North American Association of Central Cancer Registries, Inc. (NAACCR)

Discharge and Claims Data Best Practices Guide

November 2023

Acknowledgments

Thank you to the North American Association of Central Cancer Registries (NAACCR) Discharge and Claims Data Task Force for their dedication and hours in preparing this document.

This activity is supported in part by Cooperative Agreement Number 5NU58DP006458 from the Centers for Disease Control and Prevention. Its contents are solely the responsibility of the authors and do not necessarily represent the official views of the Centers for Disease Control and Prevention.

Suggested Citation: NAACCR Discharge and Data Claims Task Force (eds). *Discharge and Claims Data Best Practices Guide.* Springfield, IL: North American Association of Central Cancer Registries, September 2023

Discharge and Claims Data Task Force Roster

Jenna Deniaud (Chair)
Wisconsin Cancer Reporting System
jenna.deniaud@dhs.wisconsin.gov

Carrie Bateman
Utah Cancer Registry
carrie.bateman@hsc.utah.edu

Angela Eckstrand
Alberta Cancer Registry
angela.eckstrand@albertahealthservices.ca

Jennifer Hafterson
Fred Hutchinson Cancer Center
jhafters@fredhutch.org

Kim Herget
Utah Cancer Registry
kim.herget@hsc.utah.edu

Megsys Herna
Florida Cancer Data System
mherna@med.miami.edu

Annette Hurlbut
Elekta
annette.hurlbut@elekta.com

Sandy Kwong
California Cancer Registry
sandy.kwong@cdph.ca.gov

Brent Mumphrey
Louisiana Tumor Registry
bmumph@lsuhsc.edu

Susana Perez
Texas Cancer Registry
susan.perez@dshs.texas.gov

Valentina Petkov
National Cancer Institute
petkovvi@mail.nih.gov

Chris Poon
Fred Hutchinson Cancer Center
wpoon@fredhutch.org

Devbani Raha
Nova Scotia Health Cancer Care Program
devbani.raha@nshealth.ca

Tasha Robbins
trtrobbins@msn.com

Katrina Swett
Connecticut Tumor Registry
katrina.swett@ct.gov

Alfreda Woods
District of Columbia Cancer Registry
alfreda.woods@dc.gov

Table of Contents

Introduction

1 Purpose and Scope

The North American Association of Central Cancer Registries (NAACCR) assembled the Discharge and Claims Data Best Practices Guide Task Force, charged with updating the *Discharge and Claims Data Best Practices Guide* (2015). The task force consists of staff from central registries, standard-setting agencies, and software vendors.

The *Discharge and Claims Data Best Practices Guide* is a document that provides information and guidelines for use of hospital discharge data (HDD) and claims data by central cancer registries (Registries) with a focus on informatics and operations. The goal of this guide is to help Registries improve current processes, determine which new processes are feasible, and determine how to implement them.

Every Registry is unique; thus, methods used to obtain and integrate data vary among them according to laws, organizational structures, and resources. With this in mind, topics in this document are covered at a high level. To supplement, stories and examples from Registries are included throughout to demonstrate use cases and provide additional detail. The content of this manual will generally address Registries in the United States, with distinct subsections provided for Canada.

2 Background

The use of HDD and claims data in Registries is made possible through coding system standardizations and electronic transmission of data. These standardizations date back to 1996, when the Health Insurance Portability and Accountability Act (HIPAA) was enacted and, with it, a mandate to establish and utilize national standards for electronic transactions in health care. These standards continue to be updated and maintained and are enforced for all HIPAA-covered entities.

The national standard for transactions is the Accredited Standards Committee X12 (ASC X12) standard format for health insurance transactions. A transaction is defined as an electronic exchange of information between two parties to carry out financial or administrative activities related to health care. Transactions encompass claims and encounters, eligibility inquiries and responses, claim status inquiries and responses, referrals and prior authorizations, and health care payment and remittance advice.

Health transactions utilize standardized code sets. Table 1 includes the current U.S. and Canadian standards for code sets used in hospital discharge and claims data.

Table 1. Standards for Code Sets Used in Hospital Discharge and Claims Data

Activity	Standard
Diagnosis Coding	International Classification of Diseases, 10th edition, Clinical Modification (ICD-10-CM)—United States International Classification of Diseases, 10th edition, Canada (ICD-10-CA)—Canada
Hospital Inpatient Procedure Coding	International Classification of Diseases, 10th edition, Procedure Coding System (ICD-10-PCS) Canadian Classification of Health Interventions (CCI)—Canada
Outpatient Procedures and Physician Services	Current Procedural Terminology (CPT)—United States CCI—Canada
Supplies and Procedures not included in CPT	Healthcare Common Procedure Coding System (HCPCS)—United States

Sources: Adopted Standards and Operating Rules [CMS], Discharge Abstract Database Metadata [CIHI], National Ambulatory Care Reporting System (NACRS) Metadata [CIHI]

2.1 Hospital Discharge Data in the United States

HDD consist of discharge abstracts containing administrative, clinical, and demographic information on hospital discharges. The primary source of HDD is inpatient discharge abstracts. Other sources for discharge abstracts include hospital-affiliated emergency departments, hospital-affiliated ambulatory care sites, and, occasionally, free-standing ambulatory surgery sites.

HDD abstracts include various data items, such as diagnosis and procedure codes, demographics (specific data included varies by state), expected payment source, total charges, admission and discharge status, encounter dates, length of stay, and hospital identifiers. The International Classification of Diseases, 10th edition, Clinical Modification (ICD-10-CM) diagnosis coding system is used in discharge abstracts for inpatient and outpatient medical encounters, and the International Classification of Diseases, 10th edition, Procedure Coding System (ICD-10-PCS) procedure coding system is used for inpatient hospital procedures.

Primary and secondary diagnoses are identified using ICD-10-CM codes. “Present on admission” codes are included with these clinical diagnoses, identifying whether the condition was present at the time of admission to the hospital (comorbidities in Registry abstracts). Comparing these codes with discharge codes allows identifying hospital-acquired conditions, such as infections. Other related data elements, such as discharge status codes, allow determining whether the patient died in the hospital, was released to home, or was transferred to another hospital or facility.

2.2 Claims Data in the United States

Claims data are administrative data generated by a variety of health care facilities and providers for billing clinical services provided to patients. Claims are generated from a variety of sources, including inpatient, ambulatory surgery, emergency department, outpatient, clinic, pharmacy, and physician office services.

