Data Quality Documentation, Discharge Abstract Database—Multi-Year Information
Our Vision
Better data. Better decisions. 
Healthier Canadians.

Our Mandate
To lead the development and 
maintenance of comprehensive 
and integrated health information 
that enables sound policy and 
effective health system management 
that improve health and health care.

Our Values
Respect, Integrity, Collaboration, 
Excellence, Innovation
# Table of Contents

Abbreviations ................................................................................................................................... iii

1 Introduction ........................................................................................................................................... 1  
  1.1 An Overview of the Discharge Abstract Database ......................................................................... 1

2 Coverage ............................................................................................................................................... 3  
  2.1 Population of Reference for the DAD ......................................................................................... 3  
  2.2 The DAD Frame .......................................................................................................................... 3

3 Data Collection and Standards ........................................................................................................... 4  
  3.1 Data Collection ........................................................................................................................... 4  
  3.2 Data Quality Control .................................................................................................................. 6  
  3.3 Standardization .......................................................................................................................... 9  
  3.4 Linkage ....................................................................................................................................... 10  
  3.5 Equivalency ................................................................................................................................ 11

4 Major Changes to the DAD ............................................................................................................... 11  
  4.1 Historical Comparability ............................................................................................................. 11

References ............................................................................................................................................ 13

Contacts ............................................................................................................................................. 14
Abbreviations

CACS  Comprehensive Ambulatory Classification System
CAD  Clinical Administrative Databases
CCI  Canadian Classification of Health Interventions
CIHI  Canadian Institute for Health Information
CMG  Case Mix Group
CSR  client services representative
DAD  Discharge Abstract Database
ED  emergency department
ELOS  expected length of stay
HCN  Health Care Number
HMDB  Hospital Morbidity Database
ICD-9  International Statistical Classification of Diseases and Related Health Problems, 9th Revision
ICD-9-CM  International Statistical Classification of Diseases and Related Health Problems, 9th Revision, Clinical Modification
ICD-10-CA  International Statistical Classification of Diseases and Related Health Problems, 10th Revision, Canada
NACRS  National Ambulatory Care Reporting System
NCAD  National Clinical Administrative Databases (steering committee)
PCCF  Postal Code Conversion File
PHAC  Public Health Agency of Canada
1 Introduction

1.1 An Overview of the Discharge Abstract Database

The Discharge Abstract Database (DAD) is a national database for information on all separations from acute care institutions, including discharges, deaths, sign-outs and transfers, within a fiscal year (April 1 to March 31). Over time, the DAD has also been used to capture data on day surgery procedures, long-term care, rehabilitation and other types of care.

More than 3.2 million abstracts are submitted annually to the DAD, representing approximately 75% of all acute inpatient separations in Canada. Quebec does not submit data to the DAD; Quebec’s acute inpatient separations are reported to the Hospital Morbidity Database (HMDB) and usually account for 25% of total inpatient separations in Canada.

About 2.4 million day surgery abstracts are submitted to the Canadian Institute for Health Information (CIHI) annually; approximately 35% are sent to the DAD and 65% are sent to the National Ambulatory Care Reporting System (NACRS).
Figure 1 depicts data flow from institutions to the DAD and the HMDB.

**Notes**
1. Data submission/correction via the DAD—coded summary not including name or street address.
2. Data submission via the HMDB—coded summary not including name or street address.
3. Data submission/correction via the DAD (Manitoba and Alberta)—coded summary not including name or street address.
4. Registry/database-specific subset of data.
2 Coverage

2.1 Population of Reference for the DAD

The population of reference includes all separations from acute inpatient care and day surgery institutions in Canada (excluding stillbirths and cadaveric donor cases) from April 1 to March 31. All acute care data except that from Quebec is submitted to the DAD; Quebec acute care data is submitted via Quebec’s ministère de la Santé et des Services sociaux once per year and is included in the HMDB. Day surgery data from Ontario, Alberta and Nova Scotia is submitted to NACRS.

