Information Management

A System We Can Count On

Reabstraction Study of the Ontario Case Costing Facilities

for Fiscal Years 2002/2003 and 2003/2004

Health Results Team for Information Management

November 2005

CHIMA Healthcare Professionals for the Information Age
Canadian Health Information Management Association
Canadian College of Health Record Administrators

CIHI ICIS Canadian Institute for Health Information
Institut canadien d'information sur la santé

Ontario

Acknowledgements

The findings of this report are a result of the significant effort of several members from various organizations: the Ministry of Health and Long-Term Care, the Canadian Institute for Health Information, CHIM Information Consulting Inc., IT/NET, and Salumatics. In addition to these organizations, we sincerely acknowledge the hospital staff that accommodated this large study by pulling charts and providing workspace for the reabstraction specialists.

We would also like to thank the many health information professionals who contributed tremendously to the outcome of the study. Without their hard work and dedication, this study would not have been possible.

Health Information Professionals
Marikay Bailey, Usha Chidambaram, Debbie Cullen, Shebnum Devji, Lisa Dillman, Yvonne Duhaney, Lori Goodwin, Lisa Gray, Susan Hamilton, Karen Higgins, JoAnne Lokun, Angeline Luteyn, Jodi McClarty, Stella Mesina, Stephanie Olsen, Yvonne Peekhaus, Cyndee Pettifer, Laurel Pickett, Debra Ann Tetreault, Kim Torvinen, Catherine Wade, Linda Wilson, Leah Woida, Deborah Young

Table of Contents

List of Tables and Figures

Executive Summary

The Ontario Ministry of Health and Long-Term Care (MOHLTC), the Canadian Institute for Health information (CIHI) and CHIM Information Consulting Inc. (CHIM) jointly conducted a reabstraction study to review the clinical coding practices of Ontario's ten case-costing hospital corporations.

This reabstraction study is the largest ever conducted in Canada. Approximately 14,500 discharges were reabstracted from the 18 sites within Ontario's ten case-costing hospital corporations for the 2002–2003 and 2003–2004 fiscal years. This study also includes the double reabstraction of approximately 800 discharges to specifically measure the inter-rater reliability of the health information professionals who participated in the study.

Ontario's case-costing hospital corporations use a standardized methodology to collect patient-specific cost data. These data include all direct and indirect costs incurred by the hospital in the provision of patient care. The patient-specific cost records are combined with the clinical records submitted to CIHI by hospitals. The resultant data set is used for many important initiatives, including CIHI's development of a new acute inpatient grouping methodology and of the associated Resource Intensity Weight (or RIW™) using International Statistical Classification of Diseases and Related Health Problems, 10th Revision, Canada (ICD-10-CA) and Canadian Classification of Health Interventions (CCI). The quality of clinical data from these hospitals is therefore of particular importance.

Encouraging findings from the study showed the following:[1]

- Non-medical data (for example, birth date, gender and admission date) continue to be accurately coded. Agreement rates for non-medical elements range from 97 to 100%.
- Interventions and their associated attributes were accurately coded. Agreement rates of 86% in FY 2002–2003 and 91% in FY 2003–2004 were observed.
- Some improvement in coding quality between fiscal years was noted for intervention coding, assignment of significance, assignment of complexity (where the complexity overlay is applied), and the net change in expected length of stay (ELOS). Although the upward trend for the noted variables is encouraging, the overall level of coding consistency is still low.

The study also identified six priority areas for data quality improvement initiatives.

(1) *Diagnosis significance.* Identifying diagnoses that had a significant impact on patients' length of stay or the resources required for their care was an

[1] The figures in the Executive Summary for the main data set have been adjusted for coder effect for all data elements with the exceptions of non-medical, ELOS, and RIW. The figures for the inter-rater data set include adjustments for case-mix effect.

™ Registered trade-mark of the Canadian Institute for Health Information.

issue for both the original coders and the reabstractors. When considering all conditions that were identified as significant by the original coder and/or reabstractor, only 62% of the conditions were deemed significant by both in FY 2002–2003 and 67% in FY 2003–2004.

(2) *Chart documentation.* Chart documentation was the single-largest contributor to the noted discrepancies in both the main study and inter-rater results. There were many circumstances that would lead the reabstractor to indicate that a noted discrepancy was the result of chart documentation. The following five reasons were most prevalent:

- Different interpretation of the chart by the reabstractor and the hospital coder;
- Documentation in the chart was incomplete at the time of original abstraction;
- Conflicting documentation in the chart led the reabstractor and hospital coder to select different diagnosis codes;
- Information on the chart was missed by the hospital coder; and
- The reabstractor determined that the specificity of the diagnosis code assigned by the hospital coder was not supported in the chart.

(3) *Selection of the Most Responsible Diagnosis.* Low agreement rates on the codes selected as the most responsible diagnosis (MRDx) were observed. Agreement rates were approximately 75% in each fiscal year. These discrepancies affected the assignment of the Major Clinical Category (MCC) for about 5% of the discharges.

(4) *Inter-rater reliability.* The findings from the inter-rater data (that is, comparison between reabstractors) were generally similar to the findings from the main study data (that is, comparison between original coder and reabstractor). This was a particularly surprising result given the rigour involved in the selection process for the study reabstractors, the intensive one-week training and the support and the resources available to the reabstractors during data collection.

The few exceptions to this were the following:

- The selection of code for the MRDx had a higher agreement rate in the inter-rater data set than in the main data set by an estimated 12% in FY 2002–2003 and 18% in FY 2003–2004;
- The assignment of significance to a condition had a higher agreement rate in the inter-rater data set by an estimated 7% in FY 2003–2004;
- Improvement between fiscal years was significant in the inter-rater data set but not in the main data set for selection of diagnosis code; and
- There was improvement in the selection of intervention codes between fiscal years in the main data set but not in the inter-rater data set.

Any bias arising from the variability in agreement rates between reabstractors was adjusted throughout the report and is referred to as the "coder effect."

(5) *Variation in coding practice across facilities.* Case-costing facilities generally do not have similar coding practices. Extensive analysis revealed that, even after adjusting for the different types of conditions present in each facility's patient mix (noted as case-mix effect), substantive differences were observed across facilities.

(6) *Impact of observed variation on case-mix grouping.* The impact of the observed discrepancies in the coding of diagnoses, interventions and assignment of significance to diagnoses affected the output variables from CIHI's grouping methodology in the following ways:

- About 14% of the discharges in FY 2002–2003 and 16% in FY 2003–2004 changed assignment of Case Mix Group (or CMG™).
- Discrepancies associated with diagnosis typing resulted in 13% of the discharges in FY 2002–2003 and 10% in FY 2003–2004 changing complexity (or Plx™) level assignment. This was increased to 21% and 15% for the respective study years when the cases with embedded complexity were removed (for example, obstetrics, neonates and mental health). High agreement rates were observed for Plx 1 cases (93% in FY 2003–2004). Lower agreement rates were observed for the other Plx levels. In FY 2003–2004, the agreement rates for Plx 2, Plx 3 and Plx 4 were 50%, 42% and 59%, respectively.
- The net change in total ELOS was a decrease of 7.3% in FY 2002–2003 and of 4.5% in FY 2003–2004.
- The net change in total RIW value was a decrease of 4.3% in FY 2002–2003 and of 2.8% in FY 2003–2004.

Recommendations: This study identified important issues with the quality of clinical coding at Ontario's case-costing hospitals. The following recommendations were designed to address the identified issues and to improve the quality of the clinical data in the Discharge Abstract Database:

1. Review the current concept of diagnosis typing with a view to improving the consistency of implementation.

2. Conduct further analysis on the FY 2004–2005 clinical data to assess the extent to which initiatives launched in 2003–2004 have had an impact on reducing the discrepancies noted in the study.

3. Establish Local Data Management Partnerships.

4. Establish a Physician Documentation Expert Panel to engage physicians in addressing chart documentation issues.

5. Conduct detailed analyses of the discrepancy rates within the case-costing hospitals to determine the specific factors contributing to the observed results.

™ Registered trade-mark of the Canadian Institute for Health Information.

1 Introduction

The Ontario Ministry of Health and Long-Term Care (MOHLTC), the Canadian Institute for Health information (CIHI), and CHIM Information Consulting Inc. (CHIM) jointly conducted a reabstraction study to review the clinical coding practices of Ontario's ten case-costing hospital corporations.

This reabstraction study breaks new ground in Ontario. It is the largest reabstraction study ever conducted in Canada and is among the largest studies ever conducted internationally. Approximately 14,500 discharges were reabstracted from the 18 sites within Ontario's ten case-costing hospital corporations. The discharges were sampled equally from the 2002–2003 and 2003–2004 fiscal years. This study therefore assesses the quality of data from the first two years of data abstracted using the International Statistical Classification of Diseases and Related Health Problems - Tenth Revision, Canada, and Canadian Classification of Health Interventions (ICD-10-CA and CCI). Furthermore, this study provides very timely information, as 2003–2004 is the most recent full fiscal year of hospital discharge abstract data currently available.