Claims data use several different coding systems. The Healthcare Common Procedure Coding System (HCPCS) and the Current Procedural Terminology (CPT) coding are usually coded in the same field and indicate what procedure was provided. The HCPCS codes will typically cover such services as chemotherapy, hormone therapy, or immunotherapy and will start with a letter of the alphabet, whereas the CPT codes will cover such services as radiation, chemotherapy administration, and surgery procedures and are a five-digit number. These coding systems are updated annually as new treatment options are approved and old codes are retired.

For pharmacy claims, the National Drug Code (NDC) is used to identify which drugs are prescribed and in what form and quantity. The Surveillance, Epidemiology, and End Results (SEER) Program developed a Cancer Medications Enquiry Database (CanMED) that maintains a list of cancer-relevant HCPCS and NDCs and how SEER would classify each treatment. SEER is currently developing a similar database to flag International Classification of Diseases, 9th edition (ICD-9), International Classification of Diseases, 10th edition (ICD-10), and CPT codes.

Claims data will typically have three types of elements: member characteristics, member enrollment, and unique claims. Member characteristics can be used to conduct patient-level linkages between the Registry and claims data. Member characteristics variables include such patient identifiers as name, date of birth, and sex, which are important in conducting a linkage. They usually include a Social Security number (SSN), but each state will have different policies regarding the use of SSN in conducting the linkage. Additional demographic data may include race, ethnicity, or residential information.

The member enrollment files will allow a Registry to identify when a cancer patient might have claims available within the claim's files. Because claims data are usually limited in scope, perhaps excluding Medicare claims, it is important for a Registry to recognize that not every cancer patient will be included in the claims data for their time frames of interest. These files include start and end dates of insurance coverage, indicators for primary or secondary insurance, type of insurance, and level of insurance. This file will also indicate if someone belongs to a Preferred Provider Organization (PPO), a Health Maintenance Organization (HMO), or some other type of plan. This is important to note because the claims that are available in these different types of plans will vary. In some states, HMOs are far more common than in other states, but the data available under HMO structures is often more limited in use by Registries because claims are often paid in one lump sum for a combination of treatments, so Registries cannot always identify exactly what individual treatments were provided.

Unique claims data highlights the individual treatments a member receives. This is where Registries can identify the specific types of treatment a patient had. Multiple files may be available for different types of claims (medical claims, pharmacy claims, dental claims, etc.). Each claim that appears in a claims file will be an individual claim for an individual service, so a single encounter with a doctor may produce several unique lines of claims. The claims will feature an event date for when the procedure was provided, the service that was provided, the ICD diagnostic code for the reason the service was provided, and some information about the provider and the cost associated with each claim. Pharmacy claims will include the date the prescription was filled, the drug that was provided, and the amount of medication provided.

2.3 Hospital Discharge Data and Claims Data in Canada

The Canadian Institute for Health Information (CIHI) maintains two commonly used administrative datasets: the Discharge Abstract Database (DAD) and National Ambulatory Care Reporting System (NACRS). DAD contains discharge records for all patients hospitalized in Canada for all provinces and territories except Quebec. NACRS contains records for hospital and community-based ambulatory care, such as day surgery, outpatient clinics, and emergency department visits. Both DAD (inpatient) and NACRS (outpatient) capture demographic, clinical, and administrative information.

DAD and NACRS include various demographic data items, such as name, health insurance number (unique health system identifier), patient mailing address, sex, and date of birth. However, no race information is currently collected or provided, though discussions are underway in Canada on how to collect race/ethnicity information. The databases also include clinical and administrative data items, such as diagnosis codes, procedure codes, length of stay, vital status at discharge, and hospital identifiers. The CIHI website contains full lists of DAD and NACRS data elements.

The ICD-10-CA diagnosis coding system, developed by the World Health Organization and modified for Canadian context, is used in both the DAD and NACRS databases to identify primary and secondary diagnoses. The Canadian Classification of Health Interventions (CCI) procedure coding system is used in both databases to identify the health intervention or service received. The detailed data collected for inpatients and outpatients allow for linkage to existing Registry data to identify whether the discharge information is about a new primary (in either an existing or new patient) or whether the discharge was related to the treatment of an existing cancer (or metastasis).

In Canada, billing or practitioner claims data are based on fee-for-service payments and tend to be more specific to each province/territory. Claims data are not typically used as a data source for Canadian Registries. However, many provinces and territorial Registries have access to electronic medical records or clinical information systems that provide a comprehensive patient record that includes information for medical oncology and radiation treatment.

2.4 Limitations

HDD and claims data are rich sources of data that can provide Registries with information to fulfill their reporting requirements. Even though these data have been used extensively, weaknesses remain that present challenges for accessing, linking, and integrating linked data.

The availability of data sources continues to be an issue, as participation and reporting requirements for HDD, claims data, and Registries differ among jurisdictions. Requirements can range from mandated to voluntary to nonexistent. While most states will have some level of participation of HDD databases, as indicated by the Healthcare Cost and Utilization Project (HCUP), less than half have a similar level of participation of claims data available through an All-Payer Claims Database (APCD).

Data element availability may also vary, even among data sources that mandate participation or have a high degree of volunteer participation. For example, SSN, which is a unique patient identifier, may be available in some state and territorial HDD, but not others. When SSN or other unique patient identifiers are included in HDD, patient linkage has a higher chance of success, and questions related to readmissions can be readily answered because tracking of hospital stays across hospitals can be completed. When only local identifiers, such as hospital-assigned IDs, are available, only readmissions to the same hospital can be tracked. When unique patient identifiers are not available, other data items, such as readmission flags, may be available and are used to indicate whether the specific discharge was for a readmission after a prior hospitalization.

The level of detail for claims data can be affected by reimbursement. For example, a claim for an HMO may have a lower level of detail than a claim for a PPO. This should be considered in states with a high portion of HMO-insured populations. Like HDD, claims data and Registry records have a many-to-many relationship. A patient may have many claims for a single tumor, for example. The level of detail can also be affected by the number of payors. Claims may be submitted through supplemental insurance, which may cause Registries to get only partial information if one payor (such as Medicare or Medicaid) is the primary payor but is not included in the claims dataset but the supplemental insurance is included.

The primary intent of health transactions is payment. Diagnosis (ICD-10-CM) and procedure (ICD-10-PCS and CPT) codes in discharge and claims data impact payment. Registries should consider this when reviewing or utilizing claims data, especially high-severity codes, which garner larger payments. Although it is an illegal practice, Registries may encounter "upcoding" in claims data—a practice in which an inaccurate billing code is assigned to increase reimbursement and enhance revenue.

Both discharge and claims data include comparable information, and the decision about which to use and when will depend on the intended outcome and resources available. For example, claims data, particularly those retrieved through APCDs, include both inpatient and outpatient services and often include actual payment information as well as charges, allowing the determination of the real cost for the payer. However, the varied specificity of coding can result in high yield and low gain. Discharge data, although less robust, do provide information on length of stay and accompanying hospital charges, including cases where payment was made by the patient (identified as "self-pay").