Information about the HMDB and NACRS can be found in *Data Quality Documentation, Hospital Morbidity Database* and *Data Quality Documentation, National Ambulatory Care Reporting System*. This document refers to the set of records submitted to the DAD only.

The population of reference for the DAD can be identified by either the Analytical Institution Type Code or the Institution Type Code; the former is used most often. The Analytical Institution Type Code was introduced to the DAD to minimize the impact of differences between level-of-care definitions across provinces/territories and to facilitate comparative reporting across Canada. It is a CIHI-defined data element that is assigned when the Institution Type assigned to an Institution Number is known to differ from the type of care provided. **Users are advised to use Analytical Institution Type Codes to identify acute inpatient and day surgery separations.**

2.2 The DAD Frame

The frame for the DAD is an inventory of institutions that is used to ensure the collection of all units in the population of reference. Since the provinces and territories determine which institutions will be included in the DAD and all Institution Numbers are identified in advance, the DAD frame is validated by individual provinces and territories. If data is not received from a particular institution, that institution is contacted by CIHI if necessary.
3 Data Collection and Standards

3.1 Data Collection

Figure 2 summarizes the three stages of information flow to the DAD. Patient flow (stage one) into a health care facility leads to information flow at the institution level (stage two). Information at the institution level then feeds into information flow at CIHI (stage three).

![Figure 2: Discharge Abstract Database—Information Flow](image)
3.1.1 Abstracting and Data Submission

The DAD abstract is a record of hospital activity that is completed for each instance of a hospital separation (discharge, death, sign-out or transfer of the patient to another facility). The data collected on each abstract includes coded diagnostic, intervention and patient demographic and administrative information. The format of the DAD abstract was changed in 2001–2002 to accommodate the adoption of the ICD-10-CA/CCI classification systems in some provinces and territories. The International Statistical Classification of Diseases and Related Health Problems, 10th Revision, Canada (ICD-10-CA) is an enhanced version of the 10th revision of the International Statistical Classification of Diseases and Related Health Conditions (ICD-10) and was developed by CIHI for morbidity classification in Canada. The Canadian Classification of Health Interventions (CCI) is the Canadian standard for classifying health care interventions. Since 2004–2005, all provinces and territories have submitted data to the DAD using the ICD-10 abstract. In 2007, the DAD abstract and the standard suite of reports were modified to support the new Case Mix Group methodology, CMG+. This methodology is designed to aggregate acute care inpatient cases with shared clinical and resource-utilization characteristics and to take advantage of the increased clinical specificity of ICD-10-CA and CCI. Section 3.3.1 describes these classification systems in detail.

The Discharge Abstract Database Abstracting Manual\(^1\) is available as a PDF on CIHI’s website through CIHI’s Core Plan service package. The manual provides data element definitions, collection instructions, valid data values, validation rules and edits. The manual is used by clients, researchers and abstracting software vendors.

Adherence to the data submission and abstracting standards described in the manual helps to ensure that CIHI’s reports accurately reflect the institution’s activities. Adherence is obtained through the application of hard and soft edits, education sessions and ongoing client support.

3.1.2 Data Submission Timeline

All data must be submitted to the DAD prior to the year-end deadline. The published submission deadline is July 31, which is four months after the end of the reference period (March 31).

3.1.3 Completeness of Data Submission

A number of reports are generated upon data processing that show the number of abstracts in the database by province/territory. These reports are used to monitor data submission throughout the year and to follow up with facilities or provincial/territorial representatives when an unusually high or low number of abstracts is submitted for any period.
3.2 Data Quality Control

In addition to the above, quality control for the DAD also occurs through the following channels:

3.2.1 CIHI Production System Edits and Correction Process

More than 900 data element edits are applied to each abstract as it is processed at CIHI to ensure that the data in each field is in the expected format, falls within a specific range of values and has a logical relationship to other data elements. For most data elements, when errors are detected, a standard default value of Z is substituted into the data field (for hard errors), or the field is flagged with a warning message. For some data elements, blanks or numeric values are used to represent missing or invalid data. The client receives an electronic report that provides the details of all abstracts and fields that were defaulted or flagged and is asked to submit corrections. The correction and editing steps are repeated until either the client successfully corrects the abstracts or the database closes at the year-end deadline. Before the end of the fiscal year, clients can submit previously missing abstracts or delete duplicate abstracts. Any uncorrected hard errors that remain in the database can be identified by the standard default value of Z.