The quality of hospital clinical data used for performance measurement, planning, and resource allocation is critical to all hospital system stakeholders. Previous studies of Ontario discharges have identified discrepancies between original and reabstracted records that could lead to spurious measurements of hospital activity. This study's large sample size, compared to previous studies, will enhance the ability to draw meaningful conclusions about the nature and extent of observed variation in discharge abstract coding in Ontario case-costing corporations.

It is important to note at the beginning of this report that this study is the result of a collaborative effort between the MOHLTC, CIHI, and CHIM. While this partnership was certainly crucial to the operational success of the study, it also represents the commitment of each group to achieve the highest quality clinical data. The MOHLTC, CIHI, and CHIM look forward to continuing their synergistic collaboration and to engaging other data quality stakeholder groups.

2 Background

Ontario's ten case-costing hospital corporations use a standardized methodology to collect patient-specific cost data. These data include all direct and indirect costs incurred by the hospital in the provision of patient care. The patient-specific cost records are combined with the clinical records submitted to CIHI by hospitals. The resultant dataset is used for many important initiatives, including CIHI's development of a new ICD-10-CA and CCI case mix grouping methodology and of the associated resource intensity weights. The quality of clinical data from these hospitals is therefore of particular importance.

Several reabstraction studies with fewer selected discharges have been performed recently in Ontario. The discrepancies observed in these previous reabstraction studies were caused most often by differences in the significance assigned to diagnoses, and differences in the application of CIHI's coding standards[2]. The MOHLTC's 2003 Pilot Clinical Data Quality Audit, which examined five specific case mix groups (CMG) at a small sample of Ontario hospitals, showed that coding variation caused changes in the allocation of patients to CMG groups and major clinical categories (MCC). Variation in the assignment of co-morbid diagnoses also caused deviation in the results of the application of CIHI's complexity overlay to the CMG grouping methodology. Changes in MCC categories and CMG groups are important because they are used extensively in health services research and performance measurement. Furthermore, both CMG groups and complexity levels can independently affect measures of hospital output such as expected lengths of stay and resource intensity weights.

Previous reabstraction studies and other ongoing data quality improvement initiatives identified the following potential causes of observed coding variation:

- Inherent subjectivity in clinical coding
- Manipulation of data to maximize output measures from CIHI's grouping methodology
- Coding contrary to CIHI standards
- Non-comprehensive coding standards in some areas
- Incomplete documentation on the patient's chart
- Lack of clarity on chart documentation

One of the central challenges to improving the quality of clinical data is to quantify the relative impact of these possible causes on data quality. This reabstraction study will help address this challenge by closely describing the extent and nature of any observed coding variation. These findings will feed back into the MOHLTC-led,

[2] Hospital Medical Records Institute, Ministry of Health and Long-Term Care
"Data Quality of the Discharge Abstract Database Following the First-Year Implementation of ICD-10-CA/CCI", Canadian Institute for Health Information, Sept 2004.
"Accuracy of Complexity Measurement in Ontario: Results of the Clinical Data Quality Pilot Audit", Ministry of Health and Long-Term Care, 2003.

multi-stakeholder, data quality improvement committee, as well as informing the work performed by CIHI and CHIM.

By including a representative sample of all discharges from the participating case-costing hospitals, the study results will help identify any specific medical conditions that are particularly prone to coding variation. This study also includes the double reabstraction of approximately 800 discharges to specifically measure the inter-rater reliability of the health information professionals who participated in the study. The results of this inter-rater reliability sample will help establish baseline measures of variation in coding. They will also contribute to the identification of coding standards open to interpretation, coding standards that are not well applied, and confirm the impact on coding variation associated with current practices for chart documentation.

3 Objectives

There are four main objectives to this report:

1. To evaluate the quality of coding of clinical and non-clinical information at Ontario's case-costing hospitals
2. To assess the impact of any observed coding variation on measures of hospital output and resource utilization
3. To identify coding issues that arise as a result of the observed coding variation
4. To assess the findings from the inter-rater reliability dataset and to discuss the associated implications for the study results.

It is important to differentiate between the objectives of this report and the broader objectives of the reabstraction study. The aim of this report is to address each of the objectives outlined above. However, the results of the reabstraction study will serve numerous other purposes. CIHI and the MOHLTC will undoubtedly have cause to refer to the reabstraction datasets to support many aspects of their routine operations and a myriad of other special projects. This report therefore does not aim to provide a complete analysis of all the potential applications of the reabstraction study. Rather, CIHI and the MOHLTC intend for this report to provide a thorough assessment of the quality of abstracted data from Ontario's case-costing facilities. CIHI and the MOHLTC also hope that this report will facilitate informed decisions about the fitness of use of 2002–2003 and 2003–2004 DAD data for various purposes.

4 Subject Overview

This section provides information on the terminology used in this report relating to health data, coding practices, and inter-rater reliability.

4.1 Discharge Abstract Database

The Discharge Abstract Database (DAD) contains data on inpatient hospital discharges across Canada. CIHI receives DAD data either directly from participating hospitals or from the participating provinces. All provinces and territories, except Quebec, participate in the DAD as of April 1, 2005. The DAD contains demographic, administrative and clinical data for hospital discharges (inpatient acute, chronic, rehabilitation) and day surgeries.

4.2 Non-medical data

A selection of non-medical data is included in the study. They are essential data elements in the DAD that provide demographic and administrative information for each discharge. Some data elements directly impact the grouping methodology, such as birth date and weight for neonates. Refer to Appendix B for a complete listing.

4.3 Diagnosis Coding and Typing

Diagnosis coding relates to the practice of reviewing a patient's chart to identify pertinent health information and reporting it in a standardized format. Generally performed by health information professionals, the diagnostic information contained in the patient chart is assigned diagnosis codes using the International Statistical Classification of Diseases and Related Health Problems - Tenth Revision, Canada (ICD-10-CA). ICD-10 is developed and maintained by the World Health Organization. CIHI is charged with ensuring that the version is appropriate for Canadian use through the development and maintenance of enhancements for morbidity coding in Canada.

Diagnosis typing is an inherent part of the Canadian Coding Standards that must be applied to ensure consistent coding practices. It is used to indicate the relationship of a diagnosis to the patient's stay in a hospital. A diagnosis type is required for every ICD-10-CA code collected on the DAD abstract. The purpose of typing is to differentiate conditions that influence the patient's length of stay and/or resource intensity from those that do not. Typing also flags significant conditions that either coexist at the time of admission (pre-admit co-morbidity) or develop subsequently in hospital (post-admit co-morbidity). The appropriate typing of diagnoses is also important as this can affect CMG assignment. The diagnoses types which are included in CIHI's grouping methodology include type (M), (1), (2), (0), (W), (X) and (Y). Refer to Appendix D for a listing of diagnosis types.

4.4 Intervention Codes and Attributes

Intervention data are coded using the Canadian Classification of Health Interventions (CCI). CCI was developed by CIHI to complement ICD-10-CA and is the national standard for inpatient intervention coding. CCI has an expanded scope to encompass a broad spectrum of interventions to meet the needs across the continuum of health services in Canada. Interventions are grouped largely into therapeutic, diagnostic, and obstetrical, in addition to other interventions.

Attributes are separate data elements that provide extra detail about an intervention. Attributes are related to the intervention code and include: status, location, extent and mode of delivery. While most attributes are optional for DAD submission, some interventions have mandatory attributes. CIHI ensures that mandatory attributes are captured for codes that belong to particular intervention code sets in CCI upon data submission. Mandatory attributes can impact grouper outputs. They are also required for comparability to ICD-9-CM/CCP, the former intervention classification.

4.5 Coding Standards for ICD-10-CA and CCI

Coding standards are used to facilitate accurate and consistent coding practices. Within ICD-10-CA and CCI are coding guidelines, conventions and notes embedded in the classification to provide direction for basic coding. The Canadian Coding Standards for ICD-10-CA and CCI were developed by CIHI and clarify the general coding rules to provide additional guidance. The coding standards are vetted by a National Coding Advisory Committee that consists of representatives from every province and territory, except Quebec.

The Canadian Coding Standards include diagnosis typing definitions and provide guidance in determining whether a condition qualifies as a co-morbidity. The standards also address specific coding topics, grouped by the relevant chapter in ICD-10-CA. Coding standards are updated periodically and different standards may apply to different fiscal years.

4.6 CIHI Grouping Methodology Outputs

CIHI's grouping methodology produces output variables that enable aggregation of the data into clinically similar and cost homogeneous groups and provide national indicators for reporting resource utilization. The CIHI grouping methodology outputs discussed in this report are defined below.

