Quality Assurance Processes

Applied to the Discharge Abstract and Hospital Morbidity Databases

Canadian Institute for Health Information

Institut canadien d'information sur la santé

August 2002

Quality Assurance Processes Applied to the Discharge Abstract and Hospital Morbidity Databases

August 2002

Canadian Institute for Health Information
377 Dalhousie Street
Suite 200
Ottawa, Ontario
K1N 9N8

Telephone: (613) 241-7860
Fax: (613) 241-8120
www.cihi.ca

ISBN 1-55392-078-3 (PDF)

Table of Contents

Table of Contents

Introduction

This document reviews the quality assurance activities of the Discharge Abstract Database (DAD) and National Hospital Morbidity Database in accordance with the Accuracy, Timeliness, Comparability, Relevance and Usability dimensions outlined in CIHI's Data Quality Framework.

Good data quality is the outcome of a solid quality assurance process used to manage a database. But what is "good" data quality? Where does a quality assurance process begin and end? For the Discharge Abstract Database, the production of ***accurate*** and valid data begins with the ***timely*** submission of data according to pre-defined codes and data elements outlined in CIHI's DAD Abstracting Manual and using the International Statistical Classification of Diseases and Health Problems 10th revision as endorsed by the World Health Assembly of WHO. Once data are submitted to CIHI, systematic quality assurance practices begin that are designed to ensure ***comparability*** and ***usability***. ***Relevance*** is maintained through consultation with advisory committees and the dissemination of comparative and special topic reports.

The Purpose of this Report

This document is intended to serve as a single reference source for information about the data quality and quality assurance processes applied to both the Discharge Abstract and Hospital Morbidity databases.

What is the DAD?

The Discharge Abstract Database (DAD) is a national database for information related to hospital inpatient and day surgery events. Currently, over four million records are submitted to the DAD annually. Inpatient records submitted to DAD represent 75% of all inpatient discharges in Canada. Each record in the DAD captures a standard clinical, demographic and administrative data set on a patient-specific basis. The database, in its present form, includes data from fiscal years 1979/80 to the present.

The Discharge Abstract Database was originally developed in 1963 to collect data on hospital discharges in Ontario. Over time, it has expanded to provide almost national coverage (excluding Quebec and parts of Manitoba). Information from the Discharge Abstract Database is used by a variety of agencies and facilities for planning, evaluation, research, and hospital funding. Hospitals also use the data to support facility-specific utilization management decisions and administrative research. Governments use the data for funding and system planning. Given these important uses, the quality of data submitted to and produced from the DAD warrants careful attention.

What is Hospital Morbidity?

Like the DAD, the Hospital Morbidity Database (HMDB) provides a count of inpatients separated (through discharge or death) from a hospital, listed by the primary (most responsible) morbidity (disease) diagnosis.

Management and responsibility of the National Hospital Morbidity Database was assumed by CIHI from Statistics Canada in 1995 during a transfer of several health services databases. Data for Hospital Morbidity are downloaded from the Discharge Abstract Database for participating provinces. Data files for hospitals not submitted to the DAD (mainly Quebec and parts of Manitoba) are submitted annually to CIHI by the respective governments for inclusion in the National Hospital Morbidity database. Data are received from general and allied special hospitals, including acute care, convalescence and chronic facilities (except in Ontario). Data do not include any outpatient services in any hospital, or services in psychiatric hospitals.

CIHI's Data Quality Framework

An ongoing challenge for any organization producing statistical information is to ensure that the quality of the information it produces is suited for its intended uses, and that data users are provided with good information about data quality. To this end, the Canadian Institute for Health Information (CIHI) has established a comprehensive and systematic data quality program that includes the implementation and ongoing monitoring of a corporate Data Quality Framework, as well as conducting special studies that focus on data quality issues.

The Data Quality Framework (DQF) was introduced to provide a common, objective approach to assessing the data quality of all CIHI databases and registries. It also

standardizes information on data quality for users and helps to identify priority issues, which in turn leads to continuous improvements. The DQF draws on the Statistics Canada guidelines and methods, the Information Quality literature, CIHI's mandate, as well as the principle of Continuous Quality Improvement (CQI). The first version of the framework was implemented on CIHI's Ontario Chronic Care Reporting System in April 2000.

The framework is structured along five general dimensions of quality: accuracy, timeliness, comparability, usability and relevance. These five dimensions are based on 24 characteristics, which in turn are based on 86 criteria. The framework implementation is part of the larger quality cycle in which problems are identified, addressed, documented and reviewed on a regular basis. The evaluation for the database results in an assessment of appropriate, marginal, not acceptable or unknown for each dimension of quality.

Dimensions of Data Quality
Accuracy—how well information within a database reflects what was supposed to be collected.
Timeliness—examines whether the data are available for user needs within a reasonable time period.
Comparability—refers to the extent to which a database can be properly integrated within the entire health information system at CIHI.
Usability—describes how easily the storage and documentation of data allows one to make intelligent use of the data.
Relevance—incorporates all of the above dimensions to some degree, but focuses specifically on value and adaptability.

The implementation of the Framework invokes the quality cycle:

- determine time period of assessment, i.e. fiscal year of data;
- determine date of availability of data;
- evaluate data quality;
- document results;
- assign priorities;
- implement enhancements; and
- re-evaluate.

Figure 1 provides a summary of the Data Quality Framework version 1.

CIHI will be updating the Data Quality Framework based on the first round of implementation and evaluations by the CIHI data holdings. A plan will be developed to articulate the data limitations and overall data quality of each data holding, to external users. CIHI also plans to share a simplified version of the updated Data Quality Framework with data suppliers and users.

Figure 1. The CIHI Data Quality Framework Evaluation Instrument, Version 1 (CIHI-DQF, v1)

Dimension	Characteristics	Criteria
I. Accuracy	i.1. Over-coverage	1-6
	i.2. Under-coverage	7-12
	i.3. Simple response bias	13
	i.4. Reliability	14-15
	i.5. Correlated response bias	16
	i.6. Collection and capture	17-24
	i.7. Unit non-response	25-26
	i.8. Item (partial) non-response	27-30
	i.9. Edit and imputation	31-37
	i.10.Processing	38
	i.11.Estimation	39-41
II. Timeliness	ii.1. Timeliness	Actual Release-Planned Release 42-45
III. Comparability	iii.1. Comprehensiveness	46-49
	iii.2. Integration	50-53
	iii.3. Standardization	54-57
	iii.4. Equivalency	58-59
	iii.5. Linkage-ability	60-64
	iii.6. Product comparability	65
	iii.7. Historical comparability	66-69
IV. Usability	iv.1. Accessibility	70-75
	iv.2. Documentation	77-78
	iv.3. Interpretability	79-81
V. Relevance	v.1. Adaptability	82-84
	v.2. Value	85-86

1.0 ACCURACY

Accuracy—how well information within a database reflects what was supposed to be collected.

1.1 ICD-10-CA/CCI

Classification systems in health care provide a standard mechanism for the capture and coding of diagnoses and interventions. The coding classification schemes supported by CIHI include:

- ICD-10-CA—Enhanced Canadian version of the 10th revision of the International Statistical Classification of Diseases and Related Health Problems. ICD-10-CA replaces the earlier ICD-9/ICD9-CM classification.
- CCI—Canadian Classification of Health Interventions, developed to accompany ICD-10-CA. CCI replaces the earlier CCP classification.
- ICF—International Classification of Functioning, Disability and Health (formerly known as ICIDH).

The initial version of the ICD-10-CA and CCI Coding Guidelines was released in 2001. These guidelines are reviewed, amended and enhanced annually by a Pan-Canadian Committee representing the provinces and territories. The 2002 Guidelines, of which there are 178, are available in printable document format (pdf) on the CIHI web site and may be down loaded free of charge. The Guidelines were developed:

- to ensure national data quality and consistency
- to decrease subjectivity in clinical coding
- to reinforce the rules for coding in ICD-10-CA
- to clarify the Canadian enhancements to the WHO's ICD-10
- to reinforce the coding rules for CCI
- to accurately reflect case complexity
- to aid in future grouper development
- to enhance CIHI data user confidence.

1.2 What is ICD?

The International Statistical Classification of Diseases and Related Health Problems—Tenth Revision was adopted by the World Health Assembly in 1990. It is the most recent revision of an international classification for classifying mortality and morbidity statistics.

Recently, it has been recognized that a single classification is incapable of: 1) meeting the needs of the increasing number and variety of professional groups in the health field; and 2) satisfying the demands for uniform assessment of health problems for decision making in disease prevention, provision of health care and research on particular health problems. Instead, the development of a family of classifications has been proposed. The International Classification of Diseases is the core classification.

Ongoing Maintenance and Updating

Adaptability, maintenance and updating are critical if a classification system is to be dynamic enough to be used in our rapidly changing world. Unlike previous Revisions, ICD-10 allows for enhancements to accommodate newly discovered diseases, such as AIDS. WHO has established an ongoing maintenance and updating process that ensures input from member states, including Canada, as well as from interested professional bodies. In addition, there are plans to share updates internationally by means of the latest technology. This enhances the long-term viability of the classification system. At this time there are no plans for an eleventh revision of ICD because of the introduction of this maintenance and updating process.

1.3 What is CCI?

The Canadian Classification of Health Interventions, referred to as CCI, is a multi-axial classification of health-related interventions, developed and maintained by the Canadian Institute for Health Information. It contains a comprehensive list of diagnostic, therapeutic, and support interventions (>17,000 codes), and includes a tabular listing (listing of codes in alphanumeric order), an alphabetical index, as well as relevant coding guidelines/recording instructions.

Key Features of CCI

A number of guiding principles were used to assist in the development of the new classification. These principles form the basis for some of CCI's key features. These are described below.

Service Provider and Service Setting Neutral

One of the key features of the classification is its service provider and service setting neutrality. The classification has been developed in such a way that modes of practice are not reflected in the code structure. Therefore, the same codes are intended to be applicable regardless of whether a physician, a nurse or a respiratory technologist performs the interventions, or whether the intervention is performed in an operating room, an emergency department, a clinic or a physician's office.

Multi-axial and Hierarchical Structure

The code structure design uses a multi-axial approach to identify, for example, the body system/anatomy involved, and the intervention performed, including the approach and the technique or device used. The coding structure is also designed to be hierarchical to facilitate data analysis by providing roll-up and roll-down capabilities at various levels (e.g. anatomy, intervention, qualifiers).

Comprehensive

The classification has a significantly expanded scope to meet the needs across the continuum of health services. The comprehensive range of interventions reflects the broad spectrum of providers and variety of applications beyond traditional classifications.