In addition to verifying individual data elements, the editing process checks a number of inter-relationships. Clients may receive an error message in a field when the reported value is valid but violates certain logical relationships with the data in other fields.

To ensure relevance and consistency, edits are reviewed and updated each year as new data elements are added and changes are made to the database. Test cases and specifications are created according to internal guidelines so that new edits will function correctly.

3.2.2 Abstracting Software

CIHI maintains data capture quality control measures by offering vendor support, coordinating the annual release of vendor system specifications and facilitating vendor system testing. CIHI requires vendors to test their software annually. They must submit a specified number and type of test abstracts, which are then processed in a testing environment to ensure that the format and content of the files meet the submission requirements for the fiscal year. Facilities are also required to submit test submissions after their vendor has passed an annual test.

3.2.3 Annual Database Change Cycle

Every year, enhancements are made to the database to address emerging health care issues, address client needs and improve data quality. Refinements and suggested enhancements to data elements and edits in the DAD are communicated to CIHI in several ways, including

- Routine communication from clients (both internal and external) to DAD client services representatives;
- Input from advisory committees; and
- Formal submissions for data element additions or deletions from key stakeholders.
Appendix B in *Data Quality Documentation, Discharge Abstract Database—Current-Year Information* outlines the mandatory and optional data elements. Appendix C in the same document outlines the evolution of data elements over time.

### 3.2.4 Advisory Groups

The National Clinical Administrative Databases (NCAD) Steering Committee has national representation from ministries of health, Statistics Canada and the Public Health Agency of Canada (PHAC). This committee assists in the annual database change cycle by providing feedback on requested changes, such as proposals for new data elements and edits, as well as revisions to current data elements and edits. The feedback helps to ensure national comparability.

### 3.2.5 Client Services Representatives

CIHI assigns specific client services representatives (CSRs) to provide support for data collectors in each province and territory. The CSRs answer questions related to DAD products; assist in the development and delivery of education programs; provide data quality expertise; and build relationships with provincial/territorial data consultants, health organizations and data users.

### 3.2.6 CIHI Education Program

Through the CIHI education program, instructional sessions are provided to clients on coding and abstracting, how to manage submission errors and corrections, the CMG+ methodology and other related topics. These sessions are one mechanism to ensure standardized data collection coding practices and adherence to CIHI’s data submission and collection requirements. The instructional sessions are provided through elearning courses, self-study packages and web conferences and can be accessed on the Events and Education page of CIHI’s website.

CIHI’s eQuery application is a shared knowledge base that helps registered users find answers and ask questions related to the DAD on topics such as data file submission, institution file updates, application access, reports, abstracting and data quality. It can be used to submit new questions if clients cannot locate answers to their questions in the knowledge base. When the Clinical Administrative Databases (CAD) team learns of a new DAD-related issue, the CSRs will post a question and answer in the eQuery knowledge base in anticipation of clients’ questions.

