Canadian Institute
for Health Information

Institut canadien
d'information sur la santé

Discharge Abstract Database
Data Quality
Re-abstraction Study

Combined Findings for
Fiscal Years 1999/2000 and 2000/2001

December 2002

Discharge Abstract Database
Data Quality
Re-abstraction Study

Combined Findings for
Fiscal Years 1999/2000 and 2000/2001

December 2002

For further information, please contact the Canadian Institute for Health Information at:

377 Dalhousie Street
Suite 200
Ottawa, Ontario
K1N 9N8

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

ISBN: 1-55392-102-X

Discharge Abstract Database
Data Quality Re-abstraction Study

Table of Contents

Appendices

Discharge Abstract Database
Data Quality Re-abstraction Study

Table of Contents (cont'd)

Introduction

This report provides the combined national findings from the first two years of the Discharge Abstract Database (DAD) Data Quality Re-abstraction study being done by the Canadian Institute for Health Information. The focus of the study is the measurement of the accuracy of selected administrative and clinical data contained in the DAD.

This report contains an overview of the background, specific objectives and methodology of the study. A summary of the combined national level findings from the first year (DAD fiscal year 1999/2000) and the second year (DAD fiscal year 2000/2001) by health indicator and for demographic, administrative and clinical (diagnoses and procedures) data is included. The report concludes with a summary of the study findings, and presents recommendations for consideration by program managers and data suppliers.

This report[1] for the DAD Quality study presents the combined findings for the first and second year of the study. A further report will be produced for the Case Mix Group/Complexity (CMG™/PIx™) Data Quality Study.

Background

Established in 1994, CIHI is a national, not-for-profit organization that plays a critical role in the development of Canada's health information system. CIHI's mandate is to co-ordinate the development and maintenance of a comprehensive and integrated approach to health information in Canada. CIHI's diverse data holdings are playing an increasingly important role in supporting public debate and decision-making about the Canadian health system.

In 1998, the Federal Minister of Health's Advisory Council on Health Infostructure, the Canadian Institute for Health Information, and Statistics Canada brought together over 500 people throughout the spectrum of the health care sector to identify health information needs for Canada. These national consultations resulted in the Health Information Roadmap—an initiative that outlined a national Canadian vision for modernizing health information and funded projects for improvements to the health information system, including enhancements to the Discharge Abstract Database.

Discharge Abstract Database (DAD)

The Discharge Abstract Database (DAD) is a national database containing 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[2]% of all patient discharges in Canada. Each record in the DAD contains standard clinical, demographic and administrative data for the health services provided for each inpatient stay. Health records staff at hospitals code the abstract data from the discharge summary

[1] The preliminary findings from the first year of the study are contained in *"Discharge Abstract Database Data Quality Study, Preliminary Year 1 Findings",* www.cihi.ca, Canadian Institute for Health Information (2002),

[2] Facilities in Quebec and some in Manitoba do not currently submit to the DAD.

and other information contained in the patient chart. On a monthly basis, the abstracted data are forwarded to CIHI where the information is processed and edited. Default reports are provided to hospitals for analysis and correction of erroneous data.

While the primary responsibility for data quality lies with individual data suppliers, CIHI provides direct client support through a number of activities related to the DAD products. CIHI's Support Services Representatives (SSR) and Classification Specialists liaise with data suppliers to provide quality assurance support for consistent coding and abstracting. Other activities include assisting in the development and delivery of educational programs, providing coding and other expertise, and building relationships with provincial/territorial data consultants, health organizations and data users.

In addition to the abstracted data submitted by hospitals, the DAD contains value-added outputs such as the Case Mix Groups and related resource consumption indicators developed by applying CIHI's grouping methodologies and costing algorithms.

A revised DAD abstract was implemented in fiscal 2001/2002 to accommodate the ICD-10-CA/CCI national classification system (described below) and to adapt to the evolving health information needs of stakeholders. The revised DAD abstract was designed to improve inter-provincial standardization, facilitate linkages among databases and registries, and improve the reporting of specific data elements, as well as to add new and delete no-longer required data elements. The DAD Abstracting Manual is provided to clients, in either PDF or html format, to provide guidance for the abstracting of each data element as mandated by each provincial/territorial jurisdiction.

Uses of the Discharge Abstract Database (DAD)

The DAD data are used by a variety of stakeholder groups including health service providers, policy and decision makers, governments (federal, provincial and territorial, regional and local) and researchers. The data are used extensively to monitor utilization of acute care health services, conduct analyses of health conditions and injuries, and support the development of value-added outputs such as Case Mix Groups and related resource consumption indicators. The clinical and administrative data contained in each abstract are used to derive many CIHI health indicators, such as the rates of hip replacements and coronary artery bypass graft surgery. The DAD is also being used increasingly to track patient outcomes. It is a major data source used to produce various CIHI reports and publications, including its annual report on the performance of the health care system, and for seven of the health indicators (see Appendix A) adopted by the federal, provincial and territorial governments to meet the performance reporting commitments agreed to by the First Ministers.

Classification Standards

During fiscal year 1999/2000 and 2000/2001 two classification systems were in use for diagnosis coding, namely, the International Statistical Classification of Diseases, Injuries and Causes of Death, Ninth Revision (ICD-9) and the ICD-9-Clinical Modification (ICD-9-CM). For procedure coding the Canadian Classification of Diagnostic, Therapeutic,

and Surgical Procedures (CCP) and Volume 3 of ICD-9-CM were in use. Starting in fiscal 2001/2002, Canada introduced new classifications for diagnosis and interventions, the International Statistical Classification of Diseases and Related Health Problems, Tenth Revision—Canadian Modification (ICD-10-CA) and the Canadian Classification of Health Interventions (CCI). The implementation is being phased in by province starting in fiscal 2001/2002 and ending in fiscal year 2004/2005.

CIHI plays a central and key role in the development of guidelines for coding of diagnoses and procedures involving these classification systems. The guidelines specify the use of the classification systems and include examples to further illustrate their appropriate application. Under the direction of the National Coding Advisory Committee[3], their development has been improved as of March 2001. Effective April 1, 2003, the Coding Guidelines will become known as the Canadian Coding Standards for ICD-10-CA and CCI. The standards are being updated on an annual basis, based on coding queries, hospital feedback and the results of the DAD Data Quality Study and may be amended, if needed for additional clarification. Endorsement of CIHI coding standards by the Provincial/Territorial Ministries of Health will facilitate their proper adherence by hospitals.

CIHI provides support to the health records staff and their managers at hospitals through on-going education workshops and by responding to coding queries through the web based Coding Query Database. Communication media such as the CIHI Bulletins are used on an as needed basis. The Bulletins are kept for reference purposes on CIHI's Web site as well.

Data Quality

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 users are provided with good information about data quality. To this end, CIHI established a comprehensive and systematic data quality program. Its purpose is to enhance the quality of existing data holdings and ensure that new data holdings and information products meet standards of quality consistent with CIHI's program objectives and CIHI's commitment to excellence. The data quality program involves the implementation of a data quality framework[4] and special studies focusing on specific data quality issues.

The first special data quality study undertaken by CIHI is of the DAD, given its size, coverage and importance. The DAD Data Quality (DQ) Re-abstraction Study is a large-scale multi-year study that uses a statistical sampling methodology to reliably measure the accuracy of selected non-medical and clinical administrative data contained in the DAD at a national level.

[3] The National Coding Advisory Committee is composed of hospitals and/or ministry representatives from each provincial/territorial jurisdiction. Their mandate is to review and approve with 100% agreement all standards before they are published by CIHI.

[4] The CIHI Data Quality Framework involves 24 characteristics relating to five data quality dimensions of accuracy, timeliness, relevance, comparability, and usability.

DAD DQ Re-abstraction Study

Goal and Objectives of Study

The goal of the DAD Data Quality Study is to evaluate and measure the accuracy of selected data contained in the DAD.

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 included in CIHI's Health Indicators Framework;
4. Evaluate and measure the extent to which diagnoses and procedures are not coded according to CIHI guidelines and identify where additional coding guidelines may be required;
5. Facilitate the evaluation of the change to new diagnosis and intervention classification standards (i.e. ICD-10-CA/CCI); and
6. Assess whether any of the above evaluations have an impact on the assignment of Case Mix Group (CMG) and Length of Stay (LOS).

This report addresses the first five objectives while the sixth objective will be addressed in the forthcoming report focusing on complexity level coding.

The health indicators[5] evaluated in the first two years of the study were selected from the CIHI Health Indicator Framework through consultation with staff within CIHI's Health Reports and Analysis section. In year one, the study focused on the following diagnoses and procedures relevant to the health indicators:

- 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

The first year of the study also provided an opportunity for collaboration with the Canadian Perinatal Surveillance System (CPSS) of the Bureau of Reproductive and Child Health which is part of Health Canada's Centre for Healthy Human Development[6] (HHD). The CPSS is part of Health Canada's initiative to strengthen Canada's health surveillance

[5] Note that the health indicators reported in the CIHI Annual Reports are expressed as rates, and also may have other specific criteria. That is, in addition to the conditions that are relevant to the health indicator, further criteria are applied to case selection for the calculation of the health indicator—an example is the length of stay and the death code that is used to calculate the 30-day AMI In-hospital mortality rate.

[6] The centre was previously known as the Laboratory Centre for Disease Control (LCDC).

capacity. Its long-term goal is to create a national database that provides the data elements required to monitor a comprehensive set of perinatal indicators.
The following indicators, developed and defined by the Canadian Perinatal Surveillance System, were therefore included as part of the first year of the DAD Study:

- 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 the year 1 indicators was also randomly selected from the DAD for fiscal year 1999/2000. 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 the year 1 indicators.

In year two, the study focused on the following diagnoses and procedures (conditions) relevant to the health indicators:

- Acute Myocardial Infarction
- Hip Fracture
- Hysterectomy
- Total Knee Replacement

Similarly to year 1, a sample of charts that did not contain any of the year 2 indicators was also randomly selected from the DAD for fiscal year 2000-2001. For ease of reference, this sample is defined as: Not assigned to any of the year 2 indicators.

Details of the specific conditions and procedures included in each indicator can be found in Appendix B, Health and CPSS Indicators.

Study Methodology

Sample Design and Methodology to Produce Estimates

Over the two years of the study a total of 29 facilities[7] participated in the study, allowing for the re-abstraction of a total of 4,292 charts[8]. The overall facility response rate for participation in the study was greater than expected at 84%. The total number of re-abstracted charts exceeded the total target number of charts by 5%.

The study features a multi-stage stratified sample design. The first stage involved the random selection of facilities and the second stage involved the random selection of charts from participating facilities. For the first year of the study, the facilities were stratified by geography, size and type of hospital and the charts were stratified by the conditions or procedures relating to the first year Health and CPSS indicators. For the second year, the facilities were stratified by geography and size and the charts were stratified by the conditions or procedures relating to the second year Health indicators. Each abstract was assigned to an indicator based on the algorithm described in Appendix B, Health and CPSS indicators. In cases where an abstract could be assigned to more than one indicator, the condition with less prevalence was given priority for sample selection purposes only. During analysis of the data, all diagnoses and procedures were reviewed.