The classification provides NOS (not otherwise specified) and NEC (not elsewhere classified) categories (where required) to ensure that all interventions can be classified somewhere. The classification also clearly identifies, by the use of inclusion and exclusion terms, which interventions and their common synonyms map to each conceptual term.

Relevant

The classification has been developed to ensure that the meaning of each conceptual term is unique and clinically significant. Complex and multi-component interventions are identified, where possible, by a conceptual term which recognizes the various levels of complexity. Furthermore, selected experts from various clinical specialties participated in the developmental process to ensure that the classification is clinically relevant.

Simple
The classification's code-building logic has been, where possible, made apparent to users to enhance comprehension and improve coding practices and data quality. As well, it should facilitate data capture through controlled or natural language interfaces as part of the electronic health record. Each code description has been kept as short and simple as possible while providing maximum detail supportable by clinical documentation to meet the needs of users.

Dynamic and Expandable
Ongoing maintenance and updating of the classification will be facilitated by reserving blocks of codes which will allow for future growth or changes.

Restricted to Procedure-related Information
For the most part, diagnostic (or other non-procedure variables) are not included in the intervention code. This information will be collected elsewhere in the diagnosis portion of the abstract.

Ongoing Maintenance and Update of CCI
CCI will be revised/updated on a regular basis to reflect changes in practice and technologies used to perform the various interventions.

1.4 CIHI's Coding Query Database

In order to (a) enhance CIHI's supported coding and classification schemes, (b) maintain a link to the user community, and (c) document consistent coding practices in a single data repository, CIHI developed an on-line coding query database. CIHI's on-line coding query service helps registered users to find answers to their ICD-10-CA/CCI coding questions. The service is available to all clients who have acquired the ICD-10-CA/CCI CD-ROM. The ultimate purpose of this service is to improve data quality by creating a single repository that will influence consistent coding and thereby good data quality. Responses to queries are designed:

- to reinforce coding rules and guidelines
- to continue the education process in the use and application of the electronic product
- to facilitate the coder's understanding of the coding structure of ICD-10-CA and CCI
- to reflect current coding practices
- to alert coders to any changes or enhancements to coding practices through references to new guidelines, bulletins or research.

The coding query database is maintained by CIHI's Standards Branch.

1.5 The DAD Abstracting Manual

The DAD Abstracting manual is the tool provided to clients, in either PDF or html format, to guide them with the abstracting process of the demographic, administrative, and clinical data elements. The manual is also used by researchers and external software vendors who design abstracting systems for use in hospitals. The manual is divided into two sections: the core section with requirements applicable at a national level and the provincial variation section relating to specific requirements for a province/territory. Depending on their requirement, the data elements are defined as mandatory or optional. For each data element, a description of the data element with the corresponding valid values, an example of how the field is to be used, and applicable edits to the field, are documented in the DAD manual.

The DAD Abstracting manual has been designed around the Group (19 groups) and Field (161 fields) concept. Each data element has been designated a group and field number for easy reference. Many data elements only need to be collected once on each abstract. However, the abstract provides flexibility in instances where a data element needs be repeated, for example the 2001/02 DAD abstract can accommodate data collection on 8 provider numbers, 25 diagnoses, 20 interventions and 6 Special Care Units.

1.6 Hospital Morbidity Specifications Manual

The Hospital Morbidity Specifications Manual (1994/95-1996/97) was compiled to serve as a source of information on the steps involved in the processing and editing of the data used in the creation of the Hospital Morbidity Data Base (HMDB). It also serves as a source for the definition, description and interpretation—including code values—of the data elements in the database. In this regard the main body of the manual reflects the situation as of 1996/97. For each subsequent year, an "Addendum" was written to reflect any changes in the processing and editing that were made in that year.

The manual is in seven sections and four appendices The first (and shortest section) is a layout of the flat file record that is extracted from the database and sent to Statistics Canada annually (3 pages). The second section defines and describes the data elements—including code values—that are included in the database (24 pages). Section 3 consists of a "Process Decomposition Chart" which "blocks out" the sequence of steps Information Systems goes through in the production of the tape of records that is sent to Statistics Canada (4 pages). Section 4 (57 pages) describes the "Process Descriptions" in detail that were "blocked out" in section 3. Section 5 consists of a "Data Flow Diagram" (1 page). Section 6, "Data Maps/Input/Output Specifications," consists of a series of charts tracing each data element—on the file that is sent to Statistics Canada—back to its DAD, QDB, Quebec, PEI or Manitoba source file (26 pages). Section 7 consists of the record layouts of the source files of the provinces whose ministries of health—in 1996/97—submitted data directly to CIHI, as opposed to provinces that submitted their data through DAD (25 pages). Appendix "A" consists of the institution name, provincial hospital number and ICFMI number of all hospitals included in the HMDB for 1996/97 and indicates

whether or not these hospitals also submitted data in each of 1994/95 and 1995/96. Appendix "B" consists of a key to "Quebec Patient Service" codes. Appendix "C" consists of tables showing the grouping of ICD-9 and CCP codes into CDL (Canadian Diagnostic Listing) and CPL (Canadian Procedures Listing) categories, respectively. Appendix "D" consists of diagrams of the tables comprising the HMDB relational database.

1.7 DAD Support Services Representatives

In the context of supporting accuracy through Quality Assurance practices for consistent coding and abstracting, CIHI's Support Services Representatives serve as liaisons to data suppliers on issues with the DAD. The Support Services Representative (SSR) is responsible for providing direct client support related to the DAD products, assisting in the development and delivery of educational programs, providing site visits to clients, providing data quality expertise, and building relationships with provincial/territorial data consultants, health organizations and data users. Each Service Support Representative has been assigned specific provinces/territories for effective and efficient support and expertise. The Support Services Representatives are certified with the Canadian Health Record Association with a minimum of five years experience in the health record field.

1.8 CIHI Education Program

The Education Program is a core function of CIHI. In the context of data accuracy, education programs facilitate the understanding of health information and CIHI products and services. This Program offers its clients a series of workshops and distance learning programs (e.g. paper-based, e-learning, teleconference) to support the implementation and maintenance of national standards and reporting systems.

Sessions focusing on the interpretation, uses and application of data, indicators and other information tools are also offered. Through these initiatives, the Education Program enhances the quality of data submissions to national databases and registries as well as their correct interpretation and application.

For the Discharge Abstract Database (DAD), the education program goes a long way to promote and maintain the integrity of this CIHI database. The education program focuses on the input of data (e.g. coding), the database (e.g. submissions, errors and corrections) and outputs (e.g. report interpretation) with data quality being emphasized in all three areas. Refer to the following table for more specifics.

Focus	Session	Highlights
Annual Update:	What's New for Fiscal XXXX for the DAD	• Introduces and reviews the annual revisions to the Discharge Abstract Database (DAD).
Data Quality	Quality Coding: From Input to Impact	• Provides education around issues related to coding and data quality. • Discusses auditing procedures and quality programs to help ensure accuracy and consistency at the institutional, regional and national levels. • Focuses on the expanding role of data in decision making and the impact data quality programs can have in this process.
Input and Data Quality	Introduction to ICD-10-CA and CCI—(Self-Learning Program)	• Introduces concepts of CIHI's classification standards—ICD-10-CA and CCI
Input and Data Quality	Coding with ICD-10-CA and CCI (Workshop)	• Promotes correct coding with the new classification standards ICD-10-CA and CCI. • Encourages standardized coding practices among CIHI clients • Promotes an accurate and consistent facility, regional, provincial and national database.
Input and Data Quality	ICD-10-CA/CCI Refresher	• Reinforces the understanding of code structures, key concepts and guidelines relative to both ICD-10-CA and CCI • Improves comfort level when coding with the new standards.
Input and Data Quality	Diagnosis Typing	• Promotes standardized diagnosis typing practices to ensure an accurate and consistent facility/provincial/national database
Database and Data Quality	DAD Abstracting	• Provides a detailed overview of the (new) DAD and NACRS abstracting guidelines • Highlights changes for DAD and NACRS Abstracting. • Reviews the testing process, data submission and correction process, and future reporting requirements.

Focus	Session	Highlights
Database and Data Quality	Errors and Corrections	• Introduces and reviews the CIHI data correction system • Ensures the integrity of CIHI's database
Output and Application	Report Interpretation I & II	• Promotes standardized interpretation and application of information from monthly/quarterly reports • Enhances the overall understanding of ELOS reports, CHAP 1, 2 & 3 reports
Output and Application	Inpatient Grouping Methodologies: CMG, DPG, and Plx	• Provides a detailed look at Case Mix Group (CMG) and Day Procedure Group (DPG) Methodologies • Gives an insight into the assignment of complexity (Plx).
Output and Application	ELOS and RIW	• Focuses on the methodology, formulas, and application of RIW and Expected Length of Stay (ELOS).

The above table represents an inventory of education offerings that CIHI has provided over the last four years related to the DAD. In the last two years, the implementation of ICD-10-CA/CCI has been the primary focus and it was not appropriate to schedule all of these sessions. As ICD-10-CA/CCI takes hold, CIHI will once again turn to many of these sessions to encourage good coding practices and emphasize the importance of data quality.

Education is a bridge between CIHI products and the successful application of the content of these products. Through education initiatives, CIHI is able to work with users (e.g. health record personnel) to capture data in an accurate and acceptable manner and provides an opportunity to support the correct interpretation and application of the clinical/administrative data (i.e. to not make an incorrect assumption or conclusion). Both of these initiatives contribute to the data quality and the appropriate application of the various data holdings of CIHI, and especially, the Discharge Abstract Database (DAD).

1.9 Abstracting Software & the Role of External Software Developers (Vendors)

In order to standardize and ensure accurate data collection, CIHI's respective data suppliers hire external software vendors to install the necessary software infrastructure to enable data submission and analysis. Data collection or abstracting software is generally developed by external companies or by individual hospital Information Systems staff. At CIHI, the external companies are commonly known as 'vendors' or 'VAR' (Value Added Retailers). The abstracting software is developed according to CIHI electronic data submission specifications, standards and data holding user documentation (i.e. DAD Abstracting (User) Manual). Data providers (e.g. hospitals) select and contract vendor systems that best meet their information and internal reporting needs. The abstracting software is either "stand-alone" or interfaced with hospital systems such as the Patient Registration or Admission/Discharge/Transfer (ADT) system. Where the abstracting system is interfaced, many of the demographic data elements are downloaded into the software.

