how Provider will address Recipient concerns and questions about the Data*

6. Publication Requirements (completed by Provider)

In order to simplify cell size suppression requirements across registries, Term 7 of this Agreement requires Recipient to suppress state-specific or registry-specific cell sizes containing less than six (6) subjects. Please indicate whether the Provider accepts the standard suppression rule as written:

☐ Yes, Provider accepts the standard cell size suppression rule of less than six (6) subjects as stated in Term 7 of this Agreement.

☐ No, Provider requires the following alternative cell size suppression rule:

Enter cell size suppression rule here:

7. Acknowledgement Statement for Publications (completed by Provider)

In order to simplify the list of acknowledgements in publications including data from numerous central cancer registries, the following standard acknowledgement is proposed (and has been approved by the referenced Federal funding organizations):

Agreement ID:

"The authors would like to acknowledge the contribution to this study from central cancer registries supported through the *Centers for Disease Control and Prevention's National Program of Cancer Registries (NPCR)* and/or *the National Cancer Institute's Surveillance, Epidemiology, and End Results (SEER) Program.* Central registries may also be supported by state agencies, universities, and cancer centers. Participating central cancer registries include the following: <list of participating registries>.

Please indicate what acknowledgement statement the Provider requires the Recipient to use in publications:

☐ Provider requires the standard acknowledgement (stated above).

☐ Provider requires that the standard acknowledgement (stated above) be used for publications that DO NOT INCLUDE state-specific or registry-specific results. For any publications that DO INCLUDE state-specific results the following acknowledgement is required:

Enter acknowledgement here:

☐ Provider requires the following acknowledgement statement to be used for all publications, regardless of whether state-specific or registry-specific results are included:

Enter acknowledgement here:

8. Data Destruction and/or Retention Requirements (completed by Recipient, reviewed by Provider)

Instructions to the drafter; delete after completion of this section. *This section of this attachment should include sufficient information such that each party understands the Recipient's obligations with regards to the destruction or return of the Provider Data at the earliest time possible consistent with the purpose of the Project or upon the expiration or early termination of this Agreement. Please be sure to specify the methods and timeline for data destruction.*