Case Mix Group (CMG)

The case mix group (CMG) methodology aggregates patients into clusters based on clinical diagnoses, procedures and resource utilization. It is a methodology that describes the mix of patients treated in a hospital or jurisdiction (i.e. case mix). Acute care inpatients are assigned to CMG groups based on clinical and administrative data collected through the DAD. There are 478 CMG groupings in the 2003 CIHI grouping methodology.

Major Clinical Category (MCC)
The major clinical category (MCC) methodology aggregates patients more broadly than the CMG methodology. MCC categories generally describe a body system or specific type of clinical problem (e.g. mental disorders, neonates, burns, trauma, HIV). Each most responsible diagnosis is assigned to one of 25 MCC categories in the 2003 CIHI grouping methodology.

Complexity (Plx)
Complexity reflects the interaction of multiple diagnoses on length of stay or resources within each CMG group. Complexity overlay identifies those acute inpatients with additional diagnoses (other than the most responsible diagnosis) for which a prolonged length of stay and/or more costly treatment might be reasonably expected. These include cases with one or more chronic conditions outside of the primary focus of the acute care episode, cases with multi-system failure, and cases with iatrogenic or other complications. Cases are stratified into four complexity levels. A fifth level, level 9, is used for discharges where complexity is not applied.

1	No complexity
2	Complexity related to chronic conditions
3	Complexity related to serious/important condition
4	Complexity related to potentially life-threatening conditions
9	Complexity not applied

Resource Intensity Weight (RIW)
The resource intensity weight (RIW) methodology provides users with a tool to estimate expected resource use and relationships of costs between patient types. This methodology indicates the relative value of treating a patient compared to treating the average patient whose RIW weight is 1.0000. For example, an RIW value of 2.0000 represents a patient's visit that resulted in twice the expected resource use of the average patient. Values are calculated using actual patient-specific cost data from Ontario, Alberta, and British Columbia.

Expected Length of Stay (ELOS)
The expected length of stay (ELOS) algorithm estimates the duration of a typical acute care visit in a case mix group, measured in days. ELOS is a national average length of stay (ALOS) estimate that accounts for differences in age and complexity when these factors are found to be predictive of length of stay. Some CMG groups are refined for both age and complexity, generating as many as 12 different ELOS values (3 age groups multiplied by 4 complexity levels).

4.7 Inter-rater Reliability

Inter-rater reliability is the degree to which multiple evaluators obtain the same result and is an important measure of consistency. A high inter-rater reliability score indicates consistently coded data. Conversely, low measures of inter-rater reliability indicate high variation in the coding. Analysis of inter-rater reliability can indicate different levels of coder training, subjectivity or complexity of the subject matter, or the influence of an external factor.

4.8 Statistical Significance

A *sample* of the discharges submitted to the DAD from the case-costing hospitals was used in this study. As a result, it is necessary to determine how precisely the results of the *sample* reflect the true value if all discharges were reabstracted.

To determine how close an estimate (E) is to the true value, margins of error (ε) are included with the estimate, and are used to compute a confidence interval. The two estimates, combined together, would read $E \pm \varepsilon$ (e.g. 10% ± 5%). A confidence interval is a range of values that is likely to include the true population value. A 95% confidence interval means that if all possible samples were drawn from the population and the estimates were computed, 95% of the confidence intervals would contain the true population value.

For example, if the rate of discrepancy for item A is 10% ± 5% for FY 2002–2003 and 20% ± 6% for FY 2003–2004, we would like to know if the difference between these two years for item A is statistically significant. 'Statistical significance' means statistical analysis has revealed a difference unlikely to have occurred by chance alone.

In this example, the lower limit for item A in FY 2002–2003 is 5% (that is 10% - 5%); and the upper limit for item A for FY 2002–2003 is 15% (10% + 5%). Similarly, the lower and upper limits for item A in FY 2003–2004 are: 14% to 26%.

The estimates are said to be statistically significant if there is no overlap between the confidence intervals. In our example, the confidence interval for item A for FY 2002–2003 is [5% to 15%]; and for FY 2003–2004, the confidence interval is [14% to 26%].

As there is an overlap between these confidence intervals (i.e. one number could fall in both intervals such as 14% or 15%), the difference between the two disagreement rates would ***not be*** statistically significant.

5 Study Design

5.1 Target Population

The target population for this reabstraction study is all acute care inpatient discharges from all hospitals that submitted patient specific case-costing financial information to the Ontario Ministry of Health and Long-Term Care in FY 2002–2003 and FY 2003–2004. These hospitals are commonly referred to as case-costing hospitals. The frames from which the study sample was obtained from are the DAD for FY 2002–2003 and FY 2003–2004 as they were on closing.

Ontario's ten case-costing hospital corporations represent 18 different hospital sites. Three of these sites submit data to CIHI under one facility number, thus resulting in data being submitted under 16 *unique* facility numbers. As a result, the sample selection and data analysis was based on data submitted under the 16 facility numbers.

The following tables compare, by facility type, the case-costing hospitals to all Ontario acute care facilities. Table 5.1 illustrates that the case-costing facilities represent a small portion of all Ontario acute care facilities (around 10%). However, as illustrated in Table 5.2, these case-costing facilities submit a substantial proportion of the inpatient data to the DAD for Ontario (almost 25%).

Table 5.1: Distribution of the Ontario Case-Costing Facilities by Facility Type

Facility Type	Ontario Case Costing Facilities	All Acute Care Ontario Facilities *
Small	1	53
Community	10	99
Teaching	5	15
Total	**16**	**167**

Source: CIHI 2005

Note: * - Includes case-costing facilities

Table 5.2: Allocation of Discharges in the DAD for Ontario, by Facility Type

Facility Type	DAD Fiscal Year 2002–2003			DAD Fiscal Year 2003–2004		
	Case Costing Discharges	Total Ontario Discharges	%	Case Costing Discharges	Total Ontario Discharges	%
Small	1,249	45,966	2.7	1,255	45,037	2.8
Community	126,274	765,335	16.5	127,362	755,546	16.9
Teaching	139,267	291,771	47.7	136,755	298,874	45.8
Total	**266,790**	**1,103,072**	**24.2**	**265,372**	**1,110,042**	**23.9**

Source: CIHI 2005

5.2 Study Datasets

This study involved the reabstraction of approximately 14,500 discharges, from which about 800 discharges were reabstracted twice to measure the inter-rater reliability of the study's reabstractors. This resulted in two samples being selected: a "primary" dataset and an inter-rater dataset. These are described below.

5.2.1 Primary Dataset

The primary dataset was selected to obtain estimates of the variation in the data originally submitted by hospitals to the DAD. The number of discharges selected for the study was obtained using a stratified sampling design. The population of discharges was divided into strata, with each stratum consisting of all discharges that were in a particular institution and complexity level combination.

For the eleven larger facilities, the target sample size (n_{ij}) for each stratum was calculated using Formula 5A. Larger facilities are those that submitted more than 10,000 discharges during FY 2002–2003.

Formula 5A

$$n_{ij} = \frac{\dfrac{t^2 * (\hat{p}_j(1-\hat{p}_j))}{d^2}}{1 + \dfrac{1}{N_{ij}} * \left(\dfrac{t^2 * (\hat{p}_j(1-\hat{p}_j)}{d^2} - 1 \right)}$$

where:

- N_{ij} = the population size for facility i and complexity level j in FY 2002–2003
- d = the desired margin of error, set at 5%
- t = the value of the normal probability function that corresponds to a 95% confidence interval, which is 1.96.
- $\hat{p}_j$ = the estimated probability of an event occurring for complexity level j

Estimated probabilities for each stratum ($\hat{p}_j$) are the rates of discrepancies for either case mix group[3] or complexity level from the results for Ontario facilities from the national reabstraction study on the first year of implementation of ICD-10-CA and CCI[4]. These discrepancy rates were both used in Formula 5A, and the rate that yielded a higher target sample size n_{ij} was used in the study.

A different method was used to calculate the target sample size for the smaller facilities. Smaller facilities are those that submitted fewer than 2,500 discharges in

[3] The case mix group discrepancy rate was calculated within each original complexity level.

[4] "Data Quality of the Discharge Abstract Database Following the First-Year Implementation of ICD-10-CA/CCI", Canadian Institute for Health Information, September 2004

FY 2002–2003[5]. For these facilities, the sample size was set at 50% of the population count of discharges in each stratum, up to a specified maximum. The maximum was 100 discharges for complexity levels 1 and 9 and 200 discharges for complexity levels 2, 3 and 4.

It should be noted that the definition of stratum used for sample size determination does not divide the records by year. At the time the sample size was calculated, the number of discharges in the population for FY 2003–2004 was not known. The sample size that was determined for each stratum using the above methods was evenly divided between the two fiscal years of the study.

After the sample size was determined for each stratum, the next step was to select the sample. Discharges were selected using a probability proportional to size methodology, which uses a measure of size to determine the probability of selection. That is, the probability of selection is proportional to the size of the unit, where resource intensity weight (RIW)[6] values were used as the size indicator. This resulted in discharges with a higher RIW value having a greater probability of selection in the study.