Starting with the fiscal 2002/2003 DAD, vendors were provided detailed edit specifications that match the CIHI production system for the DAD. Prior to this, vendors developed their edits by using the DAD Abstracting Manual abstracting rules and error message descriptions. As of fiscal 2002/2003 DAD (for use with ICD-10-CA/CCI) CIHI expects that the vendor abstracting systems incorporate all of CIHI edits and consequently this should improve data quality. If the vendor software is developed appropriately, it should edit each data element at the time that data are entered by the coder and inter-field relationship edits should occur at some point prior to sending the abstract to CIHI. An example of a single data element edit is the discharge date where it must be in the valid format of YYYYMMDD. An example of inter-field edits on the discharge date check to see if the discharge date falls within the submission period, that it occurs after the admission date, etc. The use of this software allows the data to be edited "at source" and corrected prior to being submitted to CIHI where it is edited again in a 'batch mode'.

Vendors are expected to submit test files to CIHI prior to installing their system(s) at a client site or at a minimum prior to their clients submitting a 'hospital test' or their client sites submitting "live" data. There should be sufficient test records on this file to provide good feedback to the vendor (minimum of 250 records is expected). This process tests the control record, file format, file size as well as a full edit test on the file. Feedback (i.e. error/default/rejection report) is provided to the vendors outlining problem areas where changes are required. For most systems (except the DAD), CIHI expects to receive "clean" data only and therefore, if the vendor's system contains all of the appropriate edits, then there will be no error reports produced (only a submission report). CIHI staff analyze the reports and support the vendor through their system development process. Vendors are expected to correct all of their problems before submitting a second or subsequent test. Vendors are expected to test the various record types (e.g. original abstracts, corrections, additions and deletions) within their test submissions. For the DAD, vendors are expected to have a 90% error-free submission in order for CIHI to deem that the test is "successful".

CIHI does not certify vendors, but will post the vendor test results when they are deemed successful according to pre-set criteria.

CIHI/Vendor Relationship
The data provider contracts with the vendor of their choice for products and services. CIHI sets the standards and receives the data from the data providers. CIHI provides support to the vendors for the interpretation of the standards, electronic data submission and edit specifications. CIHI liaises with vendors on an ongoing basis to provide fiscal updates to specifications, identify issues and provide feedback encountered during the transmission of data from client sites.

1.10 Bulletins

Bulletins provide periodic updates for data collectors, Ministries of Health and vendors. Bulletins are generally topic-specific and directed to a specific data holding. This communication medium is used to identify issues, updates and general data quality information. These are distributed via regular mailing channels, and posted on the CIHI website. The distribution list of client groups are identified at the bottom of each Bulletin. The authoring unit and Bulletin number are identified under the CIHI address at the top left-hand side of the Bulletin.

1.11 CIHI Production System Edits

The CIHI Production System Edits are developed to ensure validity and integrity of the data submitted to a data holding. In the DAD, there are errors that are considered hard or soft. Hard errors are those which are identified in fields that are mandatory, meaning blanks are not permitted and standard values must be coded. Soft errors or warning messages apply to fields where recording is optional, where fields may be user-defined. There are approximately 700 Production System field edits in DAD. The DAD accepts erroneous data and standard values are defaulted into the field when the data are identified as incorrect. This standard default value is Z.

There are the following types of Production System edits:

a) individual field (data element) edit
   - mandatory or optional (hard or soft)
   - valid values in appropriate use (specific values, range of values)
   - format (i.e. justification, numeric, alpha or alpha-numeric fields)
b) inter-field (data element) edits (between 2 or more data elements)
c) provincial/territorial variations in correct usage
d) abstract against institution file
e) control and batch integrity testing
f) post grouping methodology edits

CIHI developed a standard layout for the types of edits listed in a), b) and c) above. These are contained in an MS-Excel workbook format listing the core edits (those that

apply to all provinces/territories) and province/territorial edits as separate worksheets within the workbook. This workbook is maintained for each fiscal year. Enhancements or modifications may be made during a fiscal year (see Change Management Policy & Procedures in this document).

1.12 Corrections submitted DAD

Before the information on the abstract is accepted for reporting and storage, it is edited for validity and consistency against the 700 CIHI defined edits. The purpose of the edit program is to identify the erroneous data and guard against both the printing of the erroneous data on reports and the storage of erroneous data on the master data file.

To facilitate the correction process, an error message appears on the default report for each erroneous data element. The error messages are divided in hard and soft/warning type of errors. When a data element is missing or invalid, but is mandatory to record, the data element is replaced by "Z"s meaning that a correction must be applied to this data element. Soft/warning edit requires verification by the abstractor/coder to ensure the accuracy of the information. Each error message consists of three sections: error identification, error description, and data as submitted. All error messages are preceded by the following identifiers: group number, field number, occurrence number (where applicable) and edit code number (number identifying the specific error message). The group and field identify the data element to be corrected and the edit code explains the cause and action to rectify the error.

The correction abstract method is designed to allow software in the hospital to create a Correction Abstract by direct access to the hospital database after a change has been made to that database. Detailed specifications describing this process are included in CIHI's software system specifications that are provided to abstracting vendors.

Errors detected by the edit system and reported on the Default Report or additional changes requested by the client can be applied at any time after the receipt of the initial monthly reports but prior to the production of the annual reports and closure of the fiscal year for DAD. During the course of a fiscal year, clients can also submit additional abstracts if previously missing at the time of submission of a period or delete duplicate abstracts when detected in subsequent analysis.

1.13 Corrections applied to the DAD

The edit program processes data as the abstracts are received by CIHI. All errors detected in the abstracts, for the same period, are reported together. After each editing process, the system ensures that the appropriate added values such as Case Mix Group (CMG™) or Day Procedure Group, Complexity Level (Plx™), Expected Length of Stay (ELOS), and Resource Intensity Weight (RIW™) are re-calculated where corrections on diagnosis and/or intervention codes were made. The corrections and editing steps will repeat until the client (i.e. hospital) successfully corrects the abstracts or the database closes as per its year-end deadline.

1.14 Quality Assurance in the National Hospital Morbidity Database (HMDB)

Manitoba and Quebec are the only remaining provinces whose ministries of health submit data directly to the Hospital Morbidity Data Base (HMDB) rather than through the Discharge Abstract Data Base (DAD). Effective April 01 2003, it is anticipated that only Quebec will do so.

The process for acquiring Manitoba and Quebec data begins with a letter that is internally referred to as a "Call for Data". After the formal requests for the data are sent out to each province's ministry of health, usually in September, the Manitoba data arrive in late October/November and the Quebec data in late December/January of that year. The same data elements which DAD contributes to the HMDB are then extracted from the tape or CD that the provincial ministry has sent in by Information Systems (IS) at CIHI and put online for checking and analysis at CIHI. Based on past experience, it is with these two provinces where errors are most likely to be detected in the Morbidity database given that their submissions have not been edited according to CIHI DAD Database specifications. The source files are checked for these two provinces for counts of records and any obvious anomalies that may present themselves. Concurrently, the DAD file for the remaining provinces is simultaneously checked against trend information (e.g. substantial changes in disease categories, etc). The data are then loaded directly into a series of "staging tables" in preparation for processing. The counts in the "staging tables" are checked against the counts obtained from the source files. If the counts are the same, a successful transfer of data elements/records is assumed to have taken place. The diagnosis and procedure codes are validated at this stage and sequence numbers are assigned to them in accordance with editing rules. A number of other edits such as dropping records with a separation date that lies outside the fiscal year, spurious hospital type codes, etc. are performed. A number of derived variables are also added. Once this has been done the "final" set of tables is produced. Counts of records are again checked between the "staging tables" and the "final" tables. From the "final" set of tables, a flat file is prepared and sent to Statistics Canada.

Upon receipt Statistics Canada begins its own data quality verification activities. This begins an iterative process between CIHI and Statistics Canada to discuss anomalies and/or questionable trends that may require further follow-up directly with the data sources (i.e. Ministries). Statistics Canada checks the data file and any questions are reported back to CIHI and appropriate steps are taken from imputation to re-processing of a file depending on the size or impact of any findings. Any findings are also corrected on CIHI's internal files. Internal files are then loaded into CIHI's Query & Analysis platforms for analysis and reporting.

1.15 Re-abstraction: Special Studies in Data Quality

Given the size, coverage and importance of the Discharge Abstract Database, CIHI is conducting a special study, designed to evaluate the accuracy of the DAD data. The DAD Data Quality Study is the first national study that uses a statistical sampling methodology to reliably measure the accuracy of the coding of selected non-medical and clinical administrative data contained in the DAD.

The study is assessing the data quality of the DAD by returning to the original sources of information (i.e. patient charts) and comparing this information with what exists in the CIHI database. The study will review two years of ICD-9/CM/CCP data and one year of ICD-10-CA/CCI data. Data collection for year one of the study was conducted in the fall 2000 (fiscal 1999/2000 data) and for year two (fiscal 2000/2001 data) was collected in the fall 2001. Due to the staggered provincial implementation of ICD-10-CA/CCI, the third year of the study has been scheduled for fiscal year 2003/2004. At that time, data will be re-abstracted for fiscal year 2001/2002.

The goal of the DAD Data Quality Study is to evaluate the accuracy of selected data, at the national level, contained in the DAD. The specific objectives of the study are to:

1. Evaluate and measure the overall accuracy of the DAD;
2. Evaluate and measure the impact of data collection from incomplete charts;
3. Evaluate and measure the coding quality of diagnoses and procedures relevant to specific Health Indicators represented in the CIHI Health Indicators Framework;
4. Identify and measure how often diagnoses and procedures are not coded according to CIHI guidelines and identify where additional coding guidelines may be required;
5. Assess whether any of the of the above evaluations have an impact on the assignment of case mix group and expected length of stay (ELOS); and
6. Facilitate the evaluation of the change to new diagnosis and intervention standards ICD-10-CA/ CCI.

The health indicators to be evaluated were identified in the first two years of the study through consultation with staff within CIHI's Health Reports and Analysis section. In year one, the study focused on the following health indicators selected from the CIHI Health Indicator Framework:

- Ambulatory Care Sensitive Conditions
- Cesarean Sections
- Coronary Artery Bypass Graft
- Hospitalization due to Pneumonia and Influenza
- Injury Hospitalizations
- Total Hip Replacement
- Vaginal Births After Cesarean Sections

In year one of the study, the sample size was increased to enable assessment of the following indicators which were developed and defined by the Canadian Perinatal Surveillance System (CPSS), Health Canada:

- Rare Congenital Anomalies
- Rare Maternal Conditions
- Rare Neonatal Conditions
- Respiratory Distress Syndrome
- Third Degree Perineal Laceration
- Other Non-rare Maternal & Neonatal Conditions

In addition to the CIHI and CPSS indicators, a sample of charts that did not contain any of these indicators was also randomly selected. This was done in order to estimate the false negative type of discrepancies. For ease of reference, this sample is defined as: Not assigned to any of above conditions.