The size of the sample of charts was calculated for each indicator by using standard formula and by assuming (as a starting point) that at the national level the proportion of charts for each condition relating to the indicator that contained a discrepancy was 15% and that the reliability required for the sample was a coefficient of variation of 16.5% (that is, a standard error of 2.5%). The minimum sample size was increased by 20% to account for possible chart unavailability and to efficiently utilize re-abstractor resources. The number of charts was allocated (as far as possible) equally among the participating facilities. Under this design, each sampled chart has an unequal probability of selection.

The sampling weight for each chart was calculated that reflected its overall probability of selection. The sampling weight was then used to calculate the estimated count of discrepancies. The discrepancy rate for each data element[9] in the study is the percentage that the estimated count is of the appropriate total. The combined percentage is the total estimated count divided by the total denominator and is a weighted average of the year 1 and year 2 discrepancy rates.

[7] The target population for the study included all acute care facilities submitting data to the DAD. Facilities from Quebec and from Manitoba were excluded from the study because there is no provincial requirement for facilities to submit abstracts to the DAD. Submitting facilities from the 3 territories were excluded for travel/cost reasons. The DAD for fiscal year 1999/2000 was used for the first year of the study, and the DAD for fiscal year 2000-2001 was used for the second year of the study.

[8] See Table 1, Response Rates, in Appendix F, Detailed Tables Combined Year 1 and Year 2, Appendix G, Detailed Tables Year 1, and Appendix H, Detailed Tables Year 2, for additional specific details.

[9] The calculation of the percentage discrepancy for the conditions relevant to health indicators was finalized to use the estimated counts for the denominators, as well as the numerator.

The estimates are subject to sampling and non-sampling error. Since errors may occur at every phase of a study, considerable time and effort was spent to minimize non-sampling errors by implementing quality assurance procedures throughout the study.

Data Collection (Re-abstraction of Charts)

CIHI classification specialists[10] re-abstracted the data for the study by returning to the original source of the data on site at each facility. They re-abstracted at each facility for a one-week period during September to November 2000 for the first year of the study and during September to November 2001 for the second year of the study. The specific data elements that were re-abstracted are listed in Appendix C, Re-abstracted Data Elements. The re-abstracted information was then compared with the information contained in the original submission to the DAD at the highest level of code specificity.

Privacy, Confidentiality and Security

In order to respect personal privacy and to safeguard the confidentiality of individual records and facilities, a number of procedures were developed and adhered to throughout the study. CIHI classification specialists signed confidentiality agreements with the participating hospitals and CIHI agreed not to release the names of the participating facilities without their permission. All results, other than the reports provided to the participating facilities, are discussed in an aggregate form only, so that it is not possible to identify individual patients, physicians or institutions included in the study.

Identification of Discrepancies and Reasons

All clinical information such as diagnoses and procedures was re-abstracted blindly (i.e. without viewing the original abstracted data) at the highest level of code specificity. Objective non-medical information (such as date of admission, date of discharge, etc.) was viewed and compared to the original data, and then a match or a discrepancy was identified. If a discrepancy occurred, the non-medical data were re-abstracted.

For each discrepancy, both medical and non-medical, the type of discrepancy and a possible reason were assigned by the re-abstractor. More than one reason could be assigned per discrepancy. A complete list of discrepancy types and reason codes can be found in Appendix D, Discrepancy Codes, and Appendix E, Discrepancy Reason Codes, respectively. The example in Table 1 below illustrates this for Acute Myocardial Infarction (AMI).

[10] CIHI *Classification Specialists* are certified with the Canadian College of Health Record 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.

Table 1: Example

A patient was admitted to hospital with a diagnosis of Acute Myocardial Infarction (AMI). Consultation from cardiology states Non Q Wave M I.		
	Original abstract	**Re-abstract**
Most Responsible Diagnosis ICD-9-CM	410.91 Acute Myocardial Infarction, Not Otherwise Specified	410.71 Non Q Wave Myocardial Infarction
A discrepancy is identified such that the original DAD submission had an error that apparently was the result of the coder missing information when the chart was originally coded: • Discrepancy 5—Different diagnosis code used to identify same condition • Reason code P—Information on chart missed		

The discrepancy occurred in the above example because the codes were verified to the highest level of specificity. Although the original coder and the re-abstractor both agreed that the patient suffered an acute myocardial infarction the re-abstractor was able to find documentation within the patient record to specifically identify the type of injury. The original coder did not find this information and selected "unspecified".

It is to be noted that while there is a discrepancy for the diagnosis codes there would not be a discrepancy for the AMI indicator. This is because both the original and re-abstracted diagnosis codes indicate AMI (as shown in the algorithm in Appendix B, Health and CPSS Indicators).

In some cases the reason for the discrepancy between the original data and the re-abstracted data may be due to the unavailability of supporting documentation at the time it was originally submitted to the DAD; or conversely portions of the source documentation may not have been available to the re-abstractor. In other situations, the reason for the discrepancy may have been of a less critical nature, but it was captured because of its potential benefit in coding guideline development. These findings that are called "Type B Discrepancies" in Appendix E, Discrepancy Reason Codes, are not included in the findings contained in this report.

National Findings

This section highlights the combined findings from the first year (DAD fiscal year 1999/2000) and the second year (DAD fiscal year 2000/2001) for the demographic, administrative, and clinical (diagnoses and procedures) data. The findings[11] for the conditions related to the health indicators were not combined, as the health indicators were specific to each year of the study.

Detailed tables of the findings are contained in Appendix F, Detailed Tables Combined Year 1 and Year 2, Appendix G, Detailed Tables Year 1, and Appendix H, Detailed Tables Year 2.

Discussion and general observations based on the findings are presented in the summary to this section of the report.

Health Indicators

The conditions relevant to the health indicators related to pre-booked procedures such as coronary artery bypass grafts, hip and knee replacements, and those where the classification is less complex, as in cesarean sections and vaginal births after cesarean section, were the most accurately coded. Diagnoses with more complex treatment protocols, those that are less easily defined such as acute myocardial infarction, pneumonia and influenza, and injuries, and those that include a large number of different kinds of diagnoses such as ambulatory care sensitive conditions (ACSC) showed a higher degree of discrepancies.

The following conditions had fewer than 5% discrepancies: cesarean section; coronary artery bypass graft; hysterectomy, total knee replacement, vaginal birth after cesarean; and total hip replacement. These conditions had greater than 5% discrepancies: diagnoses relevant to ACSC (10.7%); acute myocardial infarction (8.9%), hip fracture (6.0%); hospitalization due to pneumonia and influenza (6.9%); and injury hospitalization (5.3%).

[11] Discrepancies for the diagnosis and procedures related to Health Indicators are false positives. These occur when the diagnosis and procedure codes in the original abstract met the criteria for inclusion but the criteria were not met by the re-abstracted codes. That is, the original diagnosis and procedure codes were part of the algorithm (listed in Appendix B, Health and CPSS indicators) for the indicator whereas the re-abstracted diagnosis and procedures were not. The percentage of false negatives was also calculated—these are not included in the report as the study was not designed to measure these reliably. It is to be noted that at the aggregate level there may be a counteraction of the false negatives and positives. Other related measures such as sensitivity and specificity are being examined and are available upon request by contacting dataquality@cihi.ca. Further it should be noted that while these conditions are used in the calculation of many indicators, the discrepancy rates should not be interpreted as discrepancy rates for the health indicators, since there are additional criteria that are applied in the calculation of the health indicator.

Through additional analysis of the national findings it was found that a small proportion of facilities with unusually high discrepancy rates were contributing substantially to the national estimates of the discrepancy rate. That is:

- For the 10.7% national discrepancy rate for diagnoses relevant to ACSC, one of the 18 facilities had a discrepancy rate of 36.4% while three quarters of the 18 institutions had discrepancy rates of 10% or less.
- For the 8.9% national discrepancy rate for acute myocardial infarctions (AMI) conditions, one of 11 facilities had a 16.9% discrepancy rate, while the remaining ten facilities had discrepancy rates of 10% or less.
- For the 6.0% national discrepancy rate for hip fractures, one of 11 facilities with a discrepancy rate greater than 7%, (the same institution with the unusually high AMI discrepancy rate) while four of the 11 institutions had no hip fracture indicator discrepancies.
- For the 6.9% national discrepancy rate for hospitalization due to pneumonia and influenza, and the 5.2% national discrepancy rate for injury hospitalizations, 8 of 18 facilities and 9 out of 18 facilities respectively had no discrepancies. Three of the 18 facilities for both indicators had high discrepancy rates ranging from 18.2% to 50.0%.

Demographic and Non-Medical Discrepancies and Reasons

For the data elements that are mandatory[12] in all provinces, many demographic and non-medical data elements had a discrepancy rate of less than 1%. See Table 3A, Estimated Demographic and Non-Medical Discrepancies by Data Element, for these data elements. Three mandatory data elements had a discrepancy rate between 1 and 5%: Institution to, Institution from and Admit hour. Discrepancies for *Admission Category* occurred 15.5% of the time and for *Discharge Hour* for 10.0% of the time.

Many of the *Admission Category* discrepancies arose from situations where patients were admitted to hospital through the Emergency department. In some cases hospitals simply identified all of these patients as "*Emergent*" when only patients with life-threatening conditions should be designated as such. In addition, there was some difficulty in identifying proper admission codes for obstetrical patients. While 17 of the 29 facilities had discrepancy rates of less than 5%, 4 of the 29 had discrepancy rates greater than 40%. The revised DAD abstract, which is being phased in by province with ICD-10-CA/CCI starting in fiscal year 2001/2002, has merged "*Urgent*" and "*Emergent*" for the *Admission Category* data element. This merger should reduce the national discrepancy rate. At the same time though, it is recognized that there is a need to add more general examples in the DAD abstracting manual.

[12] Findings for non-mandatory fields will be reported separately, as these are not mandatory in every provincial jurisdiction.

Discrepancies in the *Discharge Hour* were commonly found to be the result of discharge times being electronically downloaded into the abstract, which did not match the time that the patient actually left the floor as indicated by the nursing notes in the chart. It should be noted that the demographic and non-medical discrepancies were similar for both years of the study with the exceptions of postal code and entry code. *Postal Code* discrepancies were: Year 1–9.0%; Year 2–0.7%; Year 1 + Year 2–5.0%. One of 18 facilities had a postal code discrepancy rate of 73.1% in year 1—this field was being used incorrectly for the facility's internal purposes. *Entry Code* discrepancies were: Year 1–6.5%; Year 2–3.6%; Year 1 + Year 2–5.1%. In the first year of the study, one of the 18 facilities had a discrepancy rate of 49% that contributed substantially to the national estimate, with the other facilities all having discrepancy rates of less than 10%. In the second year all 11 facilities had discrepancy rates of less than 10%.