The probability of selection for a discharge (p_{ijk}) is further illustrated in Formula 5B.

Formula 5B

$$p_{ijk} = n_{ij} * \frac{s_{ijk}}{\sum_{k=1}^{N_{ij}} s_{ijk}}$$

where:

n_{ij} = the sample size for facility *i* and complexity level *j*
N_{ij} = the population size for facility *i* and complexity level *j*
s_{ijk} = the RIW value for discharge *k* in facility *i* and complexity level *j*

Once probabilities were assigned to each discharge, a sample was selected using systematic sampling with a random start.

Tables 5.2.1.1 and 5.2.1.2 show the *target* number of discharges selected for the Case Costing study for each facility, by complexity level. Note that these figures were increased by 20% to obtain the number of discharges that were *sampled* to account for those discharges that were unavailable during field collection.

[5] Note that there were no facilities that submitted between 2,500 and 10,000 discharges.
[6] Resource intensity weight was calculated using 2002v3 CIHI grouping methodology on the original data submitted to the DAD for FY 2002–2003, and the 2003 CIHI grouping methodology on the original data submitted to the DAD for FY 2003–2004.

Table 5.2.1.1: Case Costing Target Sample Sizes for the Primary Dataset for FY 2002–2003

Facility	Original Complexity Level					
	1	2	3	4	9	Total
Mount Sinai Hospital	64	139	135	143	58	**539**
St. Michael's Hospital	65	158	152	161	58	**594**
Arnprior & District Memorial Hospital	50	48	19	9	15	**141**
Credit Valley Hospital	65	137	123	113	58	**496**
London Health Sciences Centre	65	165	165	174	58	**627**
University Health Network	65	164	165	177	52	**623**
Trillium Health Centre	65	159	152	155	58	**589**
William Osler - Etobicoke	64	130	109	97	58	**458**
Lakeridge Health - Oshawa	65	144	131	134	58	**532**
Quinte - Belleville	64	130	119	101	57	**471**
Quinte - North Hastings	50	21	8	4	7	**90**
Quinte - Prince Edward County	50	63	27	10	50	**200**
Quinte - Trenton	50	100	43	13	10	**216**
William Osler - Brampton	65	144	119	107	58	**493**
William Osler - Georgetown	50	57	35	19	50	**211**
Ottawa Hospital, General Campus	65	152	150	164	58	**589**
Total	**962**	**1,911**	**1,652**	**1,581**	**763**	**6,869**

Source: CIHI 2005

Table 5.2.1.2: Case Costing Target Sample Sizes for the Primary Dataset for FY 2003–2004

Facility	Original Complexity Level					
	1	2	3	4	9	Total
Mount Sinai Hospital	65	141	136	145	59	**546**
St. Michael's Hospital	66	159	153	162	58	**597**
Arnprior & District Memorial Hospital	50	48	20	10	16	**144**
Credit Valley Hospital	65	137	124	114	59	**500**
London Health Sciences Centre	66	166	166	175	58	**631**
University Health Network	66	165	167	178	53	**629**
Trillium Health Centre	66	161	153	155	59	**594**
William Osler - Etobicoke	66	130	110	99	58	**463**
Lakeridge Health - Oshawa	65	145	132	135	58	**535**
Quinte - Belleville	64	131	120	101	58	**474**
Quinte - North Hastings	51	21	10	5	9	**95**
Quinte - Prince Edward County	50	65	28	11	50	**204**
Quinte - Trenton	50	101	44	14	11	**219**
William Osler - Brampton	66	146	120	109	59	**499**
William Osler - Georgetown	50	58	35	20	50	**213**
Ottawa Hospital, General Campus	65	153	151	166	59	**594**
Total	**971**	**1,927**	**1,666**	**1,598**	**774**	**6,935**

Source: CIHI 2005

5.2.2 Inter-rater Dataset

A sample of discharges from the primary dataset was randomly selected for inclusion in the inter-rater reliability study. At each facility, 5% of the primary sample discharges were randomly selected without regards to complexity level. For these discharges, two reabstractors were randomly assigned to abstract the same discharge. It should be noted that one randomly selected copy of the inter-rater discharge was retained in the primary dataset. The number of inter-rater discharges selected for each fiscal year is shown in Table 5.2.2.1.

Table 5.2.2.1: Case Costing Sample Sizes for the Inter-rater Dataset

Fiscal Year	Original Complexity Level					
	1	2	3	4	9	Total
2002–2003	58	115	99	95	46	**413**
2003–2004	58	116	100	96	46	**416**

Source: CIHI 2005

6 Data Collection

The data collection activities for the Ontario Case Costing study consisted of three broad events: reabstraction application modification, recruitment and training, and field collection.

6.1 Reabstraction Application Modification

Data for this study were collected on reabstraction software previously developed by CIHI, with changes incorporated to meet the needs of the study. The software application allowed reabstractors to easily enter reabstracted codes, and to compare these against the original DAD data previously loaded into the application. Hard edits were built into the system at data entry to minimize errors in the reabstracted data. For example, reabstracted diagnosis codes that do not exist were not accepted.

Two types of data capture were applied: verification and blind reabstraction. Verification refers to entries required due to a disparity with the original DAD value. This mode was chosen for the non-clinical fields (such as gender) to both reduce the reabstractor's response burden and the potential for data entry errors. In contrast, blind reabstraction was used for the clinical fields. Here, reabstraction of the clinical information had to be completed before the original DAD data was revealed. Reabstractors then linked the original and reabstracted clinical fields (see section 6.2.3). The application automatically output any discrepancies between the original and reabstracted data elements and the reabstractors were then prompted to enter a reason for each discrepancy.

Before field collection began, the reabstraction software was tested on a pilot facility. Three CIHI classification specialists reabstracted approximately 100 discharges of varying complexity to ensure the functionality of the software application.

6.2 Recruitment and Training

CHIM Information Consulting Inc. advertised the opportunity to participate in this study to health information professionals across Canada. Applicants were assessed in an interview performed by CIHI and CHIM. Upon successful completion of this screening process, applicants were then invited to a five-day training session conducted by CIHI.

Training consisted of a review of coding standards and training on the reabstraction software. Reabstractors were instructed to refer to the appropriate fiscal year of coding standards when participating in the study. Reabstractors were also provided with detailed guidelines to facilitate consistent reabstraction techniques.

Table 6.2.1 summarizes the results of the recruitment of health information professionals for the Case Costing study. Health information professionals were recruited from across Canada. Those that participated in the study came from Ontario, British Columbia, Alberta and Nova Scotia.

Table 6.2.1: Recruitment Results for the Ontario Case Costing Study

Interviewed	41
Passed interview	30
Trained	29
Passed training	25
Participated in study	24

Source: CIHI 2005

6.2.1 Training for Diagnosis Codes

For the purpose of this specific study, coders were instructed to capture certain information from the discharge as explained below. ***This is important to bear in mind while reviewing the results as only those items reabstracted as per the training instructions are being studied.***

- All conditions that met the criteria of significant diagnosis types were reabstracted. *Please refer to Appendix E for the criteria used to determine significance of a condition.*
- Secondary conditions (i.e. diagnosis type (3)) were only reabstracted in specific cases. These include ICD-10-CA asterisk codes, primary neoplasm coded as present or history of, outcome of delivery codes (Z37), and additional diabetes codes applicable in pop-up boxes. These diagnoses were termed “mandatory” for the purpose of the study and were based on coding standards or directives within the ICD-10-CA classification.
- Optional conditions (i.e. diagnosis type (0)) were reabstracted when the coding applied to a newborn chart. Codes from category Z38 were also reabstracted.
- Conditions assigned a diagnosis type (9) were reabstracted only when capturing external cause codes and place of occurrence. Activity codes (U99) were not reabstracted.

6.2.2 Training for Intervention Codes

For this specific study, coders were provided with the following guidelines to facilitate consistent reabstraction of intervention codes in CCI. ***This is important to bear in mind while reviewing the results as only those items reabstracted as per the training instructions are being studied.*** A list was provided in training of interventions with an intervention type code less than 50 that affect CMGs in both fiscal years.

- All therapeutic interventions (section 1 of CCI) were reabstracted when the intervention type code was 50 and greater, with some exceptions.

Interventions with an intervention type code of less than 50 that affect CMG assignment were also reabstracted.
- All diagnostic interventions (section 2 of CCI) relating to inspections and biopsies were reabstracted, in addition to those that affect CMG assignment.
- Cardiac catheterizations (section 3 of CCI) were reabstracted in both fiscal years. For fiscal year 2003–2004 only, CT scans, angiograms, and MRI were reabstracted.
- Obstetrical and fetal interventions (section 5 of CCI) were reabstracted when the intervention type code was greater than 45. As per coding standards, induction of labor was also coded.
- Attributes that were mandatory for DAD submission in 2002–2003 and 2003–2004 were also reabstracted.