The following health indicators were sampled for the second year of the study:

- Hysterectomy
- Hip Fracture
- Total Knee Replacement
- Myocardial Infarct

Sampling Methodology
The study used a multi-stage sampling approach to identify which charts would be re-abstracted. The first sampling stage randomly selected facilities across Canada[1] stratified by geography and size.

The second sampling stage randomly selected charts from each facility. All abstracts in the fiscal 1999/2000 and fiscal 2000/2001 DAD were assigned to one health indicator where possible (refer to objectives of study). In cases where an abstract could be assigned to more than one indicator, for selection purposes only, the condition with less prevalence was given priority. During analysis of the data, other diagnoses and procedures in the abstract were reviewed.

Eighteen facilities participated in the first year of the study, allowing for the reabstraction of 2,737 charts. Eleven facilities participated in the second year of the study, resulting in the re-abstraction of 1,555 charts.

In order to get an optimal sample design, it was assumed that at the national level, the proportion of charts for each indicator that contained a discrepancy was 15%. The reliability required for the sample was a coefficient of variation (C.V.) of 16.5% (that is, a

[1] The target population of the Study is provincial acute care facilities reporting to the DAD. Facilities in Quebec and some in Manitoba do not submit data to the DAD. Facilities in the Yukon, Northwest, and Nunavut Territories were not included in the study population for cost reasons.

standard error of 2.5%). Using this assumption and reliability requirement, a minimum sample size was then determined for each indicator. This sample size was increased by 10% to account for chart non-response (unavailability) and a further 10% for possible situations of better than expected productivity by the re-abstractors. There were 150 charts randomly selected at each participating facility—however the number of charts selected for each indicator varied among the participating facilities, so that the overall minimum number of charts for that indicator overall of the facilities was achieved as far as possible. Note that each sampled chart has an unequal probability of selection under this design.

Collection of Study Data

CIHI Classification Specialists[2] re-abstracted the data for the study by returning to the original source of the data on-site at each facility for a one -week period. All the original information from the DAD was downloaded to a laptop application immediately prior to the collection week. The computer-assisted application was designed and developed by CIHI to facilitate the collection of the study data. The application featured the use of pull-down lists of discrepancy codes and reasons for the discrepancy, as well as a comment field that allowed entry of additional information pertaining to the discrepancy. Additional reference material that would ordinarily be available was also loaded onto the laptop. The Classification Specialists entered all of the re-abstracted, discrepancy and reason data directly into the application.

Next Steps

The combined results of the first two years of the DAD Data Quality Study will be available in a report, in the fall 2002. Analysis of the reabstracted data is currently underway for the second year of the study. Preliminary results of year one of the study are available on the CIHI web site, in the document *Discharge Abstract Database Data Quality Study—Preliminary Year 1 Findings*. The third year of the DAD Data Quality Study will be conducted in fiscal 2003/04. A national report will be produced approximately one year following the actual reabstraction of the charts.

1.16 Case Mix Group/Complexity (CMG/Plx) Special Study

Another Roadmap project is the Grouper Redevelopment Project. With the current implementation of the new classification standards the International Statistical Classification of Diseases and Related Health Problems, Tenth Revision, Canada—Canadian Modification (ICD-10-CA) and the Canadian Classification of Health Interventions (CCI), CIHI will be redeveloping the Case Mix Group Complexity (CMG/Plx) Grouper. In preparation for the Grouper redevelopment, CIHI will be reviewing the complexity component of the Grouper, including Grade lists. One aspect of this review will involve the re-abstraction of actual charts.

[2] CIHI *Classification Specialists* are certified with the Canadian College of Health Records Administrators; are responsible for developing, interpreting and teaching classification systems; are well experienced in various hospital settings; and have expert knowledge of medical terminology and diagnosis and procedure classification standards.

Scope

The scope of this project includes the re-abstraction of medical records from acute-care inpatient facilities across Canada. This information will then be compared with what exists in the DAD. The target population of the study is provincial acute care facilities[3] reporting to the DAD.

The goal of the Case Mix Complexity (CMG/Plx) Data Quality Study is to evaluate the data quality of selected clinical and administrative data for statistical purposes from CIHI's DAD. The Study will assess the data quality of the DAD by returning to the original sources of information and comparing this information with what exists in the CIHI database. While a facility level report on general findings is being provided to each facility, it should be noted that the study is not an audit of an institution's coding. The primary use of the data collected will be to contribute to the assessment of the DAD data quality at a national level. The objectives of the study are to:

- Evaluate and measure the overall data quality of the DAD CMG Grouper variables;
- Evaluate and measure the coding quality of diagnoses and interventions relevant to CMG/Plx assignment;
- Facilitate the development of the ICD-10-CA/CCI CMG Grouper; and
- Facilitate the ongoing development of coding guidelines for the new classification standards (ICD-10-CA and CCI).

Study Approach

Patient charts were re-abstracted and compared to the CIHI database information. The Study used a multi-stage sampling approach. Data collection occurred at each participating facility during the spring/summer 2002. The first sampling stage randomly selected 15 facilities. The second sampling phase randomly selected charts from within the selected facilities. Charts were randomly pre-selected from the database to concentrate on selected complexity levels (refer to Goals/Objectives). A minimum of 55 charts was established as the sample to be re-abstracted from each facility.

Each sampled facility was visited for one week. A CIHI Classification Specialist re-abstracted into a computer application on a laptop provided by CIHI. For each of the data elements re-abstracted:

- all subjective clinical information such as diagnoses and interventions was re-abstracted blindly, i.e. without viewing the original abstracted data; and
- selected non-medical information such as institution was not re-abstracted but was compared with the original abstracted data and flagged as a match or a discrepancy. If a discrepancy occurred, the non-medical data was re-abstracted.

[3] Acute care facility is defined by the institution type flag of the DAD. This does not include Rehab, Chronic Care, Nursing Homes, Psychiatric, Home Care, Same Day Surgery or Emergency facilities reporting to the DAD.

For each discrepancy identified, both medical and non-medical, the re-abstractor assigned the type of discrepancy and a possible reason. There was no reconciliation process with the original hospital abstractor and the identity of the original abstractor was not collected for this Study.

Next Steps

Analysis of the CMG/Plx Data Quality Study is planned for late summer and early fall. A national report is planned for release in the fall 2002.

2.0 TIMELINESS

Timeliness—examines whether data are available for user needs within a reasonable time period.

Special studies are often conducted on the Discharge Abstract or Morbidity databases. In the majority, studies are conducted using DAD and HMDB data as sources of information for analytical or clinical research studies. However, sometimes studies are also conducted that are specifically oriented to enhancing characteristics of data quality, quality assurance or timeliness of the DAD and HMDB databases. In the summer of 2000, CIHI released a report that described a national survey on factors affecting the timely submission of data to the Discharge Abstract Database. The abstract from that report is presented below with the full version available at www.cihi.ca.

Improving Timeliness of the Discharge Abstract Database (2000)
"*You cannot manage what you cannot measure*" was the opening statement made by the Canadian Institute for Health Information's (CIHI) Board Chairperson Michael Decter during the launch of CIHI's Annual Health Report, 2000. This statement serves as a guiding principle for CIHI in its maintenance of Canada's Health Information databases. Implied in this statement is the fact that access to timely information is critical to ensuring effective management of our health systems. Late in the fall of 1998, the CIHI Board identified for attention the timeliness of data, specifically the Discharge Abstract Database (DAD) data. Initial analysis of the data management pathway indicated potential areas for improvement in processes related to data submission, compilation, correction, analysis, reporting and dissemination. Of these, reducing data submission time delays and improving the efficiency of examining the error correction process held the most potential for improving timeliness of DAD data (see data flow diagram of information for DAD on page 24).

In 1999, a national survey was developed and distributed to Canadian acute care facilities in selected provinces regarding the timeliness of data submission by hospitals to the Discharge Abstract Database. The purpose of the survey was to examine data collection and submission processes in hospitals to determine what variation existed in practices in the underlying process related to data submission and collection of DAD data such as documentation and coding required to complete the DAD abstract. Objectives of the timeliness survey were to identify best practices in the timely submission of data and, based on the results, to initiate a nationally oriented change process in data submission and reporting of hospital inpatient service events. The resulting report, available at www.cihi.ca, describes the purpose, objectives, methodology, results and a process that developed recommendations to enhance the availability of timely inpatient hospital data.

This project began with the objective of improving the timeliness of DAD data collection and submission processes. The report conclusions suggest a number of possible policy directions. Encouraging is the fact that many of the findings fail to represent insurmountable obstacles. Most can be addressed through process changes

at the local level that would be facilitated by hospital or provincial policies around data submission that are enforced.

CIHI's Provincial and Territorial Steering Committee for the DAD/Hospital Morbidity Database met in the Spring of 2001 and agreed to explore the following recommendations:

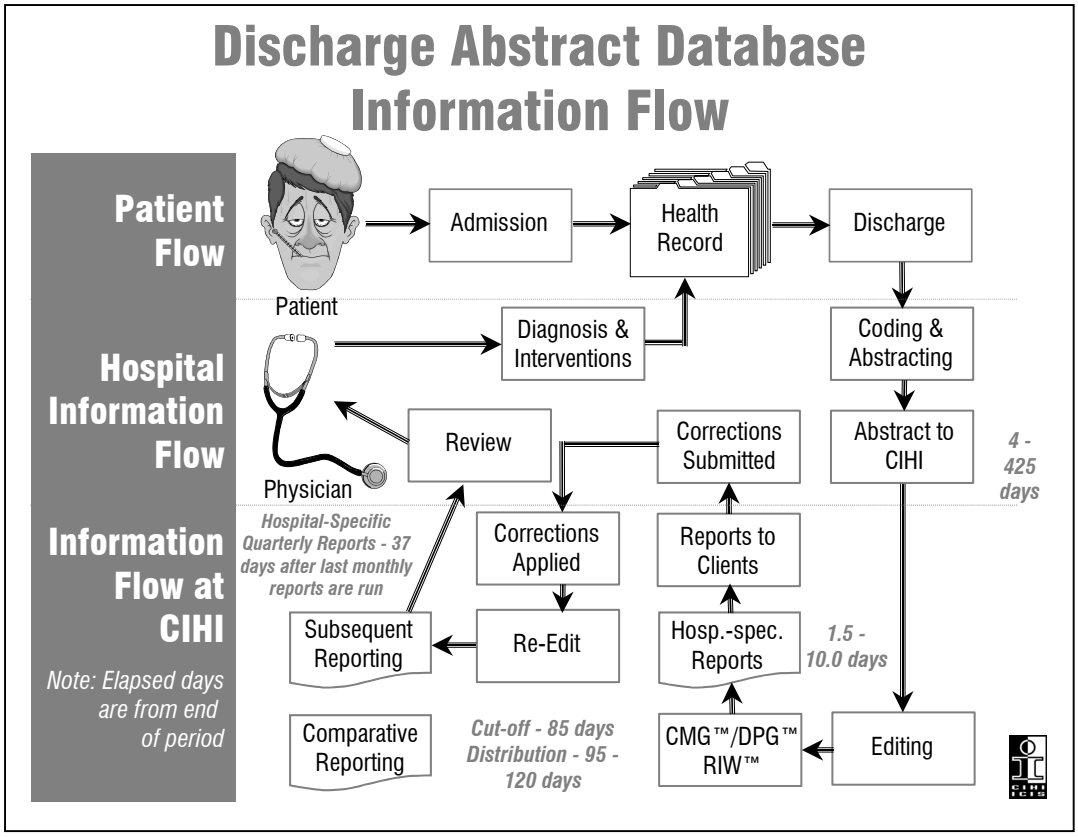
1. Encourage provinces and territories to establish policies requiring data suppliers to submit data and required corrections to CIHI within 31 days following the end of the period.
2. Encourage provinces and territories to establish financial incentives/penalties that relate to the submission of data within specified deadlines.
3. That CIHI revise the deadline for data submissions for comparative reports (i.e. CHAP reports) to 30 days (from 85 days) after the end of the first 3 quarters, with the 4th quarter remaining the same, and distribute reports 25 days (from 37 days) following the deadline.
4. CIHI should investigate the feasibility and advisability of revising its pricing structure of core and non-core plan subscribers to include financial incentives for early data submission.

At the next meeting of this committee in 2002, the Provinces and Territories ability to adopt these recommendations will be discussed. National cooperation is necessary to improve the Timeliness of the National Hospital Morbidity and Discharge Abstract Databases.

CIHI Submission and Correction Deadlines

Timeliness of the Discharge Abstract and Morbidity databases is monitored regularly according to a production schedule that coordinates all submissions and corrections to data submitted. At the beginning of each fiscal year, the Health Services Information branch at CIHI determines fiscal year-end submission and correction deadlines. These deadlines are established in consultation with the various provincial and territorial ministries of health. Although different jurisdictions may set deadlines for data submission within their own jurisdiction (e.g. DAD data must be submitted within 30 days following the end of the period), CIHI sets year-end deadlines only (e.g. F2001/02 data submission—June 28, F2001/02 corrections—July 26). CIHI determines database closure dates at which time final ministry of health files are run and ad hoc requests for the data year may be run. The year-end closure date for DAD is generally September 30.

CIHI also establishes data submission dates for the quarterly Comparison of Hospital Activity Program (CHAP) reports. These dates are used as a 'cut-off' date when data are extracted and the CHAP reports produced. Hospitals with incomplete data are included in the reports and are identified with an asterisk (*) beside their name. The extraction date is usually set at 85 days after the end of the third period in the quarter. The CHAP reports are cumulative.

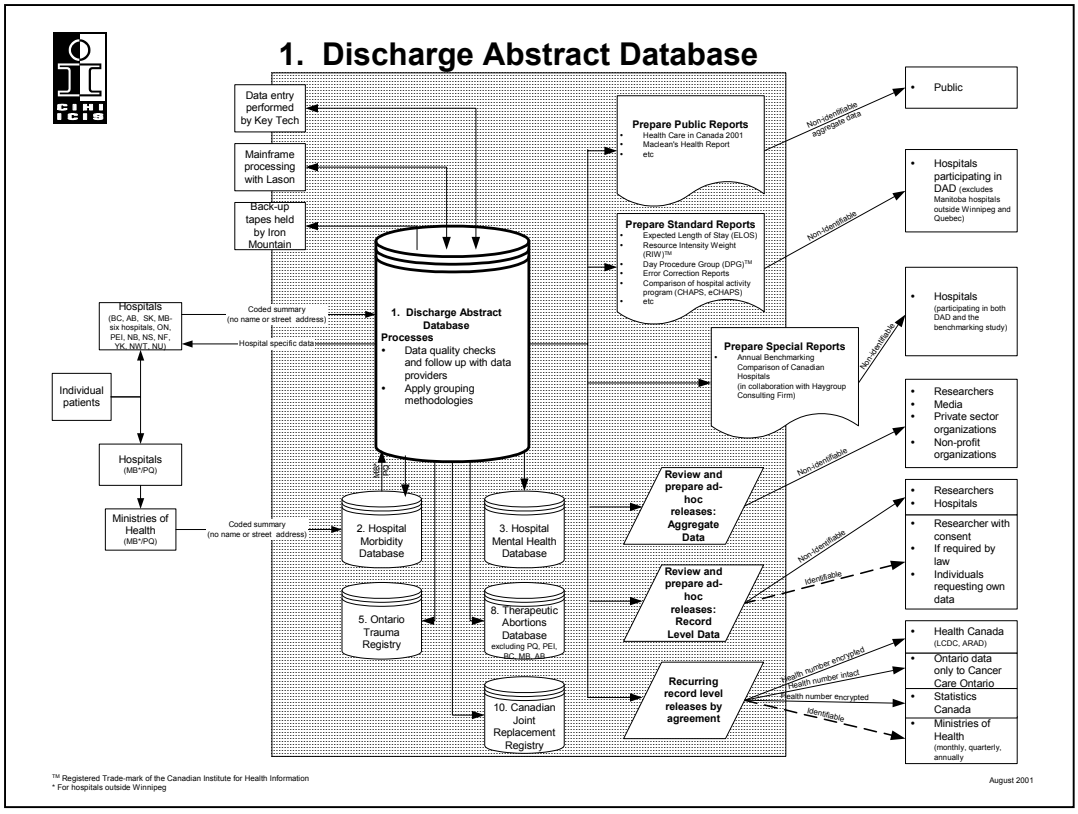

Discharge Abstract Database
Information Flow
Patient Flow
Hospital Information Flow
Information Flow at CIHI
Note: Elapsed days are from end of period
Patient
Admission
Health Record
Discharge
Physician
Diagnosis & Interventions
Coding & Abstracting
Review
Corrections Submitted
Abstract to CIHI
4 - 425 days
Hospital-Specific Quarterly Reports - 37 days after last monthly reports are run
Corrections Applied
Reports to Clients
Subsequent Reporting
Re-Edit
Hosp.-spec. Reports
1.5 - 10.0 days
Comparative Reporting
Cut-off - 85 days
Distribution - 95 - 120 days
CMG™/DPG™ RIW™
Editing

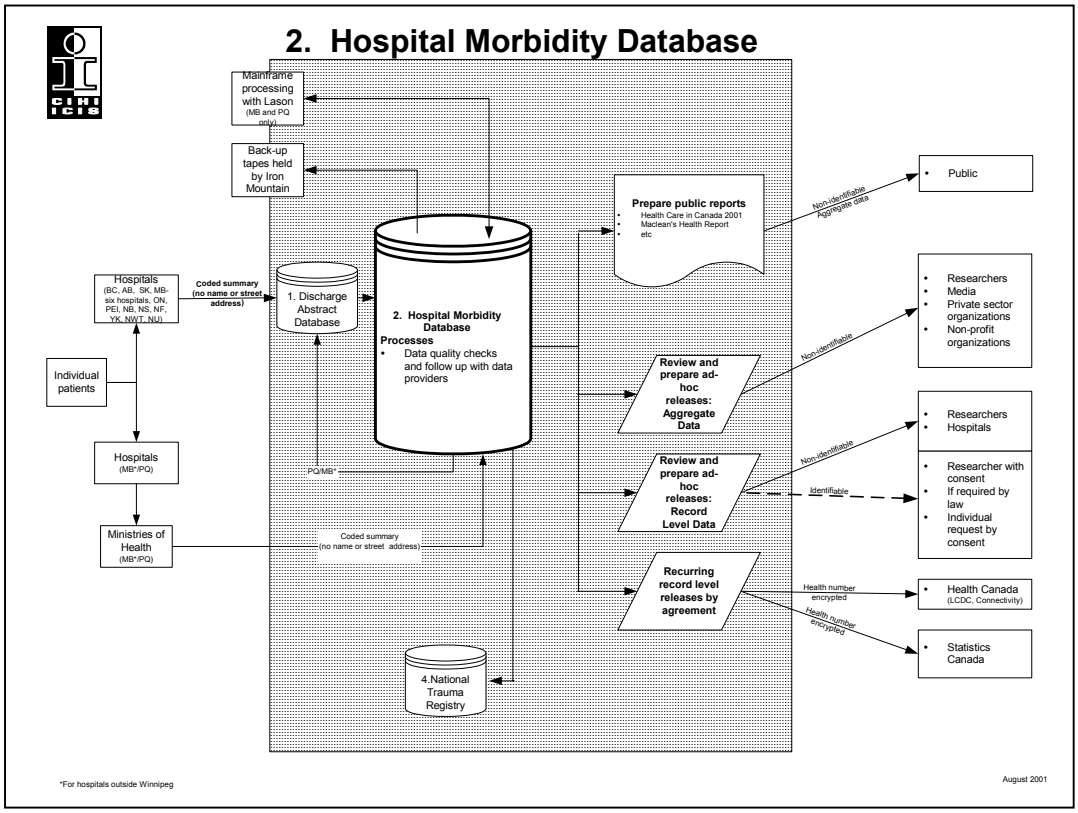
3.0 COMPARABILITY

Comparability—refers to the extent to which a database can be properly integrated within the entire health information system at CIHI.

3.1 Comparability Dimension

For the most part, the DAD and Morbidity databases fulfill the criteria of comparability according to CIHI's Data Quality Framework. The DAD is the central database at CIHI as it includes detailed information on all hospital discharges in Canada (with the exception of Quebec and parts of Manitoba). Consequently, it is the prime resource to investigate issues around acute care and hospital activity. It also constitutes the primary source of the following CIHI databases: Hospital Morbidity Database, Hospital Mental Health Database, National Trauma Registry, Ontario Trauma Registry and the Therapeutic Abortions Database (see data flow diagram on page 26). A limitation is the lack of documentation on the historical comparability of data elements. This is now being addressed through a special initiative to occur in fiscal 2002-03.