Of the total reasons for all of the non-medical and demographic data elements discrepancies, the top three reasons were:

	Year 1	Year 2	Year 1 + Year 2
The original coder missing information that was documented on the chart	24.6%	35.4%	28.7%
Inconsistent/conflicting information	2.6%	11.5%	6.0%
Incorrect data download	31.1%	1.6%	19.8%

Details are available in Table 3B, Estimated Demographic and Non-Medical Discrepancies by Reason Code in Appendix F, Detailed Tables Combined Year 1 and Year 2

Diagnosis Code/Typing Discrepancies and Reasons

The diagnosis codes in this study were compared using four different elements: the prefix, the actual code, the suffix, and the diagnosis type. The following diagnosis discrepancies were identified in the study:

- *Most Responsible Diagnosis (MRDx)* 12.8%,
- *Comorbid Condition (CC) Diagnosis typing* 15.5%,
- *Comorbid Condition (CC) Diagnosis* 23.2%, and *Any diagnosis different* 6.0% (regardless of whether MRDx or CC)
- E-code, for external cause of event for trauma cases, had discrepancies of 10.2% based on year 2 of the study only.

It is to be noted that the findings for *Comorbid Condition (CC) Diagnosis typing*, and *Comorbid Condition (CC) Diagnosis* discrepancies were not similar for each year of the study. For *Comorbid Condition (CC) Diagnosis typing* there were 11.0% discrepancies in year 1 and 18.9% in year 2. For *Comorbid Condition (CC) Diagnosis* discrepancies there were 31.2% discrepancies in year 1 and 17.0% year 2.

Through the additional analysis of the national findings it was again found that a small proportion of facilities with unusually high discrepancy rates were contributing substantially to the national estimates of the discrepancy rate. That is:

- For the 12.8% national discrepancy rate for *MRDx*, four of 29 facilities had a discrepancy rate of between 20 and 42.8%, while most (19 of 29) of the facilities had discrepancy rates between 5 and 20% and 6 of the 29 institutions had a discrepancy rate less than 5%.
- There was a substantial difference in the *CC Diagnosis typing* discrepancy rates for the first and second years of the study, with 8 of the 18 facilities in the first year of the study having less than a 5% rate and only 2 of the 11 institutions in year 2 having a discrepancy rate less than 5%. The *CC Diagnosis typing* discrepancy rates ranged from less than 1% up to 48.2%.
- The discrepancy rates for *Comorbid Condition (CC) Diagnosis* also varied in both years, ranging from less than 1% to well over 50%. Seven of the twenty-nine institutions had a discrepancy rate of less than 10%.

It should be noted that almost half of the *MRDx* discrepancies occurred when the re-abstractor coded a diagnosis as the *MRDx* when it had been coded as another diagnosis type in the original abstract.

For the other diagnosis discrepancies, the majority of the discrepancies fell into one of three areas:

- The original coder captured a condition that the re-abstractor did not feel was significant;
- The re-abstractor coded a significant condition that the original coder did not; and
- The re-abstractor and original coder used a different code to represent the same condition (as illustrated in the example in Table 1, in the Identification of Discrepancies and Reasons section).

Of the total reasons for all the diagnosis discrepancies, the three main reasons were:

	Year 1	**Year 2**	**Year 1 + Year 2**
The re-abstractor disagreeing that the diagnosis significantly impacted on the treatment and/or length of stay	20.8%	30.8%	25.8%
Different interpretations of the documentation	17.2%	19.8%	18.5%
The original coder missing information that was documented on the chart	23.6%	10.9%	17.0%

Details are available in Table 4B: Diagnosis Discrepancies by Reason Code in Appendix F, Detailed Tables Year 1 and Year 2

Procedure Code Discrepancies and Reasons

Discrepancies related to procedures[13] were divided into categories. The following discrepancy rates were identified: *Principal Procedure (PP)* 7.0%, *other procedures* 19.7%, *differences in the procedure code itself* 4.9%.

The most common procedure discrepancies were:

- The original coder captured a procedure that the re-abstractor did not;
- The re-abstractor coded a procedure that the original coder did not; and
- The re-abstractor and original coder used a different code to represent the same procedure.

It is to be noted that the findings were not similar for *Principal Procedure (PP) and other procedures* for the two years of the study. For *Principal Procedure (PP)* the combined year 1 and year 2 rate is 7.0%, where the year 1 rate is 10.0% and the year 2 rate is 3.3%. For *other procedures* the combined year 1 and year 2 rate is 19.7%, where the year 1 rate is 23.3% and 14.6% in year 2.

A large portion of the difference between the year one and two estimates for principal procedures can be attributed to one institution in year one that had a 75.6% discrepancy rate for principal procedures. Many of the discrepancies at this institution are the result of obstetrical procedures being coded incorrectly. In year 2, no institution had more than 10% discrepancies for principal procedures. For other procedures, the institution discrepancy rates varied significantly in both years, with values ranging from less than 1% up to 77.9%. In year 1, seven of the 18 institutions had more than 30% discrepancies while 4 of the 18 had less than 5% discrepancies. In year 2, one of 11 institutions had a discrepancy rate greater than 30%.

Of the total reasons for all of the procedure discrepancies, the top three reasons were:

	Year 1	Year 2	Year 1 + Year 2
The original coder missed information that was documented on the chart	34.7%	40.3%	36.5%
Different interpretations of the documentation	24.1%	38.0%	27.5%
Coding Error—original coding did not follow code book properly	12.1%	6.3%	10.4%

Details are available in Table 5B: Procedure Discrepancies by Reason Code in Appendix F, Detailed Tables Year 1 and Year 2

[13] Findings for Anaesthetic fields will be reported separately, as these are not mandatory in every provincial jurisdiction.

Discussion and Recommendations

In summary, the combined findings provide a baseline measurement of the accuracy of the DAD for fiscal years 1999/2000 and 2000/2001, prior to the implementation of the ICD-10-CA/CCI classification systems and revisions to the DAD abstract introduced in fiscal year 2001/2002. While the revisions and additional edits should result in improvements in accuracy and consistency of the DAD, one of the objectives of the third year of the DAD Data Quality Study will be to evaluate the extent to which such improvements have been achieved.

While the extent of coding discrepancies is substantial in some cases, it is to be noted that there are many hospitals where the discrepancies are minimal and coding guidelines are being followed appropriately. Such organizations are invaluable for identifying best practices.

The study identifies specific areas that require improvement. There are a number of common reasons identified by the re-abstractors across the spectrum of discrepancies that relate to individual hospital practices. Many of these relate to the variation in hospital use of CIHI guidelines (which will become known as standards as of April 2003) for clinical coding, as well as the variety in the quality of the documentation contained in the patient charts. The cause or source of error may be due in some cases to individual hospitals misunderstanding, misinterpreting or not being aware of the coding guidelines, while in other cases hospitals may be choosing not to follow the current guidelines in favour of internal guidelines. In the interest of national data standardization and comparative reporting, internal facility guidelines should not supersede CIHI coding guidelines (standards).

The study findings also indicate that documentation in some charts may have been incomplete[14] at the time that the abstract data was submitted to CIHI. In other cases it would appear that that the abstract was coded from a complete chart and the original coder missed the information.

However, there is clearly a need for:

1. Hospitals and Ministries of Health to address the data quality issues that are present either at their facility, or at the facilities within the provincial/territorial jurisdictions.
2. Adherence to coding standards to be monitored systematically and results to be provided to the hospitals and ministries.
3. CIHI to continue to enhance and develop guidelines/standards for clinical coding including diagnosis typing.

[14] It was noted in an earlier investigation done by CIHI into the timeliness of the DAD that 30% of large hospitals were submitting such abstracts.

4. CIHI education workshops:
    - To be enhanced and/or revised to emphasize the
        - importance of standardizing individual facility internal data collection processes;
        - importance of maintaining high quality documentation on inpatient charts; and
        - the provincial/territorial and national use of their individual hospital data .
    - To be provided on the availability and use of current CIHI tools;
    - To be more easily available through either teleconference, video or the Internet.
5. Standards to be established for acceptable levels of data quality.

It is only by addressing these items that the usability of the DAD can be enhanced.

In addition to providing valuable input into the annual coding guideline development process and enhancing the content of our education workshops, the study findings will also be used to develop strategies and action plans aimed at improving the overall quality of the DAD data. More specifically, CIHI will be:

- Establishing an internal working group to identify additional CIHI DAD edits[15] and improvements to the DAD abstracting manual, as well as explore options to deal with the submission of abstracts based on incomplete charts.
- Developing data quality documentation relating to the use of the health indicators included in this study and publishing such documentation in all relevant publications, including CIHI's Health Care in Canada annual report.
- Examining the feasibility of determining, through relevant research, acceptable levels of quality for health data and their determinants (or factors).
- Examining the feasibility of using the study methodology for ongoing quality monitoring (subject to the availability of resources).

In addition to the above, CIHI is also recommending that:

A national working group, with representation from CIHI, health records professionals, Ministries of Health, be established to examine and explore options related to the ongoing monitoring of compliance with CIHI Coding Standards. More specifically, this group would assist in identifying:

- Appropriate mechanisms to secure endorsement of CIHI coding guidelines/standards by the provincial/territorial Ministries of Health;
- Criteria (or indicators) that can be used to flag hospitals that may have data quality issues or that may not be complying with established coding guidelines/standards;
- Appropriate mechanisms (such as a systematic re-abstraction program) to monitor compliance with coding guidelines/standards;
- Enhancements to the coding guidelines/standards and/or coding practices;

[15] For example, editing diagnosis typing for manifestation/asterisk codes or other combination code scenarios.

- Best coding practices and effective strategies for disseminating these among facilities;
- Gaps in provincial/territorial policies relating to quality of health data; and
- Specific hospital responsibilities for ensuring quality chart documentation at the time of submission of the abstract to CIHI.

References

- Brown A., and Richards, J., *" The Data Quality Study of the Canadian Discharge Database",* Proceedings of Statistics Canada Symposium 2001, Achieving Data Quality in a Statistical Agency: a Methodological Perspective, (May 2001).

- Canadian Institute for Health Information*, "DAD Abstracting Manual"*, www.cihi.ca

- Canadian Institute for Health Information (June 1, 2000), *Bulletin, "DAD Data Quality Study"*, www.cihi.ca

- Canadian Institute for Health Information (2000), "*Improving Timeliness of the Discharge Abstract Database Data Quality Study"*, www.cihi.ca

- Canadian Institute for Health Information (2001), "*Health Care in Canada",* www.cihi.ca

- Canadian Institute for Health Information (2001), "*Products and Services Catalogue",* www.cihi.ca

- Canadian Institute for Health Information (2002), "*Discharge Abstract Database Data Quality Study, Preliminary Year 1 Findings",* www.cihi.ca

- Ontario Hospital Association, Ontario Ministry of Health, Hospital Medical Records Institute (April 1991), *"Report of the Ontario Data Quality Re-abstracting Study".*

- Ontario Clinical Data Quality Task Force, August 2002, *"Understanding and Improving the Quality of Health Care in Ontario—A Report and Recommendations by the Ontario Clinical Data Quality Task Force"*

Appendix A

PIRC Indicators

Performance Indicators Reporting Committee (PIRC)

Indicators provided by CIHI

30 day acute myocardial infarction in-hospital mortality rate
30 day stroke in hospital mortality rate
Total hip replacement rate
Total knee replacement rate
Hospital re-admission rate for acute myocardial infarction
Hospital re-admission rate for pneumonia
Ambulatory care sensitive conditions