6.2.3 Training for Linking Original Data to Reabstracted Data

The application software allowed for a one-to-one linking of original and reabstracted diagnosis and intervention codes. If the same condition or intervention was described, regardless of differing codes, the original and reabstracted codes could be linked. When a condition or intervention was present only in the original data or only in the reabstracted data, no link would be assigned, and a discrepancy of “original only” or “reabstracted only” was generated.

6.2.4 Training for Discrepancy Reasons

The reabstraction software allowed up to four reason codes when a discrepancy was identified between the original and reabstracted data. For a full description of reason codes and the scenarios in which they were used, refer to Appendix C.

Standards/Codebook/Manual: This reason was assigned when the reabstractor attributed the discrepancy as a result of the original coder not following a coding standard, a coding directive in ICD-10-CA/CCI, or a specification within the DAD Abstracting Manual.

Significance: This reason was used when the original data was typed as significant (i.e. diagnosis types (M) (1) (2) (W) (X) (Y)), and the reabstractor did not code it as such, yet agreed the condition existed. The reason was also applicable in the reverse scenario, when the original data was assigned a diagnosis type of secondary (i.e. (3) (0) (9)) but the reabstractor coded it as significant.

Chart Documentation: This reason was assigned when the reabstractor attributed the discrepancy as a result of incomplete or conflicting chart documentation, when the reabstractor’s interpretation of the information contained in the discharge was different from what was submitted to the original DAD, information on the chart was missed, or the specificity of the code selection was not supported in the chart.

Optional/Not Wrong: This reason was used in three scenarios. First, if the original diagnosis or intervention code was present, but was not reabstracted as per the training instructions. These original clinical fields are excluded from analysis. The other two scenarios that use this reason code *are* included in the study results. These are when the reabstractor entered a different diagnosis code, but both the original and reabstracted diagnosis code were correct, and when a discrepancy had already been assigned to a particular coding difference and it was unnecessary to flag the same error twice[7].

6.2.5 Training Test

At the end of training, applicants were given an inter-rater coding test based on a CIHI gold standard. The gold standard was developed by five CIHI classification specialists who independently coded the same twenty discharges. The team then met to discuss results and achieved a consensus that became known as the gold standard. Final coding decisions were based on coding standards and other CIHI resources including education material and directives from the CIHI Coding Query Database.

The gold standard was used as a benchmark against which applicants were assessed when coding these same twenty discharges using the reabstraction application. Instead of viewing the discharge in its original form, applicants were provided with only the pertinent information from the discharge in a typed format. This was done to reduce effort of reading extraneous chart documentation that was unrelated to the inter-rater test. It also removed errors related to illegible handwriting. All discharges used in the inter-rater test had sufficient chart documentation to assign each condition the correct diagnosis code and diagnosis type.

Diagnosis codes entered by the applicants were compared to the gold standard, and those closer to the gold standard achieved higher inter-rater scores.

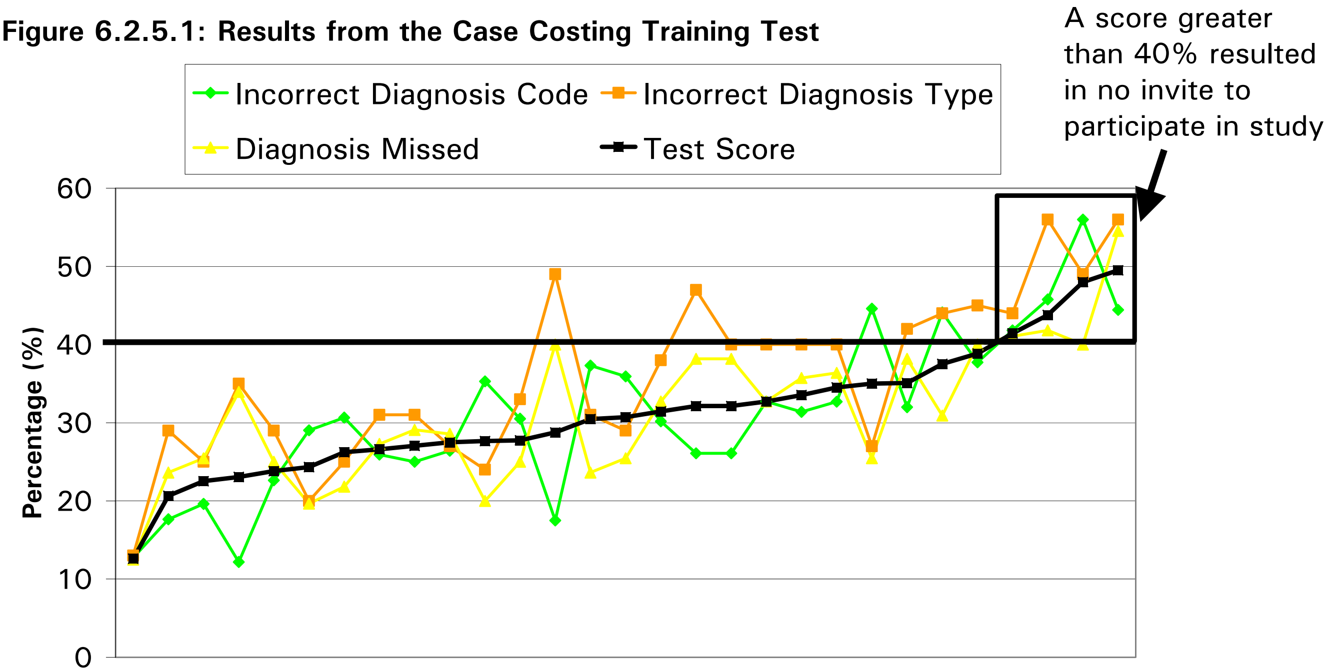
Figure 6.2.5.1 illustrates the results from the Case Costing training test from the lowest score to highest. The test scores show the disagreement rate of three variables: diagnosis code selection, diagnosis type selection, and diagnosis missed. The diagnosis code score is shown in the figure with a green line. The points that are connected by this line represent the individual coder scores. Similarly, the diagnosis type score is illustrated with an orange line, and diagnosis missed score is illustrated with a yellow line.

The final test score received by the applicants is the average of the "incorrect diagnosis code" and the "diagnosis missing" percentages, and is illustrated in the figure with the black line. As this process for screening applicants was relatively new, the pass mark was set at a 60% agreement rate. That is, applicants who received a test score greater than 40% (for disagreement) were not invited to

[7] An example of two discrepancies that relate to one error is the assignment of the non-medical data element, "Institution To". If a reabstractor entered a different institution number for this data element, it is possible that the "Institution To Type" data element will also be different. Since the two discrepancies relate to the one error in "Institution To", the second discrepancy generated for "Institution To Type" would be assigned a reason of "Optional/Not Wrong".

participate in the study. This group consisted of four participants, which are labeled in the graph with the box in the upper right hand corner.

Figure 6.2.5.1: Results from the Case Costing Training Test



Source: CIHI 2005

6.3 Field Collection

Data collection for this study took place from July 2004 to January 2005.

The 24 reabstractors that participated in the study either traveled from facility to facility (on-site), worked at CIHI's Toronto office (modified on-site), or worked from home (remote). For on-site coding, reabstractors worked at the facility. CHIM coordinated the multi-week accommodations for the reabstractors and their transportation to each facility. These facilities were required to accommodate the chart pull and spacing demands. For one facility that could not accommodate the space, temporary workstations were set up at CIHI's Toronto office. Facilities that opted for remote coding pulled the sample charts and shipped them to an external vendor for document conversion. Salumatics was subcontracted by CHIM to scan paper documents and make them available electronically through a secure virtual private network connection. This allowed reabstractors to view the discharges from home.

The distribution of on-site and remote coding facilities differed between the two fiscal years of the study. For FY 2002–2003, one of 16 case-costing facilities was reabstracted remotely. For FY 2003–2004, remote coding increased to 12 of the 16 case-costing facilities.

Reabstractors were provided with a variety of coding resources, including:

- ICD-10-CA and CCI software
- Canadian Coding Standards for ICD-10-CA and CCI
- DAD Abstracting Manual
- CIHI Coding Query Database
- CIHI training binder

In the event these resources could not help resolve a question, reabstractors contacted CIHI directly for instruction. This applied to both on-site and remote data collection. CIHI also dealt with any technical questions related to laptops or the reabstraction application.

Reabstractors backed-up the data collected each day on a USB key. This was a precaution to prevent the loss of data in the event of a technical malfunction. The USB keys were also shipped back to CIHI so that the data could be downloaded.

CIHI supplied each facility and reabstractor a list of randomly assigned sample discharges. Discharges that comprised the inter-rater reliability sample were also included on the list of discharges assigned to reabstractors. Reabstractors and hospitals were also supplied with a supplemental list that contained replacement discharges to use if discharges from the original list were not available (refer to section 5.1). Replacement discharges were to match the same sampling criteria as the missing discharge, meaning that the complexity level had to match that of the unavailable discharge.