Although both claims data and HDD provide information about the services and payment, registries should review the information presented in this document and carefully consider the best source and methods to reach their intended outcome.

3 Data Acquisition and Access

HDD and claims data availability and data access requirements vary greatly among jurisdictions. These factors, and expected coverage or impact, should be considered prior to requesting data. To determine coverage, registries may estimate the percentage of the population that is expected to be represented by the process. For example, if HDD is mandatory, it may represent a large percentage of the population. Conversely, if it is voluntarily reported and there is low participation from hospitals, a lower percentage of the population would be represented.

Similarly, the volume of data and ability to effectively process and store it should be considered. Large datasets, such as claims data in a state with high coverage, can be exceedingly difficult to process, and the time spent to retrieve, link, and integrate may outweigh the impact. To assist with determining impact, Registries may choose to conduct pilot studies to ensure there is a significant gain and sufficient resources to process the data before full scale-up.

Even when data use agreements are in place, the frequency and method of data transmission must be considered when determining feasibility. Data sources, such as electronic feeds that have a direct interface with Registry Software, may allow for significant automation but may have a high barrier to entry due to the technical requirements. Data sources that use transport mechanisms, such as secure file transfer protocols, may be easy to establish but may be limited in automation capacity. The SEER Data Management System (SEER*DMS) can store data for SEER Registries on Information Management Services (IMS) servers and IMS sets up feeds. Other registries may store data on their organizational servers. In these instances, Registries must consider their data management plan regarding the average

size of files, the frequency of transmission, the way they will be stored, and long-term and archival storage plans.

Both internal Registry staff and Information Technology and Legal teams may be required to establish agreements specifying the transmission, use, release, storage, and destruction of received data. In addition, there is often a cost for the data or a cost to establish the means of transmission or interface.

Some data sources may specify in data use agreements which data elements may be stored and for how long, with specific requirements for data destruction. Transformation of data and integration of data into a consolidated record for privacy and security reasons should be specified in the agreements.

How the Registry is funded and where it operates also impacts access if HDD and claims data are available through the state department of health. If the Registry operates within the department of health, access to the data may be straightforward but may require an internal request process and training. If the Registry is external to the department of health—in a university setting, for example—a more formal request and review process may be required. These processes can take years to develop and may require extensive ongoing review.

3.1 Canada

Discharge data should be available to most Registries in Canada, as DAD and NACRS are maintained at a national level for all provinces and territories except Quebec. Both DAD and NACRS capture demographic information (birth date, sex, health insurance number, vital status at discharge), provider information, admission/discharge information, diagnosis information (indicates the diagnoses, conditions, problems, or circumstances associated with visit), and intervention information. Data element availability and data limitations may vary by each provincial/territorial Registry. Linkages with these databases and corresponding information are used by Canadian jurisdictions differently.

4 Sources of Hospital Discharge Data

HDD may be available through multiple sources, with the most common being the state department of health. Other well-established sources are hospital associations and central distributors, such as the HCUP from the Agency for Healthcare Research and Quality, which is a family of health care databases and related tools for research and decision making. HCUP databases contain a core set of clinical and nonclinical information derived from the data collection efforts of organizations in participating states that maintain statewide data systems. Many HCUP databases contain discharge data or supplement it. The HCUP website may be useful in determining the data source for HDD if not already known.

4.1 Potential Alternative to Hospital Discharge Data—Washington

The Fred Hutchinson Cancer Center's Cancer Surveillance System is the SEER Seattle–Puget Sound Registry. The SEER Seattle–Puget Sound Registry has received electronic Disease Index data from reporting facilities since the late 1980s. The Registry proactively works with health care partners in the reporting area to establish a submission schedule to receive the Disease Index data conforming to the data specifications created and provided during data source development.

Although the SEER Seattle–Puget Sound Registry does not request some of the UB04 data elements of HDD—such as Type of Bill, Patient Discharge Status, Revenue Code and Procedure codes—the Disease Index is considered to be a close cousin to the HDD, containing other key elements that can be used to identify cases, improve follow-up, and identify potential treatment visits at treatment centers. The data elements requested include patient identifiers, service and admission/discharge dates, principal diagnosis code, and all other diagnosis codes associated with each service visit (in ICD-10-CM), as well as

the service location and department information where the patient received care. The service location and department information are not UB04 data elements but have been proven to be a critical quality assurance (QA) indicator to gauge the completeness of the Disease Index during initial data validation. These elements are reviewed to confirm that the Registry is seeing all the service departments one would anticipate for a typical health care facility. The data is also cross-checked with the health care company website to ensure the Registry is seeing all satellite sites and neighborhood clinics or treatment centers where patients are being seen.

Tips on receiving Disease Index submissions

Before negotiation: Create a data file specification and submission schedule. List clearly what you want and how you want it. Make it clear that you would like to receive the finalized diagnosis codes that have passed all internal coding audits and scrutiny.

During negotiation: Hospital registrars working at your local hospitals are excellent advocates because they understand the Registry's intention more than anyone else. They are great resources to navigate you to the right individual to start the negotiation. On the other hand, such health care facilities as stand-alone treatment centers or individual health care providers usually have fewer available resources at their disposal, and some may not have heard of cancer reporting at all. Be prepared to give a good introduction to why you do what you do. Seek to understand their challenges and identify a feasible timeline to create the data feed and establish an ongoing submission schedule. Always automate the submission when you can.

Immediately after negotiation: Maintain one primary and one backup contact who know about the data submission arrangement, ideally including a technical contact who can troubleshoot issues with the data submissions.

Ongoing maintenance after establishment of data feed: Changes are being made to ICD-10-CM codes and updated every October 1 by standard setters. Provide an updated ICD-10-CM reportable case-finding list to your technical contact once a year and have them update the query. This is also an indirect way to ensure your contact is still with the organization or can still support any maintenance work associated with the data feed. Also, in this annual communication, ask the question, "Are there any planned changes to the information system in the upcoming year that would affect the data feed?" In addition, always monitor data submission completeness internally by creating a report showing the received/imported Disease Index volume by admitting month and facility. The sooner a gap or low-volume month is discovered, the much better the chance of quickly fixing the issue and receiving a resubmission for the problematic period.