1. Discharge Abstract Database
Data entry performed by Key Tech
Mainframe processing with Lason
Back-up tapes held by Iron Mountain
Hospitals (BC, AB, SK, MB- six hospitals, ON, PEI, NB, NS, NF, YK, NWT, NU)
Coded summary (no name or street address)
Hospital specific data
Individual patients
Hospitals (MB*/PQ)
Ministries of Health (MB*/PQ)
1. Discharge Abstract Database
Processes
Data quality checks and follow up with data providers
Apply grouping methodologies
MB*/PQ
2. Hospital Morbidity Database
3. Hospital Mental Health Database
5. Ontario Trauma Registry
8. Therapeutic Abortions Database excluding PQ, PEI, BC, MB, AB
10. Canadian Joint Replacement Registry
Prepare Public Reports
Health Care in Canada 2001
Maclean's Health Report
etc
Prepare Standard Reports
Expected Length of Stay (ELOS)
Resource Intensity Weight (RIW)™
Day Procedure Group (DPG)™
Error Correction Reports
Comparison of hospital activity program (CHAPS, eCHAPS)
etc
Prepare Special Reports
Annual Benchmarking Comparison of Canadian Hospitals (in collaboration with Haygroup Consulting Firm)
Review and prepare ad-hoc releases: Aggregate Data
Review and prepare ad-hoc releases: Record Level Data
Recurring record level releases by agreement
Non-identifiable aggregate data
Non-identifiable
Identifiable
Health number encrypted
Health number intact
Public
Hospitals participating in DAD (excludes Manitoba hospitals outside Winnipeg and Quebec)
Hospitals (participating in both DAD and the benchmarking study)
Researchers
Media
Private sector organizations
Non-profit organizations
Researchers
Hospitals
Researcher with consent
If required by law
Individuals requesting own data
Health Canada (LCDC, ARAD)
Ontario data only to Cancer Care Ontario
Statistics Canada
Ministries of Health (monthly, quarterly, annually
™ Registered Trade-mark of the Canadian Institute for Health Information
* For hospitals outside Winnipeg
August 2001


2. Hospital Morbidity Database
Mainframe processing with Lason (MB and PQ only)
Back-up tapes held by Iron Mountain
Hospitals (BC, AB, SK, MB- six hospitals, ON, PEI, NB, NS, NF, YK, NWT, NU)
Coded summary (no name or street address)
Individual patients
Hospitals (MB*/PQ)
Ministries of Health (MB*/PQ)
Coded summary (no name or street address)
1. Discharge Abstract Database
2. Hospital Morbidity Database
Processes
Data quality checks and follow up with data providers
PQ/MB*
4.National Trauma Registry
Prepare public reports
Health Care in Canada 2001
Maclean's Health Report
etc
Review and prepare ad-hoc releases: Aggregate Data
Review and prepare ad-hoc releases: Record Level Data
Recurring record level releases by agreement
Non-identifiable Aggregate data
Non-identifiable
Identifiable
Health number encrypted
Health number encrypted
Public
Researchers
Media
Private sector organizations
Non-profit organizations
Researchers
Hospitals
Researcher with consent
If required by law
Individual request by consent
Health Canada (LCDC, Connectivity)
Statistics Canada
*For hospitals outside Winnipeg
August 2001

3.2 Comprehensiveness

Overall, the DAD captures its intended universe of acute care admissions and same day surgery in Canada. The omission of parts of Manitoba and Quebec are well known and discussions with each Province to become full clients of CIHI are underway. However, the Hospital Morbidity database provides 100% coverage of acute data by supplementing the DAD with data directly from these provinces.

The primary conceptual universe is that of inpatient acute admissions. However, chronic and rehabilitation care and day surgery are also included for some jurisdictions. For 1998/1999 and 1999/2000, enhancements were made for the collection of additional data on mental health inpatients and therapeutic abortions, respectively.

3.3 Integration

Although standard variables for geography, institution, time and person are in the DAD, they are not fully consistent in terms of variable names or values with other databases, making work with multiple databases unnecessarily awkward and inefficient. However, it is anticipated as databases adhere to CIHI's Data Dictionary developed in 2000, this gap will be addressed.

In addition to residence code, postal code is a common variable on almost all CIHI databases and along with the PCCF (Postal Code Conversion File), any Standard Geographical Classification (SGC) can be obtained, making it possible to make valid comparisons with other databases.

The institution code in the DAD uses the standard 4-digit provincial institution number with a provincial prefix to ensure a unique code across Canada. This is either the same or similar to the codes used in other databases.

The DAD is collected on a fiscal year basis (April 1–March 31), according to date of discharge. It also contains a number of mandatory date variables such as admit date, admit hour, procedure date, discharge date and discharge hour which can be used to calculate standard variables like Length of Stay (LOS) and age.

A number of standard fields in the DAD can be used to identify a person. Both encrypted and unencrypted health card numbers are included on the MS SQL Server tables, but the unencrypted number is removed from Query and Analysis (QnA) data sets and tools used by CIHI analysts. Other relevant fields are birth date, sex and postal code (although only sex and age are available on the QnA data sets). The above fields are not sufficient to identify a unique admission, however, since the database is event based. Fortunately, a unique admission identifier can be created by combining fiscal year, fiscal period, institution id, institution type and abstract id.

3.4 Standardization

Overall, the various data elements for geography, institution, time and person are sufficiently detailed to create equivalent concepts with other CIHI databases.

The mandatory postal code field and use of the PCCF file ensures that any SGC used in other CIHI databases can be obtained for the DAD. Province codes may not be consistent, but there is sufficient detail to create standard codes across different databases if required.

The institution number is the widely used standard and is unique within DAD which uses the provincial institution numbers with an added provincial prefix to make the number. With minor modifications, it can be made equivalent to the institution numbers used in other databases like the Ontario Trauma Registry.

Although the data are collected on a fiscal year basis, the range of date variables in the DAD (admission date, discharge date, etc.) allows the user to examine any time frame. This flexibility is especially useful in comparison with registries, which tend to be cumulative rather than separate databases for discrete years.

Persons can be identified by the encrypted health card number. This is the standard person identifier and common on most other CIHI databases. Other important person identifiers include sex, admission dates, procedure dates, etc. Although variables like birth date and postal code are collected, they are normally not made available to users.

3.5 Equivalency

Overall, the conversion between ICD-9 and ICD-9-CM is well developed and documented. For the most part, the conversion for diagnoses simply involves removing the last character. Future conversions between ICD-10-CA/CCI to ICD-9/CM/CCP will be much more complicated and care will have to be taken to document the process and any problems that may occur. Conversions between postal code and SCG are consistent, well documented and revised often.

The Postal Code Conversion File (PCCF) from Statistics Canada is updated every six months. At the time of publication the most recent version is July 2001. Within the QnA data sets, the files are updated automatically. For the most part, conversion problems are most common with census enumeration area. However, since larger areas like census subdivision and census division are typically used, any error due to conversions is negligible.

The other major use of crosswalks is between ICD-9 and ICD-9-CM for both diagnoses and procedures. The conversion tables are always done from ICD-9-CM to ICD-9. All CM codes are accounted for, although some ICD-9 and CCP codes are not included in the equivalency tables due to clinical mapping issues. The tables are fiscal year specific so they only include the valid codes for a single fiscal year.

3.6 Linkage-ability

The DAD contains the appropriate variables for geography, institution, time and person to allow for linkages with other databases. Although variables' names and codes may often differ across databases, there is typically enough detail to create equivalent matching variables.

Standard SGC coding is not actually done for the core DAD database. Instead, the PCCF file is used to link postal code to the SGC variables. Due to the number of variables in the DAD, it is more convenient to keep those geographic variables in a separate table and link them to the DAD if necessary.

The standard institution code assigned by province is used in DAD and other CIHI databases, with some minor alterations. In the DAD, a province prefix is added to institution code to make it unique. With minor modifications, it can be used to link to other databases like the Ontario Trauma Registry (OTR) that use only the standard four digits.

The standard time frame for most databases, including the DAD, is the fiscal year April 1–March 31, based on discharge date. Within DAD, the admission and discharge date variables give the flexibility of specifying records that belong to a specific time period.

The encrypted and unencrypted health card numbers facilitate linkage to other databases that have the same fields. If the health card number is not present on a particular database, it may be possible to do a probabilistic linkage using variables like birth date, sex, postal code or procedure date.

Privacy and confidentiality are important issues. Guidelines are in place at CIHI to protect privacy and confidentiality (see CIHI's Principles of Privacy & Confidentiality at www.cihi.ca). Reporting is done only on large geographic areas such as provinces and health regions and aggregate results are not released for any cell size less than five without the consent of the data supplier.

3.7 Product comparability

Comparability between products is typically very high. Most use consistent definitions like fiscal year (April 1–March 31), five-year age categories (<1, 1-4, 5-9, 10-14, etc.) and province as the standard geographical region.

3.8 Historical comparability

Historical documentation could be improved and would help explain any breaks in a series. Currently, no formal document exists that outlines all the changes over time, however, it is still possible in some cases to track changes by examining the yearly documentation that is produced (e.g. the coding manuals). Support staff are also available

to answer questions about changes. Historical comparability is verified prior to any trending being performed.

Recognizing the gap in detailed documentation to support comparability over time, a special initiative has been undertaken in 2002-03 to retrospectively create this documentation. This special initiative will result in documentation of the evolution of DAD data elements and their changes in definition from 1985 to 2002. The introduction and endorsement of data elements over time has, varied provincially and consequently tracking the evolution of data elements has at times, created mapping issues and gaps in analysis. The impact to comprehensive data element adherence will also be documented in the context of the different times when provinces participated with CIHI.

4.0 RELEVANCE

Relevance—incorporates many of the Data Quality Framework dimensions to some degree, but focuses specifically on value and adaptability.

Specific processes are in place to document suggestions for modification and to incorporate these changes. These processes are described in this section and provide information on the processes undertaken to ensure the data collected in DAD and Hospital Morbidity continue to be relevant to data suppliers and users.

4.1 How A Data Element Appears in the DAD

Researchers and stakeholders of the database often ask: "How do new data elements get suggested for inclusion in the DAD?" That is, what processes exist to ensure the data are relevant to the needs of data suppliers and its users.

Effectively there are a variety of streams for communicating refinements and suggested enhancements. These include (1) routine communication from clients to our service support representatives, (2) input from advisory committees and (3) formal submissions from stakeholders. Each of these will be described briefly below.

(1) One of the best sources of input toward enhancements is suggestions received directly from the database users or data suppliers. Health Records Professionals represent the majority of individuals charged with the task of collecting and verifying information collected from hospitals. CIHI's Client Service Representatives (described elsewhere in this document) serve as liaisons for these individuals in interpreting and communicating issues related to the DAD. In the course of their work, suggestions are often made for the addition or modification of data elements that would enhance the utility of the DAD to hospitals. Informal logs are maintained by CIHI Client Representatives and are discussed periodically with the Program Area Coordinator and Manager. These suggestions are reviewed for practicality and appropriateness and then aggregated to a list that is eventually discussed with each Provincial or Territorial Ministry of Health for potential adoption.