### 3.2.7 Special Studies

CIHI’s Data Quality department evaluates coding and abstracting accuracy in the DAD via reabstraction studies. Reabstraction involves returning to the original source of information (a patient chart) and comparing it with information in the DAD. The studies focus on data used to calculate specific health indicators, select administrative clinical data, and diagnosis and intervention coding. CIHI has conducted studies using DAD data as far back as 1999–2000.
These earlier studies can be found on CIHI’s website and include the following:

- *Discharge Abstract Database (DAD) CMG/Plx Data Quality Re-Abstraction Study*
- *Data Quality of the Discharge Abstract Database Following the First-Year Implementation of ICD-10-CA/CCI*

More recently, the Data Quality department at CIHI implemented a five-year plan for ongoing reabstraction studies, beginning with DAD 2005–2006 data and ending with DAD 2009–2010 data. The purpose of these studies was to evaluate the overall quality of clinical and non-clinical information and to identify issues associated with coding and abstracting variations. Each study year could also incorporate specific focus topics that are of interest to stakeholders. Upon completion of a study year, results were released to all institutions and provinces/territories that participated in the study, and a summary report was made available on CIHI’s website.

The 2005–2006 study included results from all provinces and territories in Canada. The focus of the study was on selected health conditions and interventions, such as ambulatory care sensitive conditions, hip replacements and percutaneous coronary interventions. The results were released in winter 2008, and a summary report has been available on CIHI’s website since August 2009.

The 2006–2007 study focused on data that is included in the CMG+ grouping methodology, such as flagged interventions and out-of-hospital interventions. Facilities in British Columbia, Alberta and Ontario were targeted. The results were released in summer 2009, and a summary report has been available on CIHI’s website since November 2009.

The 2007–2008 study included all of the provinces and territories across Canada except Quebec and New Brunswick. The focus of the study was on selected health conditions. The results were released in winter 2010, and a summary report has been available on CIHI’s website since May 2010.

The 2008–2009 study included facilities from British Columbia, Alberta, Nova Scotia, Manitoba, Saskatchewan and Ontario. This study focused on the quality of coding for stroke patients, as well as the quality of coding for the administration of thrombolytic therapy to stroke patients. A summary report has been available on CIHI’s website since September 2010.

The 2009–2010 study included all of the provinces and territories across Canada except Quebec. The study focused on the coding quality of selected health conditions, including drug-resistant organisms, palliative care, pneumonia, post-admit comorbidities, post-intervention conditions (including prefixes 5 and 6) and flagged interventions, as well as overall inpatient coding quality at a national level. A summary report has been available on CIHI’s website since March 2012.
3.3 Standardization