Table 6.3.1 gives the response rate for the primary dataset in the Case Costing study.

Table 6.3.1: Response Rate for the Ontario Case Costing Study, Primary Dataset

Number of Unique Discharges	DAD Fiscal Year 2002–2003	DAD Fiscal Year 2003–2004
Target	6,869	6,935
Selected	8,247	8,322
Re-abstracted	6,863	6,940
Percent of Target	**99.9**	**100.1**

Source: CIHI 2005

6.4 Privacy, Confidentiality and Security

CIHI policies on privacy, confidentiality and security, with respect to personal privacy and safeguarding the confidentiality of individual records and facilities, were adhered to throughout the course of the study. Information on CIHI policies can be found online at www.cihi.ca.

CHIM policies on confidentiality and code of ethics were adhered to throughout the course of the study. Information on CHIM policies can be found online at www.chima-cchra.ca.

7 Data Processing

Data processing transforms data captured during field collection into a form that is suitable for data analysis. It includes all data handling activities, whether they are automated or manual, after data collection and prior to estimation.

7.1 Edit and Imputation

Editing is the application of checks to identify missing, invalid or inconsistent entries that point to data records that are potentially in error[8]. Validity and consistency edits were applied to the Case Costing data. Validity edits verify that the entry for an individual data element is legitimate (e.g. a date of February 30th would be flagged). Consistency edits verify that the relationship between data elements is respected (e.g. a male patient cannot give birth to a baby).

The Case Costing application had limited capacity to ensure the reabstracted data were of the highest quality possible. As a result, additional edits were applied to the reabstracted data after field collection. The following is a general description of these edits.

Most edits looked to the quality of the reabstracted data (e.g. did the reabstractor enter a valid combination of diagnoses). Some of these edits are presently in place for DAD submission, meaning that the original data were also subjected to the same edits. Other edits ensured that the reabstractor complied with the various coding guidelines and standards.

The remaining edits related to data elements specific to the Case Costing application to check that the rules presented in training were followed consistently by the reabstractors. One such edit verified the linkage of original and reabstracted clinical fields. This edit identified when a link was not made when it should have been, and vice versa. Another edit checked the assignment of reason codes. For example, the assignment of the "significance" discrepancy reason is only applicable when a diagnosis changed significance upon reabstraction[9].

Imputation is a process used to determine and assign replacement values to resolve problems of missing, invalid, or inconsistent data. One imputation was applied to the Case Costing data.

As previously mentioned, reabstractors were instructed not to code most secondary conditions (diagnosis type (3)), even if they were originally submitted to the DAD as significant. Reabstractors assigned a reason of "significance" to indicate when an original significant condition was not reabstracted due to it being a type (3)

[8] Definitions used in the Edit and Imputation section are a slight modification of those provided in "Survey Methods and Practices", Statistics Canada, October 2003.
[9] For FY 2002–2003 Case Costing data, imputations to 491 reabstracted non-medical and medical data elements, 160 linkages, and 234 reason codes were performed as a result of the edit checks. For FY 2003–2004, these figures are 593, 216, and 64 respectively.

condition. If an original significant condition was not reabstracted because the condition did not exist on the discharge, the coder would have assigned a different reason for the discrepancy (e.g. “chart documentation”).

From this logic, it can be inferred that when an original diagnosis was submitted as significant, and the reabstractor did not concur and assigned a discrepancy reason of “significance”, the reabstracted diagnosis would have been a secondary condition if the reabstractor had coded it. Therefore, these instances have reabstracted diagnosis types imputed to (3).

7.2 Grouper Outputs

The patient grouping methodology was applied to both the original and reabstracted data once data editing and imputation was completed. The 2003 CIHI grouping methodology was applied to the original and reabstracted data for both fiscal years of data.

7.3 Weighting

Completed records were assigned a “weight” to ensure that the sampled discharges accurately reflect the population that is being estimated. That is, the sample weight of a discharge is the number of units in the population the sampled unit represents. Due to the study design, discharges have different probabilities of selection and must be weighted accordingly. Although discharges with higher RIW values had a higher probability of selection, the weighting ensures that records with higher RIW values are not over-represented in the sample.

Sampling weights are inversely proportional to the probability of selection, where the probability of selection is described in section 5.1. However, since the study included a 20% over-sampling of discharges and contained non-response, the weights of the records were “benchmarked” against known population totals to ensure that the correct number of discharges was being estimated for each stratum. Since RIW values were used as a size indicator, unadjusted weights would not necessarily sum to the number of discharges in the population. The weights of the records (w_{ijk}) were obtained using the Formula 7A:

Formula 7A

$$w_{ijk} = \frac{N_{ij}}{\sum_{k=1}^{R_{ij}} \frac{1}{p_{ijk}}} * \frac{1}{p_{ijk}}$$

where:

N_{ij} = the population count of discharges for facility i and complexity level j
R_{ij} = the number of completed discharges for facility i and complexity level j
p_{ijk} = the probability of selection of the individual discharge k in facility i and complexity level j

7.4 Estimation

7.4.1 Point Estimation

The estimate for the total number of units in the Case Costing study population satisfying a certain condition is calculated by summing the weights for the subset of records that meet that condition. This is illustrated in Formula 7B, known as the Horvitz-Thompson estimator for totals.

Formula 7B $$\hat{T}_c = \sum w_{ijk} * c_{ijk}$$

where:

w_{ijk} = sample weight for facility *i*, complexity level *j*, and discharge *k*
c_{ijk} = 0 if condition *c* is false for facility *i*, complexity level *j*, and discharge *k*
1 if condition *c* is true for facility *i*, complexity level *j*, and discharge *k*

The estimate for the proportion of units in the Case Costing study population satisfying condition *c* is calculated by first determining the estimate of the total number of units that satisfy that condition (as illustrated above), and then dividing this by the estimate of the population total. This is illustrated in Formula 7C, known as the Horvitz-Thompson estimator for proportions.

Formula 7C $$\hat{P}_c = \frac{\sum w_{ijk} * c_{ijk}}{\sum w_{ijk}}$$

where:

w_{ijk} = sample weight for facility *i*, complexity level *j*, and discharge *k*
c_{ijk} = 0 if condition *c* is false for facility *i*, complexity level *j*, and discharge *k*
1 if condition *c* is true for facility *i*, complexity level *j*, and discharge *k*

7.4.1.1 Coder Effect on Point Estimates

The study estimates are subject to bias due to the variability in agreement rates with the original data found between reabstractors. In particular, facility specific results are dependent on which reabstractors collected its data. By adjusting the results for coder effect, the study findings are impartial to this potential bias.

Adjustments for coder effect were determined by applying a logistic regression model to the Case Costing data. Logistic regression is a type of predictive model that can be used with categorical data elements that have exactly two categories (i.e., a binary or dichotomous variable).

The logistic model formula computes the probability of the selected outcome as a function of the values of the predictor variables. The equation used is shown in Formula 7D. The model estimates the β values given the various combinations of coders, facilities, and fiscal years of data[10]. The estimated β values were then applied to the Case Costing data and Formula 7D was solved for $\hat{p}_{ijk}$, the coder effect adjusted estimate for fiscal year *i*, facility *j*, and reabstractor *k*.

Formula 7D

$$\log\left(\frac{\hat{p}_{ijk}}{1-\hat{p}_{ijk}}\right) = \hat{\beta}_0 + \underbrace{\hat{\beta}_i\hat{X}_i + \hat{\beta}_j\hat{X}_j + \hat{\beta}_k\hat{X}_k}_{\text{main effects}} + \underbrace{\hat{\beta}_{ij}\hat{X}_{ij}}_{\text{interaction}}$$

Only the fiscal year and facility interaction was included in the model, even when it was not significant at the 5% level. The quality of data collected by a facility is expected to change between years.

Interactions involving the reabstractor were not included in the logistic regression because many reabstractors collected data from only one of the two fiscal years of data, and others reabstracted from a small number of 16 case-costing facilities. If included in the model, reabstractor interactions would produce abnormally wide confidence intervals.

Adjustments for coder effect were limited to variables of interest that have only one of two possible outcomes. Hence, some findings were not analyzed for coder effect. ***All tables and graphs will state if the results are unadjusted or adjusted for coder effect.***

7.4.1.2 Case Mix Effect on Point Estimates

Differences in the types of conditions presented by the patient population between facilities affect the comparability of results. This is referred to as case mix effect.

The Case Costing data found that facility agreement rates differed in part due to variation in the types of conditions presented. Some facilities treat a high proportion of obstetrical and neonate patients, while others treat few or no patients of these types. The codes associated with obstetrical and neonate patients were reabstracted with high agreement across all facilities. Hence, much of the difference between facilities was in the proportion of diagnoses (or discharges) that were associated with these types of patients.