Appendix B

Health and CPSS Indicators

Health[16] and CPSS Indicators

Health Indicator Algorithms for the first year of the study

Description	ICD-9 (Dx) or CCP (procedure)[17]
Ambulatory care sensitive conditions (ACSC). This indicator consists of the following conditions: diabetes, asthma, alcohol or drug psychoses, and non-dependant abuse of drugs, depression, and hypertension.	401-405, 291, 292, 303-305, 311, 300, 250, 493
Cesarean section	86.0, 86.1, 86.2, 86.8, 86.9
Coronary artery bypass graft surgery (CABG)	48.1
Hospitalization due to pneumonia and influenza	480, 481, 482, 483, 484, 485, 486, 487
Injury hospitalization consisted of the specified E-codes in any diagnosis position:	E800, E801, E802, E803, E804, E805, E806, E807, E810, E811, E812, E813, E814, E815, E816, E817, E818, E819, E820, E821, E822, E823, E824, E825, E826, E827, E828, E829, E830, E831, E832, E833, E834, E835, E836, E837, E838, E840, E841, E842, E843, E844, E845, E846, E847, E848, E880, E881, E882, E883, E884, E885, E886, E887, E888, E890, E891, E892, E893, E894, E895, E896, E897, E898, E899, E900, E901, E902, E906, E907, E908, E909, E910, E913, E914, E915, E916, E917, E918, E919, E920, E921, E922, E923, E924, E925, E926, E927, E928, E953, E954, E955, E956, E957, E958, E960, E961, E963, E964, E965, E966, E967, E968, E970, E971, E972, E973, E974, E975, E976, E978, E983, E984, E985, E986, E987, E988, E990, E991, E992, E993, E994, E995, E996, E997, E998
Total hip replacement	93.51, 93.59
Vaginal births after cesarean section (VBAC)	654.2
Not assigned to any year 1 indicator	All other abstracts not assigned to any year 1 health or CPSS indicator

[16] Note that the health indicators reported in the CIHI Annual Reports are expressed as rates, and may have other specific additional criteria, for example, 30 day AMI In hospital mortality rate.

[17] Additional details regarding the specific algorithm or the ICD-9-CM/CCI equivalent codes are available upon request by contacting dataquality@cihi.ca.

Canadian Perinatal Surveillance System Indicator Algorithms

Indicator (acronym): Description	ICD-9(Dx) OR CCP (procedure)[17]	CMG
Rare Congenital Anomalies		
Anencephalus and similar anomalies (ANEN)	740.0, 740.1, 740.2	
Anomalies of abdominal wall (AAW)	756.7	
Cleft palate (CLEFTP)	749.0	
Cleft palate with cleft lip (CPCL)	749.2	
Congenital hydrocephalus(CH)	742.3	
Down's syndrome (DS)	758.0	
Encephalocele (ENCE)	742.0	
Hypoplastic left heart syndrome (HLHS)	746.7	
Intestinal anorectal atresia and stenosis (IAAS)	751.2	
Limb reduction anomalies (LRA)	755.2, 755.3, 755.4	
Renal agenesis and dysgensis (RAD)	753.0	
Spina bifida (SB)	741.0, 741.1, 741.2, 741.3, 741.4, 741.5, 741.6, 741.7, 741.8, 741.9	
Tracheo-esophageal fistula, esohageal atresia and stenosis (TEFEAS)	750.3	
Transposition of great vessels (TGV)	745.1	
Rare Maternal Conditions		
Amniotic fluid embolism (AFE)	673.1	
Anaesthesia complications (AC)	668.0, 668.1, 668.2, 668.8, 668.9	
Cerebrovascular disorders (CD)	674.0, 430, 431, 432, 433, 434, 435, 436, 437, 438	>600 and <605 or>605 and<612
Eclampsia (ECL)	642.6	
Rupture of the uterus (ROU)	665.0, 665.1	
Obstetric septic shock (OSS)	634.5, 635.5, 636.5, 637.5, 638.5, 639.5, 669.1	
Obstetrical pulmonary embolism (OPE)	634.6, 635.6, 636.6, 637.6, 638.6, 639.6, 673.0, 673.2, 673.3, 673.8	>600 and <605 or>605 and <612)
Rare Neonatal Conditions		
Brachial plexus injury (BPI)	767.6	
Exchange transfusion (EXTRAN)	13.01	
Fracture of the clavicle (FC)	767.2	
Haemorrhagic disease of the newborn (HDN)	776.0	
Intraventricular haemorrhage (IH)	772.1	
Massive aspiration syndrome (MAS)	770.1	
Necrotizing entercolitis (NE)	777.5	
Seizures (SEIZ)	779.0	
Severe birth asphyxia (SBA)	768.5	
Respiratory distress syndrome (RDS)	769	
Third degree perineal laceration (TDPL)	664.2	
Other maternal or neonatal conditions	Other maternal or neonatal conditions not including those above	

Health Indicator Algorithms for the second year of the study

Description	ICD-9 (Dx) or CCP (procedure)[17]
Acute myocardial infarction (AMI)	410
Hysterectomy	80.2, 80.3,80.4,80.5,80.6,80.7
Hip fracture	820.0,820.1,820.2,820.3,820.8,820.9
Total knee replacement	93.41
Not assigned to any year 2 indicator	All other abstracts not assigned to any year 2 health indicator.

Appendix C

Re-abstracted Data Elements

Re-abstracted Data Elements

(Based on Fiscal 1999/2000 DAD Data Elements)

Group & Field No.	Data Element
01 11	Second Chart/Register Number
03 01	Health Care Number
03 02	Postal Code
03 04	Gender
03 05	Prov/Terr Issuing HCN
03 08	Birth date
03 09	Estimated Birth date
04 01	Admit Date
04 02	Admit Hour
04 04	Institution From
04 05	Admission Category
04 06	Entry Code
04 07	Admit by Ambulance
04 08	Readmission Code
04 09	Unplanned Readmission Code
04 10	Wait Time in Emergency (min.)
05 01	Discharge Date
05 02	Discharge Hour
05 04	Institution To

Group & Field No.	Data Element
06 01	Exit Alive
06 04-11	Death Code
07 03	Weight (0-29 days on admission)
07 04	Abstract Overflow
10 01	Diagnosis Prefix
10 02	Diagnosis Code
10 03	Diagnosis Suffix
10 04	Diagnosis Type
11 01	Procedure Date
11 02	Procedure Code
11 03	Procedure Suffix
11 10	Anaesthetic Technique
11 11	Out of Hospital Institution Number
11 12	Unplanned Return to O.R.
13 01	SCU Death Indicator
13 02	SCU Unit Number
13 03	SCU Days
17 01-07	Blood Information
18 01-05	Therapeutic Abortion Information

Appendix D

Discrepancy Codes

Discrepancy Codes

Non-medical (clinical) Data

1. **Entry missing**. Re-abstractor captured data not in database
2. **Entry not coded by re-abstractor**. Re-abstractor did not capture data that was in database
3. **Entry different**. Re-abstractor captured data that is different than the database

Diagnosis codes

4. **Diagnosis prefix/suffix different**. Either database or re-abstractor has coded prefix/suffix that the other has not.
5. **Different diagnosis code.** Different codes used to identify same condition.
6. **MRDx coded as different type.** Re-abstractor coded as MRDx but coded in database as another diagnosis type.
7. **MRDx missing.** Re-abstractor coded as MRDx but does not appear in database at all.
8. **CC diagnosis coded as type 3.** Re-abstractor coded and typed as 1 or 2 but coded in database as a type 3.
9. **CC diagnosis missing.** Re-abstractor coded and typed as 1 or 2 but does not appear in database at all.
10. **Pre-admit comorbidity typed as post-admit.** Re-abstractor coded and typed as 1 but coded in database as a type 2.
11. **Post-admit comorbidity typed as MRDx.** Re-abstractor coded and typed as 2 but coded in database as MRDx.
12. **Post-admit comorbidity typed as pre-admit.** Re-abstractor coded and typed as 2 but coded in database as a type 1.
13. **Secondary diagnosis coded as the MRDx.** Re-abstractor coded as type 3 but coded in database as MRDx.
14. **Secondary diagnosis typed as CC diagnosis.** Re-abstractor coded as type 3 but coded in database as a type 1 or 2.
15. **Diagnosis not coded, typed as MRDx.** Re-abstractor did not code, but coded in database as MRDx.
16. **Diagnosis not coded, typed as CC diagnosis.** Re-abstractor did not code, but coded in database as a type 1 or 2.
17. **Not used.**
18. **Transfer Dx missing.** Re-abstractor coded transfer Dx, but does not appear in database.
19. **Diagnosis not coded, typed as transfer diagnosis.** Re-abstractor did not code, but coded in database as a transfer diagnosis.
20. **(Year2 only) E-code different.** Different e-code used to identify same cause.

Procedures codes

21. **Procedure code different.** Different codes used to identify same procedure.
22. **Principal procedure coded as "other" procedure.** Re-abstractor coded as principal procedure but appears in database as "other" procedure.
23. **Principal procedure missing.** Re-abstractor coded as principal procedure but does not appear in database at all.
24. **Other procedure missing.** Re-abstractor coded as other procedure but does not appear in database at all.
25. **Procedure not coded, original coded as principal procedure.** Re-abstractor did not code procedure, appears in database as principal procedure.
26. **Procedure not coded, original coded as other.** Re-abstractor did not code procedure, appears in database as other procedure.
27. **Anaesthetic type different.** Re-abstractor did not identify same anaesthetic type as in database.
28. **Anaesthetic type missing.** Re-abstractor identified anaesthetic type that does not appear in the database.
29. **Anaesthetic type not identified appears in database.** Re-abstractor did not identify anaesthetic type that appears in the database.
30. **(Year 2 only) Procedure Date different.** Different dates identified for the same procedure.

Appendix E

Discrepancy Reason Codes

Discrepancy Reason Codes

Type A Discrepancies

Reason Code	Reason
A	Transcription error—errors in transcription of numbers and/or letters. Includes abstracting errors.
B	Incomplete documentation available at time of original abstraction—only when clearly identifiable
D	Lack of code specificity. A case where a non-specific or "other/unspecified" codes was used when a more specific code is supported by the chart documentation.
E	Code specificity not supported by record. Cases where a very specific code is used which is not supported by chart documentation.
F	Different interpretation of documentation. Cases where error in interpretation of documentation in original abstract has resulted in incorrect code.
I	Diagnosis coded did not have significant impact on treatment and/or LOS. Cases where code is typed as significant (1 or 2) and re-abstractor does not agree the documented treatment warranted it.
K	Other grey area coding. Other cases where different interpretation of the documentation and guidelines may lead to discrepancies.
L	Inconsistent or conflicting documentation on paper chart
M	Coding contrary to CIHI guidelines—where clearly identifiable
N	Hospital policy. Cases where, after discussion with hospital staff, it is identified that a hospital-specific rule or policy has affected the original codes chosen and caused the discrepancy.
O	Coding error—not following code book properly. Cases where discrepancy is clearly the result of incorrect or incomplete code look-ups. This includes dagger/asterisk errors.
P	Information on chart missed. Cases where a code or data was not entered in spite of clear documentation on the chart.
Q	Mathematical/counting error. Cases where a mathematical calculation error has been made such as in SCU days or Waiting Time in Emergency.
R	Downloaded incorrectly. ADT download inconsistent with the rest of the chart.
V	Other. Any identifiable reason that cannot be categorized into the other reason codes.
W	No apparent reason. When the discrepancy cannot be categorized or explained by any of the above codes.
Z	(Year 2 only) Diagnosis had a significant impact on treatment and/or LOS.