3.3.1 Classification Systems

Classification systems in health care provide a standard mechanism for the capture and coding of diagnoses and interventions. ICD-10-CA replaced the earlier ICD-9 and ICD-9-CM classifications. CCI was developed and is maintained by CIHI. It contains a comprehensive list of diagnostic, therapeutic and support interventions and replaced the CCP and ICD-9-CM intervention codes. The ICD-10-CA and CCI classification systems were implemented in a staggered fashion across the country. Please see Table 1 below for details.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>N.L.</td>
<td>ICD-10-CA/CCI</td>
<td>ICD-10-CA/CCI</td>
<td>ICD-10-CA/CCI</td>
<td>ICD-10-CA/CCI</td>
<td>ICD-10-CA/CCI</td>
</tr>
<tr>
<td>P.E.I.</td>
<td>ICD-10-CA/CCI</td>
<td>ICD-10-CA/CCI</td>
<td>ICD-10-CA/CCI</td>
<td>ICD-10-CA/CCI</td>
<td>ICD-10-CA/CCI</td>
</tr>
<tr>
<td>N.S.</td>
<td>ICD-10-CA/CCI</td>
<td>ICD-10-CA/CCI</td>
<td>ICD-10-CA/CCI</td>
<td>ICD-10-CA/CCI</td>
<td>ICD-10-CA/CCI</td>
</tr>
<tr>
<td>B.C.</td>
<td>ICD-10-CA/CCI</td>
<td>ICD-10-CA/CCI</td>
<td>ICD-10-CA/CCI</td>
<td>ICD-10-CA/CCI</td>
<td>ICD-10-CA/CCI</td>
</tr>
<tr>
<td>Y.T.</td>
<td>ICD-10-CA/CCI</td>
<td>ICD-10-CA/CCI</td>
<td>ICD-10-CA/CCI</td>
<td>ICD-10-CA/CCI</td>
<td>ICD-10-CA/CCI</td>
</tr>
<tr>
<td>Sask.</td>
<td>ICD-10-CA/CCI (partial)</td>
<td>ICD-10-CA/CCI (full)</td>
<td>ICD-10-CA/CCI</td>
<td>ICD-10-CA/CCI</td>
<td>ICD-10-CA/CCI</td>
</tr>
<tr>
<td>Ont.</td>
<td>ICD-9/CCP and ICD-9-CM</td>
<td>ICD-10-CA/CCI</td>
<td>ICD-10-CA/CCI</td>
<td>ICD-10-CA/CCI</td>
<td>ICD-10-CA/CCI</td>
</tr>
<tr>
<td>Alta.</td>
<td>ICD-9-CM</td>
<td>ICD-10-CA/CCI</td>
<td>ICD-10-CA/CCI</td>
<td>ICD-10-CA/CCI</td>
<td>ICD-10-CA/CCI</td>
</tr>
<tr>
<td>N.W.T.</td>
<td>ICD-9-CM</td>
<td>ICD-10-CA/CCI</td>
<td>ICD-10-CA/CCI</td>
<td>ICD-10-CA/CCI</td>
<td>ICD-10-CA/CCI</td>
</tr>
<tr>
<td>Nun.</td>
<td>ICD-9-CM</td>
<td>ICD-10-CA/CCI</td>
<td>ICD-10-CA/CCI</td>
<td>ICD-10-CA/CCI</td>
<td>ICD-10-CA/CCI</td>
</tr>
<tr>
<td>N.B.</td>
<td>ICD-9-CM</td>
<td>ICD-9-CM</td>
<td>ICD-10-CA/CCI</td>
<td>ICD-10-CA/CCI</td>
<td>ICD-10-CA/CCI</td>
</tr>
</tbody>
</table>

ICD-10-CA and CCI codes are reviewed regularly. Codes may be added or deactivated as requirements in the field dictate. Further information, including a description of these changes and an indication of when the changes occurred, may be found in CIHI’s ICD-10-CA and CCI Evolution Tables.¹³

3.3.2 Mandatory/Optional Status

The collection of a DAD data element within a province/territory can be mandatory, optional or vary in definition, depending on decisions made by the provincial/territorial ministries of health. Response rates are typically low for non-mandatory fields. Users should be aware of these variations when conducting data analyses. For an overview of data elements’ mandatory/optional status, consult Appendix B in Data Quality Documentation, Discharge Abstract Database—Current-Year Information,² as well the DAD Abstracting Manual.¹
When performing analyses over time or across provinces and territories, users should note that data element specifications could change between fiscal years. For example, some data elements that were optional in 2001–2002 might have been mandatory in 2002–2003. For an overview of data element evolution over time, please consult Appendix C in Data Quality Documentation, Discharge Abstract Database—Current-Year Information, as well as the DAD Abstracting Manual.

3.4 Linkage

3.4.1 Postal Code

Postal Code is a common variable in CIHI databases. If it is used along with the Postal Code Conversion File (PCCF) from Statistics Canada, any standard geographical classification can be located, and the information in databases can be compared. The forward sortation area (FSA)—that is, the first three digits of a postal code—is the lowest level of aggregation normally available to external users under CIHI’s privacy and confidentiality policy. The release of information for small geographical areas may also be restricted to ensure confidentiality. Special requests must be approved by CIHI’s Privacy, Confidentiality and Security Committee. Note that for rural areas that use post office box numbers, postal code data does not necessarily provide an accurate picture of patient residence. This is because box numbers can point to a region other than the place of residence. In addition, when rural postal codes include more than one enumeration area, it becomes more difficult to determine a specific place of residence.