When comparing data between facilities, ***the population of reference was redefined in the analysis to reduce the case mix effect***. This was also done for comparisons between the inter-rater and primary datasets. Agreement rates were recalculated for subsets of the study data that did not relate to obstetric or neonate patients. This is further described below.

- For analysis of the diagnosis data, original codes belonging to the obstetrical or neonate chapters were excluded, in addition to codes Z37-Z38. This was done for diagnosis code, diagnosis type, and the most responsible diagnosis.

[10] For example, $X(i=2002/2003) = \begin{cases} 1 \text{ if fiscal year is 2002–2003} \\ 0 \text{ otherwise} \end{cases}$

- For analysis of the CIHI grouping methodology variables, discharges originally assigned to complexity level 9 were excluded. Complexity level 9 indicates that a complexity overlay was not applied, and is assigned to obstetrical, neonate, and mental health discharges.
- Case mix effect was considered in the inter-rater analysis. When comparing inter-rater results to the primary dataset results, diagnoses (or discharges) related to obstetrics or neonates were excluded from analysis.

Non-medical and intervention data were not adjusted for case mix effect.

7.4.2 Variance Estimation

Two methodologies were applied in estimating variances. One method applied to estimates that were not adjusted for coder effect, and the second method applied to estimates that were adjusted for coder effect. These are described below.

7.4.2.1 Variances for Point Estimates without Adjustments for Coder Effect

Margins of error for the unadjusted results were calculated using a bootstrap re-sampling methodology. In bootstrap variance estimation, multiple random sub-samples are selected with replacement from the collected data. Each of these bootstrap samples is then weighted using the benchmarking methodology described in section 7.3. The variation in the estimates calculated from the bootstrap samples is used to calculate the margin of error, which is an indication of the amount of variability associated with an estimate due to the estimate being based on a sample. The margins of error are based on 95% confidence intervals. The margin of error is expressed as a "±" after the estimate (e.g. 10 ± 2). S*maller the margins of error indicate more precise the estimates.*

In the Case Costing study, there were 500 bootstrap samples generated and each sample was weighted using the same techniques described in section 7.3.

7.4.2.2 Variances for Point Estimates with Adjustments for Coder Effect

The significance of the logistic regression used to adjust estimates for coder effect were evaluated by the log likelihood test, given as the model chi-square test, based on a 95% confidence level. This is known as the Wald statistic.

The confidence intervals generated by the Wald statistic do not have upper and lower limits that are equally distanced from the point estimate. As a result, the margin of error cannot be expressed as a "±" after the estimate. Instead, measures of variance that accompany adjusted estimates will read as "10, [9, 12]", where 10 is the point estimate, 9 is the lower limit of the confidence interval, and 12 is the upper limit of the confidence interval.

8 Methodological Notes

Due to the complex nature of the subject matter and the way in which the data were collected, there are some limitations to the study data that should be noted.

8.1 Population of Facilities

The study was designed only to look at the data from the case-costing facilities in Ontario and only data from these facilities were collected. The results from the study are therefore specific to the data from these facilities.

8.2 Sample versus Estimated Results

Weighted estimates provide a better indication of the level of discrepancies in the population. *Only weighted estimates are presented in the report.* Some estimates will not add to the total due to rounding error.

8.3 Data Limitations for the Primary Dataset

8.3.1 Primary Dataset Sample Size

Although the Case Costing study consisted of 13,803 discharges, these were not allocated evenly across all major clinical categories or case mix groups. When looking at the results for these grouper outputs, it is important to note that some major clinical category and case mix group estimates may be based on small samples. The margin of error associated with the estimates gives a measure of how much confidence should be placed in the results. Estimates that are based on small sample sizes generally have larger margins of error. Refer to Appendix A for the full definition of margin of error.

8.3.2 Primary Dataset Diagnosis Codes

There are two data limitations to note relating to diagnosis codes.

The first limitation is related to linkage. Reabstractors were instructed to link diagnoses based on the condition that the codes were describing. For example, two codes describing a myocardial infarction would be linked, whether or not the same code was used to describe the condition. However, the concept of "same condition" has an element of subjectivity, particularly as two codes become less similar.

The second limitation is related to secondary conditions. Recall that secondary conditions were not reabstracted even if the diagnosis was identified as a significant condition in the original data. Here, the reabstractor would use the "significance" discrepancy reason instead of entering a code. Because of this, diagnoses codes are only compared when both the original diagnosis and the reabstracted diagnosis were identified as significant.

8.3.3 Primary Dataset Diagnosis Typing

According to the coding standards for the DAD, most conditions are only mandatory to code if they are the most responsible diagnosis or a co-morbid condition. This means that most secondary conditions (diagnosis type (3)) are optional to code. Accordingly, the reabstractors were given instructions to only code secondary conditions if they belonged to a specific subset of conditions.

For significant conditions identified as being present in the reabstracted data only, it is not possible to tell whether the original coder thought the condition was present. That is, the original coder may have felt the condition was not significant and therefore did not code the condition (as it was optional). Or, he/she may have not coded it because it was felt the condition was not present.

8.3.4 Primary Dataset Interventions

There are two limitations in the analysis of interventions.

The first limitation is related to linkage. Reabstractors were instructed to link the original and reabstracted interventions; however, the concept of "same intervention" has an element of subjectivity, particularly as two codes become less similar.

The second limitation is related to optional interventions. As mentioned in section 6.2.2, reabstractors were instructed to only code interventions and attributes that were mandatory for the study. Because of this, all "non-mandatory" interventions and attributes are excluded from analysis in the study. The code selection analysis is based solely off of the interventions and attributes deemed mandatory for the study by both the original coder and by the reabstractor.

8.4 Data Limitations for the Inter-rater Dataset

8.4.1 Inter-rater Dataset Sample Size

The sample size for estimates poses a greater limitation on the inter-rater results. The inter-rater results are based on a smaller sample than the results for the primary dataset. Inter-rater estimates often have wide margins of error.

8.4.2 Inter-rater Dataset Diagnosis Codes

Since the reabstractors compared the coded conditions to what was originally submitted to the DAD, rather than comparing them to each other, it was necessary to create links between the reabstractors' data. When the reabstractors referred to a condition that was present in the original data, the data that each reabstractor linked to the original condition was then linked to each other. When the reabstractors' data contained diagnoses that were not present in the original data submission, it was necessary for a CIHI classification specialist to examine each of the reabstractors' data to determine if the codes were referring to the same condition. For those diagnoses linked by a CIHI classification specialist, no reason for discrepancies between reabstractors can be deduced. Analysis of the reasons for discrepancies for this subset of codes will show a "cannot infer" reason code.

Similar to the results in the primary dataset, code selection comparisons are made when both reabstractors identified that the condition was significant.

8.4.3 Inter-rater Dataset Diagnosis Typing

Where only one of the two reabstractors assigned to an inter-rater chart coded a condition, it is not always possible to determine whether the second reabstractor assessed the condition as secondary, or found it to be not present on the chart at all.

The typing analysis that is done is based on conditions that both reabstractors coded as being present.

8.4.4 Inter-rater Dataset Interventions

Similar to the situation with the primary dataset, analysis of interventions is limited to those that were mandatory for the study. There were not enough mandatory attributes in the inter-rater dataset to make any comments on the coding of mandatory attributes.

8.5 Coder Effect

There are a variety of influences that determine how health information professionals view and code the same source data. Some of these factors are:

- Years of experience
- Experience working in different hospital settings
- Experience coding a variety of discharges
- Educational background (including workshops)
- Facility specific factors influencing the present work environment

The Case Costing reabstractors are also subject to these influences. This means that it is possible for facility specific results to be dependent on the selection of reabstractors that captured the data for that facility. For instance, reabstractors assigned to one facility could be more prone to assign a condition as significant than other reabstractors. Hence, if different reabstractors were assigned to that facility, the study results would be affected accordingly. The influence the reabstractors have on the facility results is called the coder effect.

The coder effect has been accounted for many of the results presented in sections 9 to 11. For details on how the adjustments for coder effect were done, refer to section 7.4.1.1.

8.6 Case Mix Effect

The unique distribution of patients treated by a facility impacts the findings of the reabstraction study results. Data collected from certain subsets of the patient population were reabstracted with high agreement across all facilities. Hence, facility specific discrepancy rates for several data elements are highly related to the types of patients treated by these facilities, which can disguise real differences in coding quality.

Case mix effect refers to any differences in observations that are explained by variations in the patient population. Most of the estimates in section 10 and 11 have accounted for case mix effect. For details on how this analysis was done, refer to section 7.4.1.2.

9 Presentation of Findings

Sections 10 to 12 present the findings of the Case Costing study data. The following highlights the information that is available in each of these sections.