5 Sources of Claims Data

A common method for Registries to retrieve claims data is through the APCDs. APCDs are large state databases that include medical claims, pharmacy claims, dental claims, and eligibility and provider files collected from private and public payers. Inclusion of Medicare and Medicaid data varies among states. APCD data are reported directly from insurers to states, usually as part of a state mandate or voluntary participation. An Interactive State Report Map indicates their current participation status and includes links to further information about their APCD efforts and any legislation, is continually updated by the APCD council.

Other sources include Medicare and Medicaid, private insurers, pharmacies, ambulatory care centers, claims processors, and data aggregators.

The choice to utilize these sources and the degree to which the data is used will vary based on the data available and the burden required to acquire and integrate it into the Registry database.

Registries should expect strict restrictions and limitations when requesting claims and discharge data. Enforcement of minimum necessary data (as defined under HIPAA) will vary by institutions, which will affect the variables provided. In addition, some institutions may utilize or require privacy preservation measures, such as unique encrypted identifiers. This is especially true for claims processors and data aggregators.

Use of data may be restricted to central Registry operations, and external release of data may require data to be transformed; in other words, the raw data cannot be retained or released for research.

Data retention policies should also be expected and are especially relevant because non-reportable cases may be included in the data.

5.1 All Payers Claims Database

The degree of implementation of APCDs varies between states, ranging from no current activity/effort, existing voluntary efforts, and full implementation with legislation mandating participation. The databases are often accessible and stored within the health department. Inclusion of Medicare and Medicaid varies by state. To access data, the National Cancer Institute may negotiate on behalf of SEER states. For example, Louisiana gets claims data from the SEER*DMS vendor. Registries that are not in a SEER state should identify the data owner and other legal stakeholders to begin talks.

5.2 Medicare and Medicaid

For more information about Medicare and Medicaid claims data, see https://www.cms.gov/Research-Statistics-Data-and-Systems/Research-Statistics-Data-and-Systems.

5.3 Private Insurers

Requesting data directly from the insurance company may be effective if the company covers a large population and is not otherwise represented in other existing datasets. The terms of the agreement may include that claims data will not be retained in native format—rather, a variety of variables will be abstracted from claims data and saved in Registries database, and original claims dataset would then be destroyed.

5.4 Pharmacies

Pharmacy data uses the NDC list published by the U.S. Food and Drug Administration. If a state has just one or two large pharmacies, negotiation and data transmission may be worth it, as the population coverage would be high. Other sources, other than major pharmacies, include pharmacy benefit management administrators.

Common Uses of Discharge and Claims Data in Registry Operations

6 Overview of Common Uses

Common uses of HDD and claims data in Registries include case finding, death clearance follow-back, and case updates.

Patient Identifiers may be the most widely used elements, as they are required for patient linkage, which is a prerequisite to many processes described in this document (see 8.2 Linkage for Death Clearance Follow-Back, for example).

Case updates may be the most intensive, but they are perhaps the most effective because they allow Registries to update and improve NAACCR data elements, include local custom data elements, and populate missing data elements through derivation or direct assignment.

The use of these datasets for research is also significant (e.g., Medicare, Medicaid, APCDs, commercial insurance).

7 Process: Case Finding

Case finding can take many forms, but common uses include operations to determine completeness by reporting facility, new and missed case identification, or follow-back.

7.1 Seattle Example

The Fred Hutchinson Cancer Center (FHCC) receives HDD monthly from all reporting hospitals in the catchment area, as well as some large multispecialty clinics, stand-alone treatment centers, and dermatology offices. All codes on the SEER ICD-10-CM Case-finding List are requested from the reporting facilities. In the FHCC case-finding program, each facility has a list of ICD-10-CM codes determined likely to identify potentially reportable cases not previously identified by real-time pathology case-finding. Facilities with less ePath coverage will have the most ICD-10-CM codes qualifying for further processing in the FHCC case-finding system. Facilities with high ePath coverage will have the fewest ICD-10-CM codes qualifying for further processing.

FHCC has identified the following core ICD-10-CM codes that are processed for most reporting facilities that have good ePath coverage: codes starting with C22, C25, C34, C61, C64, C69, C78, C79, C7B, C91, D32, D42, and D43 and the specific codes D33.2 and D44.2. Stand-alone radiation/systemic treatment centers and hematology oncology offices that have minimal pathology specimens sampled during patient visits will have more qualifying cancer-related ICD-10-CM codes for case-finding—primarily C codes and some D codes. Dermatology offices with in-house dermatopathologists have a separate skin-related cancer subset excluding non-genital basal cell carcinomas/squamous cell carcinomas, which is solely used for preliminary reportability screening to identify the relevant pathology reports that FHCC would request. If underreporting is suspected for either a specific facility or a specific cancer outside of the core codes or a disruption of normal ePath reporting, FHCC can expand the case-finding without asking the reporting facilities for a resubmission. In addition, FHCC uses all received records in linkages to update the patient name, SSN, and date of last contact.

If a Registry wants to pilot HDD case-finding, FHCC suggests obtaining files from facilities with large neurology services, requesting all codes from the SEER ICD-10-CM Case-finding List from those facilities, and limiting the case-finding processing to these core ICD-10-CM codes as a more manageable pilot volume.

8 Process: Death Clearance Follow-Back

8.1 Definition and Use

Death clearance is a process that involves matching registered deaths in a population against cases in the Registry database for two purposes: (1) ascertainment of death information for persons in the Registry (death clearance match) and (2) identification of all deaths with a reportable condition

mentioned as a cause of death that are not found in the Registry database (death clearance follow-back). For detailed information about the death clearance process and requirements, refer to the NAACCR Death Clearance Manual and Quick Reference Guide available on the Registry Operations Guidelines webpage.

HDD and claims data may be used to supplement death clearance follow-back. HDD are used more often; however, claims data can also be useful for identifying additional follow-back sources if a patient was seen in outpatient physician offices or community centers for their treatment, which may be typical for some cancers, such as urological primaries. Registries should consider their ability to identify this need and effectively utilize the data source when determining whether claims data may be useful in their death clearance process.

8.2 Linkage for Death Clearance Follow-Back

Two types of non-matches are identified through death clearance match linkage. The first is a non-match at the patient level. This is when a death certificate contains a reportable cancer, but the patient is not in the Registry database. The second is a non-match at the tumor level. This is when the death certificate contains a reportable cancer, and the Registry has the patient in their database but with a different cancer from what was indicated on the death certificate.

HDD can be linked with death certificates to identify a hospital source for follow-back of non-matches at both the patient and tumor level identified through Death Clearance match linkage. Registries may find this advantageous if they experience low response rates from coroners and medical examiners, which are traditional sources for follow-back because they are indicated on the death certificate. Hospitals that have seen the patient tend to have more resources and more information about the patient, which may yield better response rates and a lower rate of Death Clearance Only (DCO) cases.