3.4.2 Time Frame

The standard time frame for the DAD is the fiscal year (the period from April 1 of one year to March 31 of the following year). Within the DAD, a number of variables—the fiscal year, admission date/time and discharge date/time—give the flexibility of specifying abstracts that belong to a specific time period, such as the calendar year. This flexibility is especially useful when making comparisons with registries, which tend to be cumulative rather than separate databases for discrete years.

3.4.3 Institution Number

As some institutions close and others merge, a single institution can have different numbers over time. A frame of Institution Number changes is required to perform linkages over time.

Users should also be aware that the Institution Numbers for the reporting of day surgery visits are not the same in the DAD as they are in NACRS. When conducting trend analyses, mappings must be performed between the DAD day surgery Institution Numbers and the NACRS ambulatory care facility numbers.

In order to prevent potential identification, any requests for institution-identifying information require approval by CIHI’s Privacy, Confidentiality and Security Committee.
3.4.4 Health Care Numbers

Health Care Numbers (HCNs) are assigned to individuals by provincial ministries of health and territorial governments. The DAD also captures a variable representing the province or territory that issued an HCN, as the numbers are unique only within the province or territory. Combining the two variables with other relevant personal information data fields (such as Birthdate, Gender and Postal Code) allows individuals to be uniquely identified within the DAD. The HCNs facilitate linkage to other databases with the same fields.

CIHI applies standardized algorithms to encrypt all HCNs to maintain patients’ privacy and, at the same time, to enable linkage. Linkage over time, therefore, can be accomplished only by using the encrypted HCN. Health Care Number, Birthdate and full Postal Code are not normally made available to external users. Access to these restricted data elements and the use of DAD data for data linkage studies requires prior approval by CIHI’s Privacy, Confidentiality and Security Committee. Users should note that patient names and street addresses are not part of the DAD.

3.5 Equivalency

3.5.1 Day Surgery

Day surgery visits are submitted to either the DAD or NACRS. Over the years, more institutions/jurisdictions have made the switch from reporting day surgery visits to the DAD to reporting them to NACRS. There are differences between the DAD and NACRS for day surgery reporting with respect to day surgery definition, data elements and valid values, which can make it difficult to compare information between the two databases and across different fiscal years. Current work is under way to align these databases for day surgery reporting. Users are strongly advised to be aware of these differences when doing historical analysis and provincial comparisons.

4 Major Changes to the DAD

4.1 Historical Comparability

4.1.1 Case Mix Grouping

CIHI’s CMG+ methodology uses the data to derive CMGs, Comprehensive Ambulatory Classification System (CACS) groups, expected length of stay (ELOS) and Resource Intensity Weights (RIWs). CMGs categorize patients into statistically and clinically homogeneous groups based on clinical characteristics and resource use. Adjustments for patients of different levels of acuity form the basis for comparisons between health care organizations and case mix–adjusted resource utilization. Over the years, these grouping methodologies and their accompanying indicators have been used by health care facilities to plan, monitor and manage the services they provide.
All recent acute care inpatient data in the DAD is grouped to CMG+. Formerly named CMG/Plx, the CMG+ methodology is designed to aggregate acute care inpatient cases with similar clinical and resource-utilization characteristics and to take advantage of the increased clinical specificity of ICD-10-CA and CCI. Redevelopment of the CMG/Plx and relative cost weighting methodologies included an extensive review of the CMG/Plx grouper logic, complexity methodology and age groupings. More information on CMG+ can be found on CIHI’s website.

4.1.2 Historical References

The following products are useful references for users of DAD data. Users should consider both the fiscal year and classification scheme when referring to DAD documentation.

- DAD Abstracting Manual
- Quality Assurance Processes Applied to the Discharge Abstract and Hospital Morbidity Databases
- CMG+ Directory (ICD-10-CA and ICD-9-CM available)
- DAD Resource Intensity Weights and Expected Length of Stay
References


Contacts

For more information, please contact CIHI by writing to cad@cihi.ca.