(2) CIHI supports and benefits from numerous advisory committees that are multi-disciplinary in membership and geographically representative. These committees often are formed for specific advice on topical issues related to the activities of CIHI. CIHI staff in turn, are often invited to participate on external advisory committees and are afforded the opportunity to discuss health information issues as well as epidemiological surveillance issues. These groups provide an excellent forum for the generation of new data elements and coding legends that eventually find their way into the DAD. Of particular importance in this regard is the DAD/Morbidity Steering Committee (see section 4.2).

(3) Formal submissions are also another source of information where professional associations (e.g. Canadian Orthopaedic Association, Canadian Council for Health Services Accreditation, etc) will identify the DAD as a useful resource for the capture of particular information. Often letters will be addressed to the Director of the Program area

requesting CIHI's consideration for the addition of new information in support of a particular research or epidemiological surveillance initiative. These requests along with those described above are then presented to the respective Provincial and Territorial Ministries of Health for potential adoption.

4.2 The DAD-Morbidity Steering Committee

Once all of these suggestions for data element additions and enhancements are tallied from each of the sources described above, annual meetings are organized by CIHI for its DAD-Morbidity Steering Committee. The mandate of this committee is to discuss issues related to the DAD and Morbidity Databases that are operational or strategic in nature. Each Province or Territory appoints the members of the committee, with a requirement that the member possesses decision-making authority on matters related to the DAD or Morbidity Databases in their Province or Territory. The committee will decide consensually whether to identify an element as appropriate for inclusion in the database and whether its collection ought to be mandatory (to ensure national comparability), optional or a Provincial variation (specific only to selected Provinces or Territories). It is through these processes that new data elements appear in the DAD. In 1999, a significant re-development of the DAD occurred where new elements were added and others were deleted.

4.3 The Re-development of the DAD (2000-2003)

In 1998, the Federal, Provincial and Territorial Conference of Deputy Ministers of Health endorsed a vision for health information called the Roadmap initiative. The Roadmap initiative was a collaborative effort between the Canadian Institute for Health Information (CIHI), Statistics Canada, Health Canada and many other key groups at all levels—national, regional and local. The 1999 Federal Budget identified a number of specific priority projects and activities in the health information field and earmarked $95 million over four years toward their completion. These funds were used to expand and/or accelerate ongoing national health information initiatives and to support new ones.

A redevelopment of the Discharge Abstract Database (DAD) was identified as a project within the Roadmap initiative. The redevelopment project covered a number of other inter-dependent objectives, including the re-thinking of outputs (i.e. reports) generated from DAD data and the integration of the DAD and National Hospital Morbidity databases.

The 1999 DAD abstract required changes to accommodate the new diagnosis and intervention codes (ICD-10-CA and CCI). These codes are being phased in according to Provincial or Territorial adoption date of ICD-10-CA/CCI between 2001 to 2003. Objectives of the abstract redevelopment included:

- Accommodate ICD-10-CA and CCI
- Improve comparability of data through increased inter-provincial standardization
- Improve definitions and facilitate linkages among databases and registries
- Add new data elements
- Delete data elements which are no longer relevant

A report titled: An Interim Progress Report of the Re-engineering of the DAD and Morbidity Databases, available from www.cihi.ca, provides an overview of the project, a description of the process followed, and changes that were implemented for the new abstract.

4.4 Production System Processes for Change Control

Enhancements and modifications are a necessary measure to ensure continued utility and relevance to any database. In the context of database management, it is important to ensure a documented and standardized process is in place for change control. Hence enhancements or modifications for the Discharge Abstract or Hospital Morbidity databases, while logged throughout the fiscal year, are only applied to the databases' production systems at the beginning of a fiscal year.

Two primary reasons provide a rationale for only applying any enhancements or modifications at the beginning of the fiscal year. First, sufficient time is needed to notify software developers (vendors) and provide them enough time to make changes to their software and implement this updated software at their client sites. The general rule of thumb is a minimum of six months for existing systems and a minimum of nine months for new or redeveloped systems. Second, a scheduled time for changes in the production cycle ensures consistent data content for a given fiscal year (for internal and external analysis purposes). Two exceptions to the schedule of changes exist. These are:

1) To fix system problems (e.g. 'bugs') enabling the receipt of "error-free" data.
2) To add edits that were not initially identified but required to improve data quality.

Hence, with the above two exceptions noted, the following six steps define the change management policy for DAD and Morbidity:

i) All fiscal updates to existing systems should be determined by the program area in consultation with Production Systems and finalized prior to September 30.
ii) The Production Systems Coordinator or Consultant completes the 'request for change' portion of the system request forms for all of the agreed-upon changes.
iii) Production Systems update and distribute the corresponding Electronic Submission Requirements Document (i.e. 'vendor specifications') no later than October 30.
iv) The Program Area (user) develops test scenarios and creates test data for all of the agreed-upon changes.
v) Production Systems programming staff makes the modifications, tests the system (requires User sign-off on the test results) and implements the 'new' fiscal year system prior to March 1.
vi) Vendors are notified that the "vendor test system" is operational and available for them to submit vendor test files.

The following seven steps define the procedures for change management**:**

i) Client (Program Area) completes a 'system request' form (either paper or on-line) detailing the exact requirement.
ii) Client submits system request log to the Coordinator or Consultant, Production Systems (ideally, the program area should provide 'test data' along with the system request form).
iii) Production Systems Coordinator or Consultant reviews the system request to determine extent and assigns the task to the appropriate programmer.
iv) Programmer makes the modification, performs preliminary system testing.
v) Programmer completes the system request form and returns the form to the Coordinator or Consultant.
vi) Coordinator or Consultant reviews the code, preliminary test results and implements the changes in the production systems environment. The testing process requires 'user' sign-off on the test results.
vii) Coordinator or Consultant informs the 'requestor' that the change has been completed.

4.5 Discharge Abstract Reports

The best technique to ensure data are relevant to suppliers and users is through the routine provision of reports. These reports provide comparative information that is largely used for hospital efficiency comparisons. Reports also can serve administrative purposes as described below.

As referenced in the previous section, the DAD underwent a significant re-engineering in 2000 that examined both the data elements as well as the reports that would routinely be produced. Based on input from a nationally representative body of experts in utilization management, the project team proposed the following strategic directions for DAD reports:

- Move away from hospital-specific reports that do not offer value-added elements or analysis and which can be easily run on the clients' report writer purchased from software vendors;
- Move away from paper reports and increase focus on electronic reports;
- Increase focus on better comparative reports, with emphasis on providing the ability to customize queries.

Consequently, four sets of reports with no 'added value' (as defined above) were cancelled as clients can easily replicate those reports using their own report writer purchased from software vendors. The remaining reports are:

Default Report
The Default Report is a complete list of the abstracts that have had errors identified after the data submission of a period has been processed.

Correction Report
The Correction Report identifies whether a submitted correction has been accepted or rejected.

Subsequent Error Report
Like the Default Report, the Subsequent Error Report is a complete list of the abstracts still containing errors after submission of corrections. Clients must keep on submitting corrections until CIHI editing system accepts all attempts at correcting the errors on the abstract.

Outstanding Error Report
Twice a year (December and May) CIHI runs an Outstanding Error Report where abstracts still containing "Z" (erroneous data) are reported. The report is part of the Data Quality process as a reminder to the clients to correct those errors. A report is forwarded to the provincial/territorial ministries of health regarding the status of corrections required for their province/territory.

ALC (Alternate Level of Care) Report
The ALC Report identifies cases where patients have finished the acute care phase of their treatment but remain in an acute care bed waiting for a placement (extended care facility, hospice, etc.). This report is produced on a monthly basis.

CMG (Case Mix Group) 900 Series Report
CMG 900 is the category for un-groupable data. The CMG 900 Series Report lists all cases that have been assigned to case mix groups 900 to 999. The purpose of this report is to identify cases that should be reviewed for quality of the data coded and abstracted. This report is produced on a monthly basis.

Discharge Analysis Report
The Discharge Analysis Report is an executive management report displaying information about hospital practices. This report presents an overview of patterns of patient care and illustrates the utilization of resources. This report is produced on a monthly, quarterly and annual basis.

ELOS (Expected Length of Stay) Reports
The ELOS Reports summarizes a hospital's patient experience in terms of length of stay. These reports provide an analysis based on the case mix complexity. Each patient is compared to similar cases in the CIHI Expected Length of Stay Database. The reports are sequenced by percent days over/under the database match. These reports are produced either on a quarterly and annual or cumulative quarterly basis.

RIW (Resource Intensity Weight) Reports
RIW are used to standardize measurement of inpatient cases volume by recognizing that not all patients require the provision of the same type or quantity of health care resources. The RIW report supports the translation of case mix data into case costing information by measuring the allocation of weighted cases to individual services, physicians and CMG assignments. These reports are produced either on a quarterly and annual or cumulative quarterly basis.

Complexity Diskette
The Complexity diskette is an ASCII File showing abstract level data with CMG, ELOS, Plx, and RIW values added. The diskette is produced on a quarterly basis.

Day Procedure Group (DPG) Reports
The DPG Profile presents an overview of the ambulatory case mix in the facility. Case volumes allow managers to view at a glance the types of surgery being performed most frequently.

Procedures within a DPG is a companion to the DPG Profile. It prints all abstracted procedures that have been performed within each DPG. These reports are produced monthly/annually or quarterly/annually as requested by client.

Inpatient/Outpatient Comparison by Service (Patient or Provider) offers a comparison between inpatient and outpatient activity and identifies procedures that may be moved to an outpatient setting. This report also compares the hospital's percent of day surgery activity to a database percentage as a benchmark.

Inpatient DPG Listing by Service provides detailed information about inpatients who may be candidates for day surgery based on the Inpatient/Outpatient comparison by Service report.

CHAP (Comparative Hospital Activity Program) Reports
CHAP provides clients with a means of assessing the use of their resources as compared with hospitals of similar size and type based on the number of acute care beds, or teaching/paediatric designated status of the hospitals. The peer groups are:

Peer Group 0—1 to 49 beds
Peer Group 1—50 to 99 beds
Peer Group 2—100 to 199 beds
Peer Group 3—200 to 399 beds
Peer Group 4—400+ beds
Peer Group 5—Teaching facilities
Peer Group 6—Paediatric facilities

CHAP contains a series of four reports:

Based on the length of stay, CHAP 1 measures the hospital performance, assesses the patient mix in relation to peer group and demonstrates the differences in hospital practice.