Prior to initiating procedures to identify follow-back sources, as stated in the NAACCR Death Clearance Manual, it is recommended that Registries review the text—if available—on the death certificate (DC), verifying the reportable condition. For example, some DCs will list skin cancer as the reportable underlying condition, although the text on the DC may state squamous cell skin cancer, which therefore makes this example a non-reportable condition.

8.2.1 Wisconsin Example

Wisconsin utilizes HDD to identify sources for follow-back on candidate DCOs. The linkage is conducted on two files. Both files contain personally identifiable identification and unique record IDs. No Registry data files are used directly in the linkage with HDD. The linkage to identify hospital sources for follow-back uses two files:

1. A file containing an extract of death records that were determined to be non-matches at the patient or tumor level through the Death Clearance Match linkage. The source is the State Vital Records Mortality File.
2. A file containing HDD (all years). The source is the Hospital Discharge Database maintained by the Wisconsin Department of Health Services, which is a primary source of Wisconsin Hospital Association Discharge Data.

Records that successfully link proceed to the next step, which is identifying a follow-back lead and determining priority for multiple leads. Those that do not successfully link are marked as a DCO. Other states may utilize other methods, as described in the NAACCR Death Clearance Manual, for identifying follow-back sources.

Death records may link to a single hospital's discharge data or multiple hospitals' discharge data. When there is only a single hospital, the decision is simple and highly automated: Follow-back to that hospital.

In the event of multiple hospitals, manual review is conducted by staff members who are involved in data submissions or have a good understanding of Wisconsin hospitals and reporters. During this manual review, reviewers are presented with a patient event history. The patient event history includes information from the Registry database (primary site, diagnosis date, and reporting facility from all abstracts), the death record (death date and causes of death), and HDD and Claims data (encounter dates, hospital identifiers, and ICD-10 codes from discharge data) to establish a timeline of events that can be reviewed to determine which hospital would be best to follow-back with. The use of textual labels alongside coded values from discharge data, death records, and codes for primary site in Registry data is recommended to assist reviewers in selecting a facility for follow-back. Wisconsin has found it especially helpful in visualizing the frequency and recency of encounters. For example, if a patient had been seen at *Hospital A* and *Hospital B,* reviewers may select *Hospital A* if the patient had been seen there for cancer-related care several years ago, even if they were more recently seen at *Hospital B* for a single visit with a discharge status of "Expired."

8.2.2 Connecticut Example

Connecticut uses the Hospital Discharge Database to assist in death clearance. The linkage occurs after the linkage with vital records and a copy of the death certificate is available. A report is created of DCOs in SEER*DMS that contains all patient identifiers, including patient name, medical record number, SSN, date of birth, death certificate file number, admission date, discharge date, place of death, date of death, and address.

An epidemiologist performs linkage to the Hospital Discharge Database using Match*Pro. Details on methods for running SEER*DMS reports and performing linkages can be provided separately.

After manual review of the linkage results, the resulting data are sorted by patient ID, facility ID, and descending discharge date. PDF copies of the death certificate and a letter with the patient information are sent to the facilities. The letter explains the Registry goal of fewer than 1 percent of DCO cases and requests information related to the admission. The facility is asked to review the list and submit an abstract for each reportable cancer. However, if the cancer is not reportable, the facility is asked to make notations on the list or the DC and return both items so that the DC can be removed from the Registry database.

The Registry receives reports from hospitals, convalescent homes, and physicians. An administrative assistant tracks the responses in a spreadsheet. When responses to letters arrive, they are stamped, and dates are logged in the spreadsheet. If all correspondence has been received, it is given to the QA Coordinator for review. If other correspondence is still expected, what has been received is filed in a separate folder marked "waiting for responses." The spreadsheet is updated to reflect the processing of the DC.

The QA Coordinator reviews the completed responses and updates the spreadsheet if the case is ready to be coded. The Processing Technician scans the correspondence/abstract into DMS and assigns it to a Medical Record Technician to code.

9 Process: Case Updates & Follow-Up

9.1 Definition and Use

Additional data elements and improvement of existing data are made possible through integration of HDD and Claims Data. Most commonly, these sources are used to update demographic, follow-up, and death information. They are not currently used to update diagnosis information, because they are not considered precise enough. However, these sources can be leveraged for follow-back efforts.

9.2 Demographic Information

HDD is especially useful for updating demographic information. It can be used to update unknown data items, such as middle initial, SSN, and race, if available. Although the HDD file contains primary payer information, this data item cannot easily be updated without taking into consideration the date of diagnosis and the dates of admission of the HDD.

Insurance is important to hospitals, so virtually all HDD reports include information on this topic. The New Jersey Cancer Registry reported a 50 percent decline in cases coded to "Unknown" for primary payer after linkage with HDD. Some caveats apply to the use of HDD insurance information. HDD allows the inclusion of multiple payers, whereas most Registries limit the number of payers or combinations of payers. Thus, Registries must develop policies about incorporating HDD insurance data into existing abstracts. Because HDD covers all hospital visits during and after a patient's diagnosis, insurance information may include coverage that post-dates the diagnosis. Because Registries are supposed to record primary payer only at diagnosis, they will need to restrict HDD insurance information to entries that match or closely follow the diagnosis date. If HDD includes multiple insurers at diagnosis, determining which is primary is very difficult.

To update demographic information through HDD linkage, two files must be prepared. The first file should contain patient identifiers from the Registry database; the second file should contain the HDD patient identifiers.

Demographic information within matched records may then be compared. Registry fields can be updated using data consolidation rules in place at the Registry. Comparing fields may also require the use of crosswalks. Use of textual labels is recommended to assist with review.

If one patient matches several discharge records and the discharge records have differing demographic information, Registries should follow a procedure for the decision making.

9.3 Diagnosis Information

HDD and claims data from either inpatient or outpatient sources are not precise and should not typically be used to add or update diagnosis data items alone. Inpatient and outpatient data could be used to follow-back to reporting entities for missing cancer cases to improve incidence reporting. Follow-back to different reporting entities from the original source in the Registry database could also yield additional or better information.

9.4 Treatment Information

The claims data and HDD are very broad and can be useful in determining treatments a cancer patient received. The coding systems are detailed enough to help identify the specific treatment given and the date(s) the treatment was received. These data can highlight gaps in the Registry data and can potentially be used to fill in incomplete treatment information from traditional reporting methods. For example, if a hospital reported on a NAACCR abstract that chemotherapy was recommended, but no

follow-up information about the treatment was known, a claim for chemotherapy in the APCD would fill in that gap about when and what chemotherapy was given. These records can be especially useful if a patient was seen at multiple facilities for treatment.