Type B Discrepancies

Reason Code	Reason
C	Re-abstractor unable to access required information.
G	Different interpretation of documentation—either code correct. Documentation may be interpreted more than one way and it is difficult to determine which way is more correct, but neither can be said to be wrong.
H	Order of codes different—either order is correct. Cases where two or more diagnoses were of equal importance and either could have been MRDx.
J	Re-abstractor did not code procedure as it is optional.
S	Database data amended by CIHI edit. Data amended in database and different on chart.
U	Re-abstractor missed data and believes original submission was correct.
X	Not re-abstracted—not wrong to code.
Y	Not coded in DAD—not necessary to code.

Type B Discrepancy Example

A woman arrives at the hospital in labour. Her labour is augmented with syntocin, however her cervix fails to open more than 3 cm. In addition, it is noticed that the baby is having decelerations. She is therefore taken to the O.R. where a c-section is performed for dystocia, obstructed labour due to CPD and fetal distress.

Diagnosis Code	Original abstract	Re-abstract
MRDx	661.01 dystocia	660.11 obstructed labour due to CPD
Type 1 Diagnosis	660.11 obstructed labour due to CPD	661.01 dystocia
Type 1 Diagnosis	659.71 fetal distress	659.71 fetal distress

Since there are multiple reasons for the c-section, any of those above could be chosen as the MRDx, and none could be considered an "incorrect" choice. The resulting discrepancy and reason are:

- Discrepancy 6—MRDx as different type
- Reason code H—Order of codes different; either order correct

Appendix F

Detailed Tables
Combined Year 1 and Year 2

Table 1A—Facility Response Rates

Units	Combined Year 1 and 2		Year 1		Year 2	
	Count	Response Rate	Count	Response Rate	Count	Response Rate
# hospitals contacted	43		26		17	
# hospitals no response	3		2		1	
# hospitals declined	4		2		2	
# hospitals accepted	36	84%	22	85%	14	82%
# hospitals participating	29		18		11	

Table 1B—Summary Counts of Re-abstracted Charts

Combined Years 1 and 2

Indicators	Target Sample Size	Initial Sample Size[1]	Reabstracted Charts	% Response Rate[2]	% of Target[3]
Total All Years	**4,075**	**4,900**	**4,292**	**87.6**	**105.3**

Year 1

Indicators	Target Sample Size	Initial Sample Size[1]	Reabstracted Charts	% Response Rate[2]	% of Target[3]
Health Indicator					
Ambulatory care sens. cond.	250	296	272	91.9	108.8
Cesarean section	251	281	264	94.0	105.2
Coronary artery bypass graft	226	274	176	64.2	77.9
Hospitalization pneumonia	249	283	261	92.2	104.8
Injury hospitalization	251	290	269	92.8	107.2
Total hip replacement	242	279	224	80.3	92.6
Vaginal births after cesarean	236	266	209	78.6	88.6
CPSS Indicator					
Rare Congenital Anomalies	75	77	73	94.8	97.3
Rare Maternal Conditions	71	80	69	86.3	97.2
Rare Neonatal Conditions	73	93	77	82.8	105.5
Respiratory distress syn.	99	119	78	65.5	78.8
Third degree perineal lac.	207	250	208	83.2	100.5
Other Maternal Neonatal Cond.	225	350	296	84.6	131.6
Other	245	302	261	86.4	106.5
Total Year 1	**2,700**	**3,240**	**2,737**	**84.5**	**101.4**

Year 2

Indicators	Target Sample Size	Initial Sample Size[1]	Reabstracted Charts	% Response Rate[2]	% of Target[3]
Acute Myocardial Infarction	275	340	307	90.3	111.6
Hysterectomy	275	330	316	95.8	114.9
Hip fracture	275	330	310	93.9	112.7
Total knee replacement	275	330	308	93.3	112.0
Not assigned to any year 2 indicator	275	330	314	95.2	114.2
Total Year 2	**1,375**	**1,660**	**1,555**	**93.7**	**113.1**

Table Notes:

1. The initial sample is the target sample increased by 10% for non-response and a further 10% for possible situations of better than expected productivity by the re-abstractor.
2. The response rate is calculated using the initial sample size.
3. The response rate is calculated using the target sample size.

Table 2—Estimated False Positives[18] for Health and CPSS Indicators

Year 1

Indicator	Sample Count Re-abstracted	False Positives[1]					
		Sample Count	Est. Count[2]	Sample %	Est. %	Lower Bound of 95% Confidence Interval	Upper Bound of 95% Confidence Interval
Health Indicator							
Ambulatory care sens. cond.	272	31	10,200	11.4	10.7	8.7	12.7
Cesarean section	308	1	100	0.4	0.1	0.0	0.2
Coronary artery bypass graft	176	1	200	0.6	0.8	0.1	1.5
Hospitalization pneumonia	261	23	4,800	8.8	6.9	5.0	8.8
Injury hospitalization	302	18	8,300	6.7	5.3	3.0	7.6
Total hip replacement	224	2	100	0.9	0.8	0.4	1.2
Vaginal births after cesarean	233	3	100	1.4	0.7	0.1	1.2
CPSS Indicator							
Third degree perineal lac.	213	12	2,500	5.8	23.0	22.0	24.1
Respiratory distress syn.	84	21	400	26.9	22.0	13.7	30.4
Rare Neonatal Conditions	81	11	300	13.6	15.8	9.1	22.6
Rare Maternal Conditions	69	11	200	15.9	23.4	20.6	26.2
Rare Congenital Anomalies	73	11	100	15.1	12.5	6.8	18.3
Other Maternal Neonatal Cond.	1,264	1	0	0.3	0.0	0.0	0.1
Other	261	10	24,400	3.8	1.5	1.1	1.8

Year 2

Indicator	Sample Count Re-abstracted	False Positives[1]					
		Sample Count	Est. Count[2]	Sample %	Est. %	Lower Bound of 95% Confidence Interval	Upper Bound of 95% Confidence Interval
Acute Myocardial Infarction	319	20	8,800	6.3	8.9	7.4	10.4
Hip Fracture	312	11	1,400	3.5	6.0	5.1	6.8
Hysterectomy	316	2	100	0.6	0.3	0.1	0.5
Total Knee Replacement	308	3	100	1.0	0.8	0.4	1.2
Other	314	2	10,300	0.6	0.4	0.2	0.7

Table Notes:

1. Estimated false positives are calculated using both an estimated numerator and estimated denominator.
2. Estimated counts have been rounded to units of 100, so as not to imply an unintended level of precision.

[18] Discrepancies for the diagnosis and procedures related to health indicators are false positives. These occur when the diagnosis and procedure codes in the original abstract met the criteria for inclusion but the criteria were not met by the re-abstracted codes, as listed in Appendix B, Health and CPSS Indicators. See the example contained in Table 1: Example, in the report section, **Identification of Discrepancies and Reasons**. Further note that while these conditions are used in the calculation of many indicators, the discrepancy rates should not be interpreted as discrepancy rates for the health indicators, since there are additional criteria that are applied in the calculation of the health indicator.

Table 3A—Estimated Demographic and Non-Medical Discrepancies by Data Element

Combined Year 1 and 2

	Non-Medical Discrepancies (%)			
Data Element[1,2]	**Entry Missing**	**Entry Not Reabstracted**	**Entry Different**	**Total**
Admission Category	0.0	0.0	15.5	15.5
Discharge Hour	0.0	0.5	9.5	10.0
Entry Code	0.0	0.0	5.1	5.1
Postal Code	0.0	0.0	5.0	5.0
Institution From	1.6	0.6	0.1	2.3
Institution To	1.0	0.9	0.4	2.3
Admit Hour	0.0	0.8	0.9	1.6
Health Care Number	0.0	0.0	0.3	0.3
Admit Date	0.0	0.0	0.1	0.1
Supplemental Death Code	0.0	0.0	0.1	0.1
Exit Alive	0.0	0.0	0.1	0.1
Birth Date	0.0	0.0	0.1	0.1
Provincial Issuing HCN	0.0	0.0	0.0	0.0
Discharge Date	0.0	0.0	0.0	0.0

Table Notes:

1. The population (N) is 2,300,518.
2. Data elements that are mandatory for the provinces in scope to the study.

Table 3B—Estimated Demographic and Non-Medical Discrepancies by Reason Code

Combined Year 1 and 2

	Non-Medical Discrepancies (%)			
Reason Code Description	**Entry Missing[1]**	**Entry Not Reabstracted[2]**	**Entry Different[3]**	**Total[4]**
P - Information on chart missed	72.5	0.0	28.0	28.7
M - Coding contrary to CIHI guidelines	2.4	19.4	25.2	23.7
R - Downloaded incorrectly	0.0	0.3	22.0	19.8
N - Hospital policy	21.0	5.4	9.6	10.0
W - No apparent reason	3.5	0.7	6.5	6.1
L - Inconsistent or conflicting information	0.0	65.3	2.8	6.0
F - Different interpretation - disagree	0.0	7.2	3.8	3.8
V - Other	0.5	1.5	1.1	1.1
A - Transcription error	0.0	0.0	0.5	0.5
K - Other grey area coding	0.0	0.0	0.2	0.2
E - Code specificity not supported	0.0	0.0	0.1	0.1
Q - Mathematical/counting error	0.0	0.0	0.0	0.0
B - Incomplete documentation	0.0	0.1	0.0	0.0
D - Lack of code specificity	0.0	0.0	0.0	0.0
O - Coding error	0.0	0.0	0.0	0.0
Total Reasons	**100.0**	**100.0**	**100.0**	**100.0**

Table Notes:

1. The estimated population (N) of reasons for entry missing discrepancies was 60,600.
2. The estimated population (N) of reasons for entry not re-abstracted discrepancies was 63,900.
3. The estimated population (N) of reasons for entry different discrepancies was 1,081,100.
4. The estimated population (N) of reasons for total non-medical discrepancies was 1,205,700.

Table 4A—Diagnosis Discrepancies

Combined Year 1 and 2

Diagnosis Discrepancies	Discrepancies (%)	
	Sample	Estimated Population
MRDx Discrepancies[1]		
6 - MRDx different type	3.8	5.8
7 - MRDx missing	3.2	2.9
11 - Post-admit as MRDx	0.1	0.0
13 - Secondary as MRDx	0.3	0.2
15 - Dx not coded, orig MRDx	3.3	4.0
Total MRDx Discrepancies	**10.6**	**12.8**
CC Diagnosis Typing Discrepancies[2]		
8 - CC Dx as type 3	2.2	1.6
10 - Pre-admit as post-admit	0.6	0.3
12 - Post-admit as pre-admit	0.5	0.3
14 - Secondary Dx as CC	10.0	13.2
Total CC Diagnosis Typing Discrepancies	**13.3**	**15.5**
CC Diagnosis Discrepancies[2]		
9 - CC Dx missing	10.6	8.9
16 - Dx not coded, orig CC	14.5	14.0
18 - Transfer Dx missing	0.2	0.1
19 - Dx not coded orig trans	0.1	0.2
Total CC Diagnosis Discrepancies	**25.4**	**23.2**
4 - Diagnosis Prefix/Suffix Different Discrepancies[3]	0.3	0.2
5 - Any Diagnosis Different Discrepancies[3]	7.2	6.1
20 - E-code Discrepancies[4]	11.1	10.2
Total Diagnosis Discrepancies	**37.7**	**34.8**

Table Notes:

1. The denominator for % discrepancy of MRDx was 2,146 for the sample and 2,565,000 for the estimated population.
2. The denominator for % discrepancy of CC Dx was 3,738 for the sample and 3,468,900 for the estimated population.
3. The denominator for % discrepancy of all Dx was 5,884 for the sample and 6,033,900 for the estimated population.
4. The denominator for % discrepancy of E-codes was 886 for the sample and 525,200 for the estimated population. E-code discrepancies are based on year 2 results only.