CHAP 2 report helps analyze the impact of patient age in practice patterns and helps reviewing the admitting practices, unplanned re-admissions and use of the Special Care Units (SCU).

CHAP 3 report allows hospitals the ability to compare their day surgery practice with their peers to facilitate the utilization management decision-making process.

CHAP RIW can be used to evaluate approximate allocation of resources relative to peers.

eCHAP is the first set of electronic reports produced by CIHI. eCHAP is similar in format to the paper CHAP, but has additional functionality. Clients can view the reports on line from CIHI'S web site through a secure environment (i.e. internet encrypted connection). Clients can customize their reports to meet their requirements by choosing defined facilities, provinces, and/or CMG and import a flat ASCII file into an analytic tool for further analysis.

For fiscal 2002/03, all Hospital Specific reports (except for correction reports) and CHAP reports will be disseminated electronically to all clients using the 2001/02 abstracting software platform. This approach is based on clients' feedback and keeping in line with CIHI Future Directions.

5.0 USABILITY

Usability—describes how easily the storage and documentation of data allows one to make intelligent use of the data.

The relevance and quality of analysis are critical. To this end, CIHI's analytical program is designed to respond to information needs identified by health sector stakeholders through consultations and advisory panels. It incorporates data quality, validation/verification, expert peer review, and other processes to ensure that the analysis—and the data on which it is based—are sound. The presentation of methods and results, as well as the dissemination strategies, for all analytical products are designed to be appropriate for their target audience(s) to ensure that they are as accessible and easy to use as possible.

In addition, strong privacy principles, as well as effective safeguards for the confidentiality and security of personal health information, underpin all analytical projects. Analysis takes place in the context of CIHI's Privacy and Confidentiality Guidelines. Efforts continue to be made to develop and share analytical methods and tools to identify and avoid potential residual disclosure risks.

CIHI's Health Reports and Analysis Branch is responsible for many of the Institute's analytical products, including a variety of indicators and analysis derived from the Discharge Abstract Database and the Hospital Morbidity Database. Approved staff access these and other databases using a variety of analytical tools (e.g. SAS, CIHI's QnA environment, and SQL). The quality assurance framework under which these activities are conducted builds on the work described previously and subsequent to this section, designed to continually evaluate and improve the quality of the underlying databases.

5.1 Quality Assurance for Specific Analytical Products

The analysis process begins with the identification of the question to be addressed and the methods to be used. This step is designed to ensure the relevance, utility, and feasibility of the analysis for the intended target audience. It is typically undertaken with the guidance of external advisory groups, as well as clinical experts, methodological advisors, biostatisticians, and data quality and classification specialists where appropriate. It draws on the rich Canadian and international experience with, and literature on, clinical utilization and outcomes analysis.

The resultant analytical plans include well-defined quality assurance processes. Appropriate quality assurance strategies vary, depending on the type of analysis, complexity of the methodologies, data source(s), intended audience, and other factors.

For example, where methodologies are being adapted from the literature or from previous Canadian research, their appropriateness in the context of the characteristics of the relevant dataset is assessed and methodologies are revised as required. Newly developed

methods and approaches are likewise reviewed with internal and/or external expert advisors and in the context of the dataset being used.

Additional assessments of the underlying quality of the data being used for specific analyses may also be undertaken. For example, the estimation of false positive/false negative rates for specific health indicators is an integral part of the multi-year re-abstraction process currently underway for the Discharge Abstract Database.

When numerical analysis begins, two separate, independently prepared, sets of computer code are typically developed and run to extract data and perform statistical analysis. The results are then reviewed to identify and resolve any discrepancies.

An internal verification process is then undertaken. This step includes, where possible/relevant, comparing results with historical trends and/or other data sources, careful review of potential anomalies or outliers, verification of results against control totals established for the database, and much more. In some cases, automated quality assurance programs have been designed to assist with this process. Senior analytical staff also reviews preliminary results and confidentiality checks and rate stability checks are also typically applied at this point. (CIHI privacy and confidentiality policies limit the disclosure of potentially identifiable data.)

An external verification process typically follows. This involves sending preliminary results (often at a greater level of detail than will ultimately be published in order to facilitate verification) back to the original data sources/subjects. For example, preliminary regional health indicators data are shared with health regions and ministries of health. Data are accompanied by definitions, technical notes, and a request to review the results and advise CIHI of any potential issues. Through this process, a number of regions and/or ministries typically replicate the results of the analysis from independent data sources. CIHI analysts are available to provide further information and assist throughout the verification process.

Based on the results of the internal and external verification processes, CIHI analytical staff, in consultation with expert advisors as required, determines whether the information to be presented meets the organization's data quality protocols. They also review the definitions, detailed technical notes, and related materials that will accompany the release of data/analytical results to clarify any questions or issues that arose in the verification processes.

Final results and documentation are then produced. These results, as with those circulated in the external verification process, are the subject of an extensive fact checking process. This iterative process involves pairs of analysts cross-checking draft and final reports against original sources based on standardized check-lists to ensure that results and associated documentation accurately reflect the results of the analysis. Dedicated staff time is assigned to this process. Final results are also subject to senior analytical and institutional review.

5.2 Cross-Cutting Quality Assurance Strategies

The processes described above are designed to support quality assurance of specific analytical products. In addition, the Health Reports and Analysis Branch has developed cross-project strategies with application across many projects and products, including DAD/Morbidity.

5.2.1 Use of the V-File

CIHI's virtual file (or v-file), an electronic knowledge base, contains past and current information and documentation related to CIHI's data holdings, data management, analysis, and reporting. The v-file is available to all CIHI employees, with particular emphasis on analysts' needs.

There are five major sections of the v-file. The first contains information on CIHI's data holdings, including descriptions of all databases and registries, supporting documentation, contacts, relevant publications, data availability, and special notes. Special notes are posted on a variety of topics and are specific to each database or registry. For example, notes may advise analysts of exclusion/inclusion criteria, identified comparability or trending issues, and related information. Current and archival information is retained. For example, the site contains an archive of resolutions of potential data quality and other issues that analysts have been alerted to through special email distribution lists.

The second section contains operational definitions, including a glossary of terms, including a glossary of terms, list of acronyms, and a concept dictionary defining and clarifying terms specific to CIHI (for example Case Mix Group or Major Clinical Category).

The third area provides documentation on common analysis files (for example the postal code conversion file). It also relates specifically to analytic methods and tools, projects, and web-based resources. It includes sample SAS and SQL code for extracting different types of data from the Discharge Abstract Database, information on the QnA portal, and details on the results of data quality studies.

The fourth area provides analysts with information on, and links to, various CIHI and external publications, including selected reports from CIHI's data holdings.

The fifth area relates to CIHI standards. This includes important detailed documentation on ICD-9/10 coding, CCP/CCI coding, CMG, HL7, MCC, and MIS guidelines. It not only includes specific definitions and codes, but it also contains links for analysts to access more detailed information compiled by database managers.

An introduction to the v-file forms part of the orientation program for new analysts. On an on-going basis, analysts are strongly encouraged to consult the v-file before undertaking any new analysis and to contribute to the on-going maintenance and development of this important resource.

5.2.2 Developing Analytical Capacity

CIHI strongly believes in the importance of developing analytical capacity in the organization for many reasons, including supporting quality assurance processes. For example, new analysts participate in orientation sessions where database structures, content, resources, and related topics are reviewed. An introduction to quality assurance processes and resources is an important component of this orientation program. A customized series of modular training sessions on specific analytical tools and databases is also offered. Analysts attend these and other training sessions relevant for the types of work in which they will be involved. Mentorship programs are also in place where new analysts work with more experienced staff to gain further knowledge regarding CIHI's data holdings, analytical methods, and processes and protocols.

On an on-going basis, a variety of strategies is used to promote continual learning and development. These include a journal club to review recent external research, user groups for the various analytical toolsets, job rotation, collaborating with external researchers, and much more.

Summary

As stated in the introduction, the primary purpose of this document is to serve as a single reference of the quality assurance processes applied to the DAD and Morbidity databases. It is hoped that documenting these steps will enable database managers and users to step back and identify potential gaps and pitfalls in our quality assurance activities that need to be addressed to ensure continuing "fitness for use".

CIHI is committed to ensuring quality data. While there is no standard definition of data quality, there are a number of dimensions of quality that can be consistently applied to the maintenance of data quality. These include accuracy, timeliness, comparability, usability and relevance. Policy makers, health care leaders and the general public are dependent on quality data for decisions that affect the Canadian health care system. Through the ongoing data quality evaluations of CIHI's data holdings and the conduct of special data quality studies, CIHI will facilitate the continuous production of quality information. CIHI already has an established reputation for producing high quality information; the ongoing challenge is to build on that reputation by continually enhancing the quality of the underlying data.

This document is meant to be a living document, and it will be refreshed as practices evolve.

References

1. Brown, Ann and Richards, Julie. *Quality Measurements of the Canadian Discharge Abstract Database*. Proceedings of the International Conference on Quality in Official Statistics, 2001. Stockholm, Sweden.

2. Canadian Institute for Health Information. *Study Assesses Data Quality of the DAD*. CIHI directions ICIS, October/November 2000. Ottawa, Ontario, Canada.

3. Canadian Institute for Health Information. *The CIHI Data Quality Framework Version 1, April 2001. Internal Use Only*. Ottawa, Ontario, Canada.

4. Canadian Institute for Health Information. *CIHI's Commitment to Excellence Furthered through Data Quality Initiatives*. CIHI directions ICIS, October/November 2001. Ottawa, Ontario, Canada.

5. Canadian Institute for Health Information. *Discharge Abstract Database Data Quality Study—Preliminary Year 1 Findings.* March 2002. Ottawa, Ontario, Canada.

6. Canadian Institute for Health Information. *Case Mix Group Complexity Data Quality Study*. March 2002. Ottawa, Ontario, Canada.

7. Long, J.A., Richards, J.A., Seko, C.E. *The Canadian Institute for Health Information Data Quality Framework, Version 1: A Meta-Evaluation and Future Dire*ctions. Proceedings of the MIT Sixth International Conference on Information Quality, IQ-2001. Cambridge, Massachusetts, USA.

8. Richards, Julie; Brown, Ann; Homan, Craig. *The Data Quality Study of the Canadian Discharge Abstract Database*. Proceedings of Statistics Canada Symposium 2001: Achieving Data Quality in a Statistical Agency: a Methodological Perspective. Ottawa, Ontario, Canada.