Some of the information found in the claims records is more reliable than others. The primary purpose of these claims records is to transmit information about what care was received in a specific provider's office, so that the provider can be reimbursed for the cost of the care by the insurance company. The accuracy of the date and treatment given is very important to the insurance company, so the dates and CPT/HCPCS/NDC codes tend to be accurate. It is important that every Registry evaluate each new source of any claims data or HDD data, as each source may have different expectations of reporting within their health network.

Although the exact procedure or drug codes on a claim are likely to be accurate, the data need to be interpreted with care. What is considered first course treatment compared to second course or beyond is not easy to identify within the claims data. When processing the claims data, it is important to consider the dates of each claim. If a surgery date is followed immediately by chemotherapy, both would likely be part of first-course treatment, but chemotherapy that starts several months after the initial round of treatment may be indicative of a second course of therapy after testing revealed that the initial treatment was unsuccessful. The types of treatment may also be important. If a patient was receiving one type of chemotherapy for several months and then switched to a different type of chemotherapy, the second chemotherapy may be a second course of treatment. Every cancer site will have a unique standard of care and should be handled according to the typical treatment for that site.

The dates of treatment may also be helpful to flag cases for review of potential recurrence. If a patient was diagnosed with cancer months or years ago, and there is a series of claims for new relevant cancer treatments, that may indicate the patient had a recurrence, especially if no new primary cancers are reported to the Registry.

9.5 Follow-Up/Death Information

HDD can be used to populate or update the date of last contact for passive follow-up. In addition to supplementing the passive follow-up process, it may be used as a quality control measure in identifying patients incorrectly coded as deceased by comparing vital status and date of last contact in the Registry record to the date of last contact and discharge in the HDD.

9.5.1 Connecticut Example

Follow-up is a challenge for patients who were younger than 20 years old at the time of their diagnosis. They are more likely to move out of their area in adulthood and may be diagnosed and/or treated while in college in a different area from their home, and many years may have passed since their diagnosis. SEER Registries must reach at least 90 percent of follow-up on cases diagnosed at age 20 or younger, meaning cases diagnosed at age less than age 20 must have a date of last contact equal to or more recent than the latest diagnosis date for that submission year. HDD can be used to update the date of last contact for Registry cases younger than age 20. These methods can also be applied to other ages and non-SEER Registries. Within your data management system, create a dataset of identifiers for patients who have been lost to follow-up, including the name, SSN, date of birth, gender, all address history, phone numbers, and primary site.

Use the most recent year of the HDD and use Match*Pro for the linkage.

After manual review, export the data as a .csv file so they can be read into SAS or similar statistical software program. The HDD will likely have multiple records per patient. You want the most recent discharge date to update the date of last contact, so sort by discharge date and keep the last record. Example SAS code:

```
PROC SORT DATA=mymatches;BY display_id dischargedate;RUN;
DATA onerowperpatient;SET mymatches;BY display_id
dischargedate;IF last.display_id;RUN;
```

Export the data as a new .csv file to read into your data management system.

10 Common Uses in Canada

10.1 Process: Case Finding in Canada

Many Registries in Canada use HDD to supplement traditional case-finding efforts to identify possible cancer cases, but the process and degree of use varies by each provincial/territorial Registry (PTCR). Table 2 shows the differences between incorporating DAD and NACRS records for case-finding in Alberta, Canada, and in Nova Scotia.

Table 2. Differences in Case-Finding Using DAD and NACRS—Comparison Between Alberta (AB) and Nova Scotia (NS)

PTCR	Source(s)	Frequency of Linkage to Cancer Registry	Sites Considered for Case Finding
AB	DAD	Yearly	Specific cancer sites known to be missing cases from traditional case-finding methods
NS	DAD, NACRS	Monthly	https://seer.cancer.gov/tools/casefinding[†]

[†] DAD/NACRS records with only one diagnosis of cancer where the cancer type is cervical, skin, or history of cancer are excluded from the record linkage because these records led to too many "false positive" findings. Invasive cervical cancers and melanomas of the skin are identified from other data sources received by the NS Cancer Registry.

Typically, cancer cases are identified in DAD and then linked to the Registry to identify patients/tumors not already captured. In Alberta, DAD linkages are completed as a routine year-end procedure for specific cancer sites known to be missing cases from traditional case-finding methods. Once cases not in the Alberta Cancer Registry are identified, Cancer Registrars review information from other sources available, such as the provincial clinical information system and diagnostic imaging, to confirm diagnoses and register records accordingly.

DAD linkages are known to be resource intensive and often identify false cases (cases that are not reportable or already registered by the Registry); however, they can be a reliable case-finding tool for certain cancers, depending on sources of information utilized by each Registry. Because most Registries in Canada do not receive data directly from radiology facilities, the use of DAD in case finding is especially important for certain sites where pathology is not always available (e.g., non-malignant central nervous system tumors where biopsy or surgery is not always possible or recommended).[1]

Identification of cancers with DAD linkages ensures capturing diagnoses that may otherwise be missed. Many Registries have identified DAD linkage as a valuable source of non-malignant central nervous system tumors.[2,3]

10.2 Process: Other Uses in Canada

Linkage of Registries to DAD and NACRS is also commonly used by many Canadian jurisdictions to identify length of hospital stays, treatment information, vital status at discharge, and comorbidities. One example is the Canadian Cancer Treatment Linkage Project pilot, which extracted surgical treatment information by linking the Canadian Cancer Registry with DAD and NACRS.[4] DAD and NACRS are also commonly used for comorbidity data, because they document different comorbid conditions and complications that can be used as a valuable resource for researchers when looking for preexisting conditions or risk factors. Linkages to inpatient and outpatient data sources are often a valuable resource for research studies designed to identify risk factors or improvements to patient outcomes. Researchers should determine the completeness and comparability of these sources because practices may vary by province or territory.

DAD and NACRS compile large amounts of information that Registries and researchers can use to improve the quality and completeness of their information. Practices of use vary by provincial/territorial Registry.

11 Potential and Future Uses

Discharge and claims data have countless other potential uses, such as updates to recurrence and comorbidity data items. Claims data may be used to infer disease progression and recurrence, and discharge data may be used to supplement coding of treatment of the recurrence.

Efforts to use these data in identifying comorbidity are already underway, as are efforts to improve automation, reliability, and consistency.

Augmenting Registry data by abstracting and aggregating data elements from these datasets is another potential use. For example, the SEER Pharmacy Workgroup has developed business rules for populating augmented systemic therapy data items that are separate from the registrar-collected NAACCR systemic data items. The Pharmacy Workgroup is also working on how to make this pharmacy augmented systemic data available in research products. This process allows for scaling up the project to include pharmacy data from additional sources, such as regional and hospital pharmacies.