For year 1 sample and estimated counts of discrepancies please refer to Table 4A in Appendix G.
For year 2 sample and estimated counts of discrepancies please refer to Table 4A in Appendix H.

Table 4B—Diagnosis Discrepancies by Reason Code

Combined Year 1 and 2

Reason Code	Discrepancies (%)						
	MRDx[1]	CC Dx Typing[2]	CC Dx[3]	Dx Prefix/Suffix Different[4]	Any Dx Different[5]	E-Code[6]	Total[7]
I - No significant impact on treatment	6.8	62.1	24.2	0.0	0.0	0.0	25.8
F - Different interpretation - disagree	39.1	8.0	15.4	72.5	20.7	12.0	18.5
P - Information on chart missed	10.7	1.8	30.4	2.5	14.5	28.1	17.0
O - Coding error	18.3	12.9	6.9	4.8	3.0	23.6	10.0
K - Other grey area coding	8.5	11.1	7.1	3.5	3.3	0.4	7.5
M - Coding contrary to CIHI guidelines	3.6	0.6	3.3	3.2	25.9	18.6	7.1
E - Code specificity not supported	0.7	0.1	5.6	9.5	9.0	1.6	4.0
V - Other	0.2	0.0	0.0	0.0	16.3	0.6	2.9
D - Lack of Code specifity	3.3	1.0	3.1	0.7	1.4	14.9	2.6
W - No apparent reason	4.8	0.1	1.6	1.0	5.3	0.0	2.4
L - Inconsistent or conflicting information	1.4	2.0	1.2	0.0	0.0	0.0	1.2
Z - Significant impact on treatment[8]	2.5	0.2	0.8	2.4	0.0	0.2	0.8
N - Hospital Policy	0.1	0.1	0.4	0.0	0.0	0.0	0.2
A - Transcription error	0.0	0.0	0.0	0.0	0.6	0.0	0.1
B - Incomplete documentation	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Total	**100.0**	**100.0**	**100.0**	**100.0**	**100.0**	**100.0**	**100.0**

Table Notes:

1. The estimated population (N) of reasons for MRDx discrepancies was 323,600.
2. The estimated population (N) of reasons for CC Dx typing discrepancies was 537,400.
3. The estimated population (N) of reasons for CC Dx discrepancies was 804,000.
4. The estimated population (N) of reasons for Dx prefix/suffix different discrepancies was 9,700.
5. The estimated population (N) of reasons for any Dx different discrepancies was 368,000.
6. The estimated population (N) of reasons for E-code discrepancies was 53,600.
   E-code discrepancies based on year 2 data only.
7. The estimated population (N) of reasons for total diagnosis discrepancies was 2,100,400.
8. Reason code Z is based on Year 2 data only.

Table 5A—Procedure Discrepancies

Combined Year 1 and 2

Procedure Discrepancies	Discrepancies (%)	
	Sample	Population
Principal Procedure Discrepancies[1]		
22 - Principal proc as other proc	0.8	0.2
23 - Principal proc missing	3.0	3.9
25 - Proc not coded, original as PP	1.1	2.9
Total Principal Procedure Discrepancies	**4.8**	**7.0**
Other Procedure Discrepancies[2]		
24 - Other procedure missing	10.6	10.2
26 - Proc not coded, original as other	7.2	9.6
Total Other Procedure Discrepancies	**17.8**	**19.7**
21 - Procedure code different[3]	4.7	4.9
30 - Procedure date is different[3]	0.1	0.3
Total Procedure Discrepancies[3]	**16.0**	**18.3**

Table Notes:

1. The denominator for % discrepancy was 1,396 for the sample and 1,058,100 for the estimated population.
2. The denominator for % discrepancy was 1,347 for the sample and 958,300 for the estimated population.
3. The denominator for % discrepancy was 2,743 for the sample and 2,016,400 for the estimated population.

For year 1 sample and estimated counts of discrepancies please refer to Table 5A in Appendix G.
For year 2 sample and estimated counts of discrepancies please refer to Table 5A in Appendix H.

Table 5B—Procedure Discrepancies by Reason Code

Combined Year 1 and 2

Reason Code	Discrepancies (%)				
	Principal Procedure[1]	Other Procedure[2]	Any Proc. Different[3]	Any Proc. Date Different[4]	Total Proc.[5]
P - Information on chart missed	59.6	37.7	18.0	53.9	36.5
F - Different interpretation - disagree	29.7	19.8	39.8	0.0	27.5
O - Coding error	0.4	8.1	21.2	0.0	10.4
V - Other	1.0	4.4	0.3	0.0	2.3
K - Other grey area coding	1.1	8.1	2.0	0.0	4.6
M - Coding contrary to CIHI guidelines	3.0	9.4	0.2	0.0	5.0
W - No apparent reason	4.7	1.5	0.0	0.0	1.7
N - Hospital policy	0.0	2.1	6.6	0.0	3.0
E - Code specificity not supported	0.2	2.7	8.9	0.0	4.0
L - Inconsistent or conflicting information	0.3	4.6	0.7	46.0	3.6
D - Lack of Code specificity	0.0	0.2	1.3	0.0	0.5
A - Transcription error	0.0	0.0	0.9	0.1	0.3
B - Incomplete documentation	0.0	0.0	0.0	0.0	0.0
I - No significant impact on treatment	0.0	1.4	0.0	0.0	0.6
Total	**100.0**	**100.0**	**100.0**	**100.0**	**100.0**

Table Notes:

1. The estimated population (N) of reasons for principal procedure discrepancies was 73,900.
2. The estimated population (N) of reasons for other procedure discrepancies was 189,300.
3. The estimated population (N) of reasons for any procedure different discrepancies was 99,200.
4. The estimated population (N) of reasons for any procedure date different discrepancies was 11,900. Procedure data different discrepancies based on year 2 data only.
5. The estimated population (N) of reasons for total procedure discrepancies was 368,300.

Appendix G

Detailed Tables—Year 1

Table 1A—Facility Response Rates

Year 1

Units	Count	Response Rate
# hospitals contacted	26	
# hospitals no response	2	
# hospitals declined	2	
# hospitals accepted	22	85%
# hospitals participating	18	

Table 1B—Summary Counts of Re-abstracted Charts

Year 1

Indicators	Target Sample Size	Initial Sample Size[1]	Reabstracted Charts	% Response Rate[2]	% of Target[3]
Health Indicator					
Ambulatory care sens. cond.	250	296	272	91.9	108.8
Cesarean section	251	281	264	94.0	105.2
Coronary artery bypass graft	226	274	176	64.2	77.9
Hospitalization pneumonia	249	283	261	92.2	104.8
Injury hospitalization	251	290	269	92.8	107.2
Total hip replacement	242	279	224	80.3	92.6
Vaginal births after cesarean	236	266	209	78.6	88.6
CPSS Indicator					
Rare Congenital Anomalies	75	77	73	94.8	97.3
Rare Maternal Conditions	71	80	69	86.3	97.2
Rare Neonatal Conditions	73	93	77	82.8	105.5
Respiratory distress syn.	99	119	78	65.5	78.8
Third degree perineal lac.	207	250	208	83.2	100.5
Other Maternal Neonatal Cond.	225	350	296	84.6	131.6
Other	245	302	261	86.4	106.5
Total Year 1	**2,700**	**3,240**	**2,737**	**84.5**	**101.4**

Table Notes:

1. The initial sample is the target sample increased by 10% for non-response and a further 10% for possible situations of better than expected productivity by the re-abstractor.
2. The response rate is calculated using the initial sample size.
3. The response rate is calculated using the target sample size.

Table 2—Estimated False Positives[19] for Health and CPSS Indicators

Year 1

Indicator	Sample Count Re-abstracted	False Positives[1]					
		Sample Count	Est. Count[2]	Sample %	Est. %	Lower Bound of 95% Confidence Interval	Upper Bound of 95% Confidence Interval
Health Indicator							
Ambulatory care sens. cond.	272	31	10,200	11.4	10.7	8.7	12.7
Cesarean section	308	1	100	0.4	0.1	0.0	0.2
Coronary artery bypass graft	176	1	200	0.6	0.8	0.1	1.5
Hospitalization pneumonia	261	23	4,800	8.8	6.9	5.0	8.8
Injury hospitalization	302	18	8,300	6.7	5.3	3.0	7.6
Total hip replacement	224	2	100	0.9	0.8	0.4	1.2
Vaginal births after cesarean	233	3	100	1.4	0.7	0.1	1.2
CPSS Indicator							
Third degree perineal lac.	213	12	2,500	5.8	23.0	22.0	24.1
Respiratory distress syn.	84	21	400	26.9	22.0	13.7	30.4
Rare Neonatal Conditions	81	11	300	13.6	15.8	9.1	22.6
Rare Maternal Conditions	69	11	200	15.9	23.4	20.6	26.2
Rare Congenital Anomalies	73	11	100	15.1	12.5	6.8	18.3
Other Maternal Neonatal Cond.	1,264	1	0	0.3	0.0	0.0	0.1
Other	261	10	24,400	3.8	1.5	1.1	1.8

Table Notes:

1. Estimated false positives are calculated using both an estimated numerator and estimated denominator.
2. Estimated counts have been rounded to units of 100, so as not to imply an unintended level of precision.

[19] Discrepancies for the diagnosis and procedures related to Health Indicators are false positives. These occur when the diagnosis and procedure codes in the original abstract met the criteria for inclusion but the criteria were not met by the re-abstracted codes, as listed in Appendix B, Health and CPSS Indicators. See the example contained in Table 1: Example, in the report section, **Identification of Discrepancies and Reasons**. Further note that while these conditions are used in the calculation of many indicators, the discrepancy rates should not be interpreted as discrepancy rates for the health indicators, since there are additional criteria that are applied in the calculation of the health indicator.