12 Linkage Methods

Linking patients and linking tumors are the two forms of linkage required when using HDD or claims data. Linking patients is performed first to identify whether the Registry already has the patient in the database. If the patient is not in the database, but the source data indicates cancer diagnosis or treatment, a new case may be added, such as in death clearance. After patient linkage is completed, tumor linkage is conducted to identify new cases for existing patients and/or new tumor information, such as treatment.

The choice to conduct linkages internally or externally depends greatly on software capabilities and Registry resources. Internal linkage for HDD and claims data is not supported in all Registry data management systems (e.g., SEER*DMS, Registry Plus, Rocky Mountain Cancer Data System). For linkages that must be done external to the Registry DMS, Match*Pro currently is the most popular software option. Some software products may include the option of privacy-preserving linkage, which uses encrypted and tokenized records to maintain confidentiality. In-house software may be used, although it

should be used with caution unless the Registry has experts in linkage algorithms and has done rigorous testing and comparisons with software provided by standard setters (e.g., Match*Pro).

12.1 Internal Linkage

These linkages occur within the Registry DMS. An automated set of rules links data to existing records and may consolidate the data, as well. Internal linkage is utilized with constant feeds, such as pharmacy feeds in SEER*DMS. Internal linkage is common when importing new NAACCR records or other source records into a Registry database, when patient and tumor linkage is conducted.

Internal linkages depend on the capability of the Registry DMS. Some Registry DMS software cannot support HDD or claims data internally, in which case external linkage is necessary. Registries should speak with their Registry DMS vendor about the capabilities and options for linkage with these data sources, including subsequent integration.

12.2 External Linkage

These linkages occur outside of the Registry DMS. Some linkage examples that are commonly conducted externally include HDD linkage for follow-back, annual linkages with the Indian Health Service, deduplication of the Registry database, and linkages conducted for research data requests.

External linkages may be beneficial when external linkage software is superior to internal linkage software or when a data use agreement requires it. For example, Utah uses external linkages when the received set of data is likely to contain many non-matches or when a regulatory reason prevents the dataset from being imported into the Registry DMS.

With external linkages, the party conducting the linkage may be the Registry, the data owner, or a third party. Third parties are typically used when there are multiple data sources and when funding is available. After the initial linkage, the Registry may conduct further linkage either externally or internally. When using third parties, finder files may need to be provided by the Registry and other data sources for linkage, and the linkage algorithms used may vary depending on the vendor.

13 Level of Automation in Determining Match Status

Registries should consider the most practical level of automation for each linkage, which will determine the level of post-linkage review required. Factors for choosing a particular level of automation may include the number of records to be linked, the probability of records' linking, the ease of software configuration, and so forth. Manual linkage requires reviewing all linked pairs and is labor intensive. Semiautomated linkages will, by default, classify some matches and non-matches and require a subset for manual review (often called "possible matches"). Fully automated linkages will require no manual review, with the match status decision-making being determined entirely by machine algorithms.

In internal linkages—those conducted within the Registry database management system—semiautomated linkages are most common. Some linkage pairs may be flagged or queried for manual review, whereas others are disposed without any review. In external linkages—for example, those done in such products as Match*Pro or other commercial software—the level of automation can be adjusted from completely manual to nearly automatic, but most Registries default to a semiautomatic level of automation.

14 Preparing Data for Linkage

Preparing the incoming discharge or Claims Data and preparing the Registry data for linkage typically involves an extract, transform, and load procedure. Crosswalks for demographics, diagnosis, and treatment modalities may be needed, as well as annual maintenance of the crosswalks.

When preparing data for linkage, it is important to understand both the data you have received and transformations needed to perform the linkage.

To understand the data, you will need to review it for—

1. Presence of required linking variables (name, date of birth, etc.)
2. Obvious issues with the data (blank values for a variable, misplaced data, corruption, etc.)
3. Variable formatting (e.g., mm/dd/yyyy vs. yyyy/mm/dd)

To transform the data, you will need to—

1. Ensure consistency of format for variables, such as dates.
   a. If the format differs between datasets, one set will need to be modified to match the other.
2. Ensure consistency of content for all variables.
   a. An example: If one dataset has a single variable containing all name components but the other has separate first, middle, and last name variables, you will need to revise one of the datasets to match the other.

15 Crosswalks

Many crosswalks exist to assist Registries in mapping values seen in discharge and claims data with Registry records. Some crosswalks are publicly available. For example, SEER maintains the free-to-use CanMED for NDC and HCPCS code crosswalk to Registry codes, updated monthly. SEER also has crosswalks for CPT codes to Registry codes for radiation treatment, surgery, and systemic treatment, which includes ICD-10, historic ICD-9, CPT, and some HCPCS for radiation and surgery. These crosswalks are not yet publicly available, but Registries may email the Oncology Toolbox team to request them. NAACCR maintains a payer typology crosswalk. Links to these crosswalks and others are provided in Appendix B: Resources.

In addition to publicly available crosswalks, Registries may choose to modify or create their own Registry-specific crosswalks to accommodate special projects or the unique aspects of the Registry or given dataset. The time spent generating and maintaining crosswalks should be factored prior to their creation. Crosswalks should be reviewed at least annually, or more often as processes and dataset changes necessitate.

A crosswalk that maps HDD hospitals to central Registry reporting is advantageous for several operations that utilize HDD, such as follow-back to hospitals during death clearance operations. Registries should review how their provider of HDD assigns and specifies identifiers before establishing a crosswalk. Departments of Health may provide state-specific hospital association IDs and address information in their discharge data. Providers, such as HCUP, may assign several hospital IDs for a given hospital, such as an American Hospital Association ID, HCUP IDs, and other unique identifiers to meet the needs of their partners. Registries are encouraged to connect with state agencies and associations that contribute to or utilize HDD, as supplemental crosswalks may already exist and could be shared to aid the Registry in the creation of their own crosswalk.

When creating crosswalks between HDD hospitals and central Registry reporting facilities, linkage and manual review between HDD hospitals and Registry reporting facilities may be necessary. Because facility-naming conventions may vary across sources, Registries may choose to link based on geolocation of hospitals and conduct manual review on all pairs post linkage. Some HDD providers may restrict the release of hospital IDs or location information, which can ultimately be a roadblock for crosswalk creation and is why registries should evaluate these sources when planning projects.

Large health organizations that umbrella many hospitals are increasingly prevalent. Within a health system, a large metropolitan hospital may be a "parent" facility to several surrounding "child" facilities. Depending on the provider of HDD, these relationships may be indicated in data items, but how the HDD provider defines these relationships may not be consistent with how the Registry defines them for cancer reporting purposes. Therefore, it can be necessary for a Registry to develop a Registry-specific crosswalk between HDD hospitals and parent/child cancer reporting facilities, as shown in Table 3. For example, when using HDD to assist with death clearance follow-back, using such a crosswalk may be advantageous to follow-back with the "parent" facility, which houses the hospital cancer registry and has the resources to report missed incidence cases timely and accurately, rather than the rural "child" facility.

Table 3. Visualization of Crosswalk Between HDD Hospital and Central Registry Reporting Facilities with Parent/Child Relationships

Hospital Discharge Data	***Central Registry Data***	
Hospital	**Facility**	**Parent Facility**
Metropolitan Hospital Campus	Metropolitan Hospital Campus	[Reports for Self]
Farmland Critical Access Hospital	Farmland Critical Access Hospital	[Reported By] Metropolitan Hospital Campus
ABC Community Hospital	ABC Community Hospital	[Reported By] Metropolitan Hospital Campus

Note: Parent facility cells are populated only on rows that contain child facilities. Table structure and use of labels are for demonstration purpose only.

16 Integration

The ability to integrate or consolidate information into the Registry database is one of the most important considerations. Even with a high volume of high-quality data, the ability to integrate will determine its usefulness. Some data element integration can be automated, whereas other elements will require manual review. The level of manual effort or review should be assessed, along with the level of expertise required to manually review.

The level of automation directly impacts the time it takes to integrate data into the Registry database. Integration can be conducted within the Registry DMS manually, or outside the software by executing queries on the database directly, to insert or update data. Regardless of how data are integrated, audit trails and logging are recommended. Registry database software typically has these features built in, but if writing directly to the database outside of Registry software, additional measures may need to be implemented to ensure a log of changes to a record is stored and accessible. This is especially important because some data sources may require Registries to have a process in place to remove the data from

the database record upon request, making a change log and audit trail a technical requirement for removal.

The transformation or processing of data required to integrate must also be considered. Many data sources, such as claims data, may restrict raw data from being integrated but allow data that have been abstracted and transformed to be integrated.

17 Maintenance

The frequency at which linkage is conducted and thus integrated will depend on Registry resources, but the volume expected in each time frame should be considered for annual operation and maintenance purposes. At a minimum, reportability requirements change on an annual basis. Inclusion and exclusion criteria must be modified annually, and feeds or data request specifications updated accordingly.

Coding systems change and are built on over time. Maintaining up-to-date crosswalks and reference tables to coding systems used by HDD and claims data—such as ICD-10-CM, ICD-10-PCS, CPT, HCPCS, and NDC—is imperative for smooth continual operations, so costs to maintain code lists should also be considered.

In addition to codes lists, maintenance to ensure feeds and transmission are working as intended requires resources, as does regular quality control and validation. These technical aspects must be taken into consideration prior to implementing new processes.

Appendix A: References

1. Yuan Y, Amjad S, Eckstrand A, Sevick R, Scott J, Devji S, Bertrand C, King M, Brunka V, Maplethorpe E, Walker E, Muanza T, Davis FG. On Capturing Radiological Diagnoses of Brain Tumours to Provide Complete Population Data in Cancer Registries in Canada. *Journal of Registry Management.* 2018; 45: 167–172

2. NAACCR Discharge Data Work Group (eds). *Discharge Data Best Practices Guide*. Springfield, IL: North American Association of Central Cancer Registries, October 2015.

3. Eckstrand A, Shack L, Pham T, Davis F. The Impact of Hospital Discharge Linkage on Case Ascertainment of Brain Tumours in the Alberta Cancer Registry, 2010–2015. *Journal of Registry Management*. 2018; 45(3): 109–116.

4. Carrière G, Sanmartin C, Murison P, Trudeau R, Trainor C, Pelletier C, Farrar N, Snow A, Bryan S, Newman K. Canadian Cancer Treatment Linkage Project. *Statistics Canada Catalogue no. 11-633-X*, *no. 016*. Ottawa: Statistics Canada, March 27, 2018.

Appendix B: Resources

- Data Security and Confidentiality
- Match*Pro Linkage Software
- SEER Observational Research in Oncology Toolbox Linkage Software: Contains the Cancer Medications Enquiry Database (CanMED) tool with searchable crosswalks and reference tables for National Drug Codes (NDC), Healthcare Common Procedure Coding Systems (HCPCS), and applicable SEER*Rx categories.
- NAACCR Data Standards and Data Dictionary
- Adopted Standards and Operating Rules for Electronic Transactions: Webpage from the HHS containing tables with current standard formats (e.g., ASC X12) and code sets (e.g., ICD-10-CM, CPT, etc.) for electronic health care transactions.
- National Ambulatory Care Reporting System (NACRS) Metadata
- Discharge Abstract Database (DAD) Metadata
- National Association of Health Data Organizations (NAHDO)
- All Payers Claims Database (APCD) Council
- All Payers Claims Database (APCD) Interactive State Report Map

Appendix C: Abbreviations

Abbreviation	Description
AHRQ	Agency for Healthcare Research and Quality
APCD	All-Payer Claims Database
ASC X12	Accredited Standards Committee X12
CanMED	Cancer Medications Enquiry Database
CCI	Canadian Classification of Health Interventions (Canada)
CIHI	Canadian Institute for Health Information
CPT	current procedural terminology
DAD	Discharge Abstract Database
DC	death certificate
DCO	Death Clearance Only
DMS	data management system
FHCC	Fred Hutchinson Cancer Center
HCUP	Healthcare Cost and Utilization Project
HDD	hospital discharge data
HIPAA	Health Insurance Portability and Accountability Act
HMO	Health Maintenance Organization
ICD-9	International Classification of Diseases, 9th edition
ICD-10	International Classification of Diseases, 10th edition
ICD-10-CA	International Classification of Diseases, 10th edition—Canada
ICD-10-CM	International Classification of Diseases, 10th edition, Clinical Modification (Canada)
ICD-10-PCS	International Classification of Diseases, 10th edition, Procedure Coding System (Canada)
ID	identification
IMS	Information Management Services
NAACCR	North American Association of Central Cancer Registries
NACRS	National Ambulatory Care Reporting System
NDC	National Drug Code
PPO	Preferred Provider Organization
PTCR	provincial/territorial cancer registry
QA	quality assurance

Abbreviation	Description
SAS	a statistical software suite
SEER	Surveillance, Epidemiology and End Results Program
SEER*DMS	SEER Data Management System
SSN	Social Security number
UB04	Uniform Medical Billing Form