Table 3A—Estimated Demographic and Non-Medical Discrepancies by Data Element

Year 1

Data Element[1,2]	Non-Medical Discrepancies (%)			
	Entry Missing	Entry Not Reabstracted	Entry Different	Total
Admission Category	0.1	0.0	13.8	13.9
Discharge Hour	0.0	0.0	9.8	9.8
Postal Code	0.0	0.0	9.0	9.0
Entry Code	0.0	0.0	6.5	6.5
Institution To	0.4	1.0	0.3	1.7
Institution From	0.7	0.3	0.1	1.1
Admit Hour	0.0	0.0	0.5	0.5
Supplemental Death Code	0.0	0.0	0.2	0.2
Health Care Number	0.0	0.0	0.2	0.2
Exit Alive	0.0	0.0	0.1	0.1
Birth Date	0.0	0.0	0.1	0.1
Provincial Issuing HCN	0.0	0.0	0.0	0.0
Admit Date	0.0	0.0	0.0	0.0
Discharge Date	0.0	0.0	0.0	0.0

Table Notes:

1. The population (N) is 2,391,440.
2. Data elements that are mandatory for the provinces in scope to the study.

Table 3B—Estimated Demographic and Non-Medical Discrepancies by Reason Code

Year 1

Reason Code Description	Non-Medical Discrepancies (%)			
	Entry Missing[1]	Entry Not Reabstracted[2]	Entry Different[3]	Total[4]
R - Downloaded incorrectly	0.0	1.2	32.5	31.1
P - Information on chart missed	74.3	0.0	24.1	24.6
M - Coding contrary to CIHI guidelines	9.6	0.0	13.3	12.9
N - Hospital policy	0.0	20.5	12.6	12.5
W - No apparent reason	14.0	2.3	8.3	8.3
F - Different interpretation - disagree	0.0	25.7	4.6	5.0
L - Inconsistent or conflicting information	0.0	46.0	1.6	2.6
V - Other	2.1	4.0	1.7	1.7
A - Transcription error	0.0	0.0	0.8	0.8
K - Other grey area coding	0.0	0.0	0.3	0.2
E - Code specificity not supported	0.0	0.0	0.2	0.1
Q - Mathematical/counting error	0.0	0.0	0.0	0.0
B - Incomplete documentation	0.0	0.3	0.0	0.0
O - Coding error	0.0	0.0	0.0	0.0
Total Reasons	**100.0**	**100.0**	**100.0**	**100.0**

Table Notes:

1. The estimated population (N) of reasons for entry missing discrepancies was 30,400.
2. The estimated population (N) of reasons for entry not re-abstracted discrepancies was 33,900.
3. The estimated population (N) of reasons for entry different discrepancies was 1,422,400.
4. The estimated population (N) of reasons for total non-medical discrepancies was 1,486,700.

Table 4A—Diagnosis Discrepancies

Year 1

Diagnosis Discrepancies	Discrepancies	Est. # of Discr.[5]	Discrepancies (%)	
			Sample	Estimated Population
MRDx Discrepancies[1]				
6 - MRDx different type	110	157,900	4.0	6.3
7 - MRDx missing	107	79,600	3.9	3.2
11 - Post-admit as MRDx	1	100	0.0	0.0
13 - Secondary as MRDx	5	1,700	0.2	0.1
15 - Dx not coded, orig MRDx	109	97,600	4.0	3.9
Total MRDx Discrepancies	**332**	**336,700**	**12.1**	**13.4**
CC Diagnosis Typing Discrepancies[2]				
8 - CC Dx as type 3	123	62,100	2.6	2.1
10 - Pre-admit as post-admit	17	5,700	0.4	0.2
12 - Post-admit as pre-admit	26	17,700	0.6	0.6
14 - Secondary Dx as CC	195	244,100	4.2	8.1
Total CC Diagnosis Typing Discrepancies	**361**	**329,600**	**7.8**	**11.0**
CC Diagnosis Discrepancies[2]				
9 - CC Dx missing	524	432,600	11.3	14.4
16 - Dx not coded, orig CC	674	497,600	14.5	16.6
18 - Transfer Dx missing	10	5,800	0.2	0.2
19 - Dx not coded orig trans	3	1,300	0.1	0.0
Total CC Diagnosis Discrepancies	**1,211**	**937,300**	**26.0**	**31.2**
4 - Diagnosis Prefix/Suffix Different Discrepancies[3]	27	6,100	0.4	0.1
5 - Any Diagnosis Different Discrepancies[3]	550	355,900	7.4	6.5
20 - E-code Discrepancies[4]				
Total Diagnosis Discrepancies	**2,481**	**1,965,600**	**33.6**	**35.7**

Table Notes:

1. The denominator for % discrepancy of MRDx was 2,737 for the sample and 2,504,600 for the estimated population.
2. The denominator for % discrepancy of CC Dx was 4,657 for the sample and 2,999,500 for the estimated population.
3. The denominator for % discrepancy of all Dx was 7,394 for the sample and 5,504,100 for the estimated population.
4. The E-codes were not separated from other Discrepancies in year 1.
5. These estimated counts have been rounded to units of 100, so as not to imply an unintended level of precision. Note that these estimated counts may not add to the total due to rounding.

Table 4B—Diagnosis Discrepancies by Reason Code

Year 1

Reason Code	Discrepancies (%)					
	MRDx[1]	CC Dx Typing[2]	CC Dx[3]	Dx Prefix/Suffix Different[4]	Any Dx Different[5]	Total[6]
P - Information on chart missed	15.2	5.6	39.4	7.9	6.9	23.6
I - No significant impact on treatment	4.8	58.2	22.1	0.0	0.0	20.8
F - Different interpretation - disagree	40.1	12.6	9.1	20.6	19.9	17.2
E - Code specificity not supported	1.5	0.3	9.7	28.8	16.0	7.9
M - Coding contrary to CIHI guidelines	15.9	5.4	7.2	12.1	1.4	7.4
O - Coding error	3.6	0.9	3.2	6.4	28.1	7.4
K - Other grey area coding	8.5	12.7	6.3	11.1	3.2	7.2
L - Inconsistent or conflicting information	8.0	0.1	0.7	3.2	6.0	2.9
D - Lack of Code specifity	0.3	0.0	0.0	0.0	14.0	2.6
V- Other	1.6	3.2	1.3	2.4	3.1	2.0
N - Hospital Policy	0.3	0.3	0.6	0.0	0.0	0.4
W- No apparent reason	0.1	0.5	0.3	7.7	0.0	0.3
A - Transcription error	0.0	0.1	0.0	0.0	1.2	0.2
B - Incomplete documentation	0.0	0.1	0.0	0.0	0.0	0.0
Total	**100.0**	**100.0**	**100.0**	**100.0**	**100.0**	**100.0**

Table Notes:

1. The estimated population (N) of reasons for MRDx discrepancies was 336,700.
2. The estimated population (N) of reasons for CC Dx typing discrepancies was 329,600.
3. The estimated population (N) of reasons for CC Dx discrepancies was 937,300.
4. The estimated population (N) of reasons for Dx prefix/suffix different discrepancies was 6,100.
5. The estimated population (N) of reasons for any Dx different discrepancies was 355,900.
6. The estimated population (N) of reasons for total diagnosis discrepancies was 1,965,600.

Table 5A—Procedure Discrepancies

Year 1

Procedure Discrepancies	Discrepancies	Est. # of Discr.[4]	Discrepancies (%)	
			Sample	Population
Principal Procedure Discrepancies[1]				
22 - Principal proc as other proc	20	3,300	1.2	0.3
23 - Principal proc missing	55	56,900	3.4	4.9
25 - Proc not coded, original as PP	26	56,600	1.6	4.8
Total Principal Procedure Discrepancies	**101**	**116,800**	**6.3**	**10.0**
Other Procedure Discrepancies[2]				
24 - Other procedure missing	234	120,400	12.3	10.6
26 - Proc not coded, original as other	139	143,400	7.3	12.7
Total Other Procedure Discrepancies	**373**	**263,900**	**19.6**	**23.3**
21 - Procedure code different[3]	138	123,000	3.9	5.3
30 - Procedure date is different[3]			0.0	0.0
Total Procedure Discrepancies[3]	**612**	**503,600**	**17.4**	**21.9**

Table Notes:

1. The denominator for % discrepancy was 1,615 for the sample and 1,170,000 for the estimated population.
2. The denominator for % discrepancy was 1,903 for the sample and 1,130,900 for the estimated population.
3. The denominator for % discrepancy was 3,518 for the sample and 2,300,900 for the estimated population.
4. These estimated counts have been rounded to units of 100, so as not to imply an unintended level of precision. Note that these estimated counts may not add to the total due to rounding.

Table 5B—Procedure Discrepancies by Reason Code

Year 1

Reason Code	Discrepancies (%)			
	Principal Procedure[1]	Other Procedure[2]	Any Proc. Different[3]	Total Proc.[4]
P - Information on chart missed	56.0	34.1	19.1	34.7
F - Different interpretation - disagree	35.6	16.5	26.8	24.1
V - Other	1.2	5.9	0.5	3.1
K - Other grey area coding	1.3	11.0	2.9	6.3
O - Coding error	0.5	6.9	28.9	12.1
W - No apparent reason	2.6	0.5	0.0	0.9
N - Hospital policy	0.1	2.8	9.3	4.1
E - Code specificity not supported	0.3	3.7	8.7	4.4
M - Coding contrary to CIHI guidelines	2.1	10.1	0.3	5.2
D - Lack of Code specificity	0.0	0.3	1.8	0.7
L - Inconsistent or conflicting information	0.3	6.4	0.5	3.2
A - Transcription error	0.0	0.0	1.2	0.4
I - No significant impact on treatment	0.0	1.9	0.0	0.9
B - Incomplete documentation	0.0	0.0	0.0	0.0
Total	**100.0**	**100.0**	**100.0**	**100.0**

Table Notes:

1. The estimated population (N) of reasons for principal procedure discrepancies was 116,800.
2. The estimated population (N) of reasons for other procedure discrepancies was 263,900.
3. The estimated population (N) of reasons for any procedure different discrepancies was 123,000.
4. The estimated population (N) of reasons for total procedure discrepancies was 129,300.

Appendix H

Detailed Tables—Year 2

Table 1A—Facility Response Rates

Year 2

Units	Count	Response Rate
# hospitals contacted	17	
# hospitals no response	1	
# hospitals declined	2	
# hospitals accepted	14	82%
# hospitals participating	11	

Table 1B—Summary Counts of Re-abstracted Charts

Year 2

Indicators	Target Sample Size	Initial Sample Size[1]	Reabstracted Charts	% Response Rate[2]	% of Target[3]
Acute Myocardial Infarction	275	340	307	90.3	111.6
Hysterectomy	275	330	316	95.8	114.9
Hip fracture	275	330	310	93.9	112.7
Total knee replacement	275	330	308	93.3	112.0
Not assigned to any year 2 indicator	275	330	314	95.2	114.2
Total Year 2	**1,375**	**1,660**	**1,555**	**93.7**	**113.1**

Table Notes:

1. The initial sample is the target sample increased by 10% for non-response and a further 10% for possible situations of better than expected productivity by the re-abstractor.
2. The response rate is calculated using the initial sample size.
3. The response rate is calculated using the target sample size.

Table 2—Estimated False Positives[20] for Health Indicators

Year 2

Indicator	Sample Count Re-abstracted	False Positives[1]					
		Sample Count	Est. Count[2]	Sample %	Est. %	Lower Bound of 95% Confidence Interval	Upper Bound of 95% Confidence Interval
Acute Myocardial Infarction	319	20	8,800	6.3	8.9	7.4	10.4
Hip Fracture	312	11	1,400	3.5	6.0	5.1	6.8
Hysterectomy	316	2	100	0.6	0.3	0.1	0.5
Total Knee Replacement	308	3	100	1.0	0.8	0.4	1.2
Other	314	2	10,300	0.6	0.4	0.2	0.7

Table Notes:

1. Estimated false positives are calculated using both an estimated numerator and estimated denominator.
2. Estimated counts have been rounded to units of 100, so as not to imply an unintended level of precision.

[20] Discrepancies for the diagnosis and procedures related to Health Indicators are false positives. These occur when the diagnosis and procedure codes in the original abstract met the criteria for inclusion but the criteria were not met by the re-abstracted codes, as listed in Appendix B, Health and CPSS Indicators. See the example contained in Table 1: Example, in the report section, **Identification of Discrepancies and Reasons**.

Table 3A—Estimated Demographic and Non-Medical Discrepancies by Data Element

Year 2

Data Element[1,2]	Non-Medical Discrepancies (%)			
	Entry Missing	Entry Not Reabstracted	Entry Different	Total
Admission Category	0.0	0.0	17.3	17.3
Discharge Hour	0.0	1.0	9.3	10.3
Entry Code	0.0	0.0	3.6	3.6
Institution From	2.6	0.9	0.1	3.5
Admit Hour	0.0	1.6	1.3	2.9
Institution To	1.5	0.7	0.6	2.8
Postal Code	0.0	0.0	0.7	0.7
Health Care Number	0.0	0.0	0.3	0.3
Admit Date	0.0	0.0	0.3	0.3
Provincial Issuing HCN	0.0	0.0	0.0	0.0
Discharge Date	0.0	0.0	0.0	0.0
Birth Date	0.0	0.0	0.0	0.0
Supplemental Death Code	0.0	0.0	0.0	0.0
Exit Alive	0.0	0.0	0.0	0.0

Table Notes:

1. Data elements that are mandatory for the provinces in scope to the study.
2. The population (N) is 2,209,596

Table 3B—Estimated Demographic and Non-Medical Discrepancies by Reason Code

Year 2

Reason Code Description	Non-Medical Discrepancies (%)			
	Entry Missing[1]	Entry Not Reabstracted[2]	Entry Different[3]	Total[4]
M - Coding contrary to CIHI guidelines	0.0	26.4	48.0	41.1
P - Information on chart missed	71.9	0.0	35.4	35.4
L - Inconsistent or conflicting information	0.0	72.2	5.2	11.5
N - Hospital policy	28.1	0.0	3.9	5.9
W - No apparent reason	0.0	0.2	3.1	2.5
F - Different interpretation - disagree	0.0	0.5	2.2	1.8
R - Downloaded incorrectly	0.0	0.0	1.9	1.6
V - Other	0.0	0.6	0.1	0.1
K - Other grey area coding	0.0	0.1	0.0	0.0
A - Transcription error	0.0	0.0	0.0	0.0
D - Lack of code specificity	0.0	0.0	0.0	0.0
Total Reasons	**100.0**	**100.0**	**100.0**	**100.0**

Table Notes:

1. The estimated population (N) of reasons for entry missing discrepancies was 90,800.
2. The estimated population (N) of reasons for entry not re-abstracted discrepancies was 94,000.
3. The estimated population (N) of reasons for entry different discrepancies was 739,800.
4. The estimated population (N) of reasons for total non-medical discrepancies was 924,700.

Table 4A—Diagnosis Discrepancies

Year 2

Diagnosis Discrepancies	Discrepancies	Est. # of Discr.[5]	Discrepancies (%)	
			Sample	Estimated Population
MRDx Discrepancies[1]				
6 - MRDx different type	54	137,400	3.5	5.2
7 - MRDx missing	30	68,100	1.9	2.6
11 - Post-admit as MRDx	2	100	0.1	0.0
13 - Secondary as MRDx	6	7,500	0.4	0.3
15 - Dx not coded, orig MRDx	33	105,500	2.1	4.0
Total MRDx Discrepancies	**125**	**318,500**	**8.0**	**12.1**
CC Diagnosis Typing Discrepancies[2]				
8 - CC Dx as type 3	45	51,500	1.6	1.3
10 - Pre-admit as post-admit	25	18,100	0.9	0.5
12 - Post-admit as pre-admit	14	2,200	0.5	0.1
14 - Secondary Dx as CC	551	673,400	19.5	17.1
Total CC Diagnosis Typing Discrepancies	**635**	**745,200**	**22.5**	**18.9**
CC Diagnosis Discrepancies[2]				
9 - CC Dx missing	268	187,700	9.5	4.8
16 - Dx not coded, orig CC	412	472,900	14.6	12.0
18 - Transfer Dx missing	2	100	0.1	0.0
19 - Dx not coded orig trans	6	9,900	0.2	0.3
Total CC Diagnosis Discrepancies	**688**	**670,600**	**24.4**	**17.0**
4 - Diagnosis Prefix/Suffix Different Discrepancies[3]	8	13,400	0.2	0.2
5 - Any Diagnosis Different Discrepancies[3]	300	380,100	6.9	5.8
20 - E-code Discrepancies[4]	98	53,600	11.1	10.2
Total Diagnosis Discrepancies	**1,854**	**2,181,500**	**42.4**	**33.2**

Table Notes:

1. The denominator for % discrepancy of MRDx was 1,555 for the sample and 2,625,400 for the estimated population.
2. The denominator for % discrepancy of CC Dx was 2,819 for the sample and 3,938,300 for the estimated population.
3. The denominator for % discrepancy of all Dx was 4,374 for the sample and 6,563,700 for the estimated population.
4. The denominator for % discrepancy of E-codes was 886 for the sample and 525,200 for the estimated population.
5. These estimated counts have been rounded to units of 100, so as not to imply an unintended level of precision. Note that these estimated counts may not add to the total due to rounding.

Table 4B—Diagnosis Discrepancies by Reason Code

Year 2

Reason Code	Discrepancies (%)						
	MRDx[1]	CC Dx Typing[2]	CC Dx[3]	Dx Prefix/Suffix Different[4]	Any Dx Different[5]	E-Code[6]	Total[7]
I - No significant impact on treatment	8.7	63.8	27.0	0.0	0.0	0.0	30.8
F - Different interpretation - disagree	38.0	5.9	24.0	96.2	21.3	12.0	19.8
M - Coding contrary to CIHI guidelines	20.9	16.1	6.5	1.4	4.4	23.6	11.9
P - Information on chart missed	6.0	0.1	18.0	0.0	21.1	28.1	10.9
K - Other grey area coding	8.4	10.4	8.2	0.0	3.3	0.4	7.9
O - Coding error	3.6	0.5	3.5	1.8	23.9	18.6	6.5
D - Lack of Code specifity	0.0	0.0	0.0	0.0	18.3	0.6	3.3
V - Other	5.0	0.0	5.5	0.0	0.0	14.9	2.8
Z - Significant impact on treatment	2.8	2.9	3.0	0.0	0.0	0.0	2.3
L - Inconsistent or conflicting information	1.6	0.1	2.7	0.0	4.7	0.0	1.9
W - No apparent reason	5.0	0.1	1.5	0.0	0.0	0.2	1.3
E - Code specificity not supported	0.0	0.0	0.0	0.7	2.9	1.6	0.6
B - Incomplete documentation	0.1	0.0	0.0	0.0	0.0	0.0	0.0
A - Transcription error	0.0	0.0	0.0	0.0	0.1	0.0	0.0
N - Hospital Policy	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Total	**100.0**	**100.0**	**100.0**	**100.0**	**100.0**	**100.0**	**100.0**

Table notes:

1. The estimated population (N) of reasons for MRDx discrepancies was 318,500.
2. The estimated population (N) of reasons for CC Dx typing discrepancies was 745,200.
3. The estimated population (N) of reasons for CC Dx discrepancies was 670,600.
4. The estimated population (N) of reasons for Dx prefix/suffix different discrepancies was 13,400.
5. The estimated population (N) of reasons for any Dx different discrepancies was 380,100.
6. The estimated population (N) of reasons for E-code discrepancies was 53,600.
   E-code discrepancies based on year 2 data only.
7. The estimated population (N) of reasons for total diagnosis discrepancies was 2,181,500.

Table 5A—Procedure Discrepancies

Year 2

Procedure Discrepancies	Discrepancies	Est. # of Discr.[4]	Discrepancies (%)	
			Sample	Population
Principal Procedure Discrepancies[1]				
22 - Principal proc as other proc	1	20	0.1	0.0
23 - Principal proc missing	28	26,100	2.4	2.8
25 - Proc not coded, original as PP	4	5,000	0.3	0.5
Total Principal Procedure Discrepancies	**33**	**31,100**	**2.8**	**3.3**
Other Procedure Discrepancies[2]				
24 - Other procedure missing	52	74,900	6.6	9.5
26 - Proc not coded, original as other	54	39,700	6.8	5.1
Total Other Procedure Discrepancies	**106**	**114,700**	**13.4**	**14.6**
21 - Procedure code different[3]	120	75,400	6.1	4.4
30 - Procedure date is different[3]	6	11,900	0.3	0.7
Total Procedure Discrepancies[3]	**265**	**233,100**	**13.5**	**13.5**

Table Notes:

1. The denominator for % discrepancy was 1,177 for the sample and 946,300 for the estimated population.
2. The denominator for % discrepancy was 792 for the sample and 785,700 for the estimated population.
3. The denominator for % discrepancy was 1,969 for the sample and 1,732,000 for the estimated population.
4. These estimated counts have been rounded to units of 100, so as not to imply an unintended level of precision. Note that these estimated counts may not add to the total due to rounding.

Table 5B—Procedure Discrepancies by Reason Code

Year 2

Reason Code	Discrepancies (%)				
	Principal Procedure[1]	Other Procedure[2]	Any Proc. Different[3]	Any Proc. Date Different[4]	Total Proc.[5]
P - Information on chart missed	77.0	47.1	15.3	53.9	40.3
F - Different interpretation - disagree	0.3	28.4	70.8	0.0	38.0
O - Coding error	0.0	11.2	2.9	0.0	6.3
M - Coding contrary to CIHI guidelines	7.1	7.8	0.0	0.0	4.6
L - Inconsistent or conflicting	0.6	0.0	1.2	46.0	2.8
E - Code specificity not supported	0.0	0.0	9.4	0.0	3.2
D - Lack of Code specificity	0.0	0.0	0.1	0.0	0.0
K - Other grey area coding	0.0	0.6	0.0	0.0	0.3
A - Transcription error	0.0	0.0	0.2	0.1	0.1
W - No apparent reason	15.0	4.2	0.0	0.0	3.9
B - Incomplete documentation	0.0	0.1	0.0	0.0	0.0
N - Hospital policy	0.0	0.2	0.0	0.0	0.1
V - Other	0.0	0.4	0.0	0.0	0.2
Total	**100.0**	**100.0**	**100.0**	**100.0**	**100.0**

Table Notes:

1. The estimated population (N) of reasons for principal procedure discrepancies was 31,000.
2. The estimated population (N) of reasons for other procedure discrepancies was 114,700.
3. The estimated population (N) of reasons for any procedure different discrepancies was 75,400.
4. The estimated population (N) of reasons for any procedure date different discrepancies was 11,900. Procedure data different discrepancies based on year 2 data only.
5. The estimated population (N) of reasons for total procedure discrepancies was 233,100.