Section 10 – Primary Dataset Results

- For each studied data element, findings between the original and reabstracted data (primary dataset) are presented.
- Analyses by medical condition are provided for diagnosis code, diagnosis type, and the most responsible diagnosis.
- Analyses by major clinical category are provided for some of the CIHI grouping methodology output variables.
- *Adjustments for coder effect are included when the logistic regression could be applied to the data. Some results were analyzed for coder effect, and others were not.*
- *Adjustments related to case mix effect are not included in this section.*

Section 11 – Facility Specific Results

- For diagnosis data and the CIHI grouping methodology output variables, findings are graphically presented for individual facilities.
- *Adjustments for coder effect and case mix effect were applied, when possible, to the facility specific results to best reflect the variation in coding.*

Section 12 – Inter-rater Dataset Results

- For each of the studied data elements, findings between the two reabstractors' data (inter-rater dataset) are presented.
- Comparisons between the inter-rater and primary datasets are presented. *For these comparisons, adjustments for case mix effect were applied when possible. In addition, adjustments for coder effect were used, when significant.* Coder effect applied to the inter-rater data (reabstractor pair effect) accounts for the variation in the agreement rates for the pairs of reabstractors.

Coder effect accounts for the variation in the coding between reabstractors, and the effect this has on the findings. ***Case mix effect*** accounts for variation in the types of conditions presented to each of the facilities, and the effect this has on the findings.

Refer to sections 7.4.1.1 and 8.5 for more information on coder effect.
Refer to sections 7.4.1.2 and 8.6 for more information on case mix effect.

10 Primary Dataset Results

This section highlights the findings from the first two years of implementation of ICD-10-CA and CCI for Ontario's case-costing facilities for demographic, administrative, and clinical data.

This section analyzes the aggregate results from the primary dataset, as well as findings associated with selected medical conditions. Due to the breadth of coding within ICD-10-CA and CCI, discrepancy rates for the entire primary dataset can conceal information on subsets of codes that are problematic. To address this, 16 medical conditions were selected. Some were chosen as they indicated higher overall coding discrepancies, and others were selected as they were of particular interest. These medical conditions are analyzed further and are defined below.

- ***Blood disorders*** (D50-D53, D60-D69)
- ***Circulatory conditions*** (I10-I15, I20-I25, I30-I52, I95)
- ***Diabetes mellitus*** (E10-E14)
- ***Digestive conditions*** (K80-K87)
- ***Genitourinary conditions*** (N17-N23, N39, N80-N98)
- ***Injuries*** (S50-S59, S70-S79)
- ***Mental health*** (F00-F99)
- ***Metabolic disorders*** (E70-E88, E90)
- ***Neonates*** (P00-P99, Z38)
- ***Neoplasms*** (C30-C39, C76-C80)
- ***Obstetrics*** (O00-O99, Z37, Z39)
- ***Other factors*** influencing health status, known as Z-codes (Z51, Z70-Z76)
- ***Post-procedural/complications*** (E89, G97, H59, H95, J95, I97, K91, M96, N99, T80-T88)
- ***Respiratory conditions*** (J10-J18, J44, J90, J98)
- ***Symptoms*** (R00-R19, R40-R46)
- ***Surgical*** (Section 1 of CCI Therapeutic Interventions and selected Diagnostic Interventions from sections 2 and 3 which affect CMG assignment)

10.1 Non-medical Data Elements

Analysis was performed on the accuracy of coding for a subset of non-medical data elements submitted to the DAD. In general, discrepancy rates are minimal. Table 10.1.1 shows the findings for both fiscal years of data. Data elements "Admission Category", "Discharge Disposition", and "Institution To Type" have noted discrepancies in FY 2002–2003 with some improvement in FY 2003–2004. "Institution From Type" has a noticeable discrepancy rate present for both fiscal years. Discrepancies are also noted for "Institution From" and "Institution To".

Table 10.1.1: Discrepancy Counts and Rates for Non-medical Data Elements in the Primary Dataset, *Unadjusted for Coder Effect*

Data Element	DAD FY 2002–2003		DAD FY 2003–2004	
	Count	%	Count	%
Health Care Number	1,489	0.6 ± 0.5	579	0.2 ± 0.3
Gender	0	0.0 ± 0.0	0	0.0 ± 0.0
Birth Date	181	0.1 ± 0.1	2	0.0 ± 0.0
Birth Date is Estimated	0	0.0 ± 0.0	0	0.0 ± 0.0
Admission Category	7,207	2.7 ± 0.9	1,853	0.7 ± 0.4
Admission Date	84	0.0 ± 0.0	34	0.0 ± 0.0
Discharge Disposition	4,807	1.8 ± 0.6	1,671	0.6 ± 0.2
Discharge Date	71	0.0 ± 0.0	1,096	0.4 ± 0.7
Alternate Level of Care Days	1,712	0.6 ± 0.2	947	0.4 ± 0.1
Total Length of Stay	146	0.1 ± 0.0	128	0.0 ± 0.0
Acute Length of Stay	1,858	0.7 ± 0.2	1,310	0.5 ± 0.2
Institution From	5,164	1.9 ± 0.5	2,947	1.1 ± 0.6
Institution From Type	4,741	1.8 ± 0.5	4,365	1.6 ± 0.8
Institution To	5,407	2.0 ± 0.6	3,333	1.3 ± 0.6
Institution To Type	5,270	2.0 ± 0.6	2,108	0.8 ± 0.2
Weight	264	0.1 ± 0.1	514	0.2 ± 0.2

Source: CIHI 2005

Note: The denominator for percentages is 266,790 in FY 2002–2003 and 265,372 in FY 2003–2004.
Note: Variations in the version code of Health Care Number are not flagged as a discrepancy.

Both "standards/codebook/manual" and "chart documentation" accounted for most non-medical data element discrepancies. Reabstractors also assigned the reason of "optional/not wrong" for certain scenarios involving patient transfers (e.g. "Institution To Type"), as per training instructions. The improvement in discrepancy rates in FY 2003–2004 is attributed to higher compliance with the "standards/ codebook/manual", as very few discrepancies in FY 2003–2004 were assigned this discrepancy reason.

10.1.1 Coder Effect on Non-medical Data Elements

Due to the high agreement rate observed across all the non-medical data elements, the logistic regression model was unstable due to the insufficient sample of records containing discrepancies. For this reason, the coder effect could not adequately be assessed for these data elements.

10.1.2 Coding Specific Observations

Admission Category

Confusion in the interpretation of "L - elective" and "U - urgent/emergent" resulted in discrepancies for obstetric cases in FY 2002–2003, and to a lesser extent in FY 2003–2004. Other discrepancies were associated with admissions for cardiac surgery, hip replacement and chemotherapy. Some cases were originally assigned "urgent/emergent" but reabstracted as "elective". Discrepancies are due to information missed in chart, and misinterpretation of admission category definitions.

Discharge Disposition
The most common discrepancy was between "04 - discharged to home or a home setting with support services" and "05 - discharged home, no support service required". The differences are due to chart documentation.

Institution From Type
The highest proportion of transfers missed in the original data was from "1 - acute care facilities". Improvements were observed in FY 2003–2004.

Institution To Type
The highest proportion of transfers missed in the original data was to "8 - home care facilities". This is due to either information missed in discharge by the original coder, or guidelines related to home care transfers specific to a facility.

10.1.3 Improving the Quality of Non-medical Data Elements

The dominating source of error for non-medical data elements is due to issues related to chart documentation. Reabstractors only flagged "standards/codebook/manual" as a source of error for half of the "Admission Category" discrepancies in FY 2002–2003 and a third of the "Discharge Disposition" discrepancies in FY 2002–2003. In FY 2003–2004, most discrepancies were attributed to issues related to chart documentation.

Admission Category, Discharge Disposition, Institution To, and Institution From data elements need improved chart documentation.

10.2 Selection of Intervention Code

Analysis was performed on the code selection for mandatory interventions in both the original and reabstracted data. "Mandatory" interventions refer to those interventions identified in the Case Costing study as mandatory for the purposes of reabstraction. Certain interventions were reabstracted because they impact the CIHI grouping methodology outputs such as resource intensity weight and expected length of stay. Refer to section 6.2.2 for details.

Code comparisons for interventions can result in one of the following.

- **Exact Code Match**, all 10 characters of the intervention code match
- **Rubric Match Only**, where the *rubric* is the first 5 characters of the intervention code
- **Type of Intervention Only**, where *type* describes the intervention type and is the 4th and 5th character of the intervention code
- **Part of Anatomy Only**, where *anatomy* describes the same region or area of focus and is the 2nd and 3rd character of the intervention code
- **Different**, if the code comparison is none of the above

There is no statistically significant difference in the quality of code selection between fiscal years, as illustrated in Table 10.2.1. Exact agreement of intervention codes is about 88% for both fiscal years. When comparing intervention codes up to and including the rubric level, the agreement rate increases to more than 93%.

Table 10.2.1: Comparison of Codes for Mandatory Interventions in the Primary Dataset, *Unadjusted for Coder Effect*

	DAD FY 2002–2003		DAD FY 2003–2004	
	Count	**%**	**Count**	**%**
Mandatory in Original and Reabstracted	**236,606**		**274,672**	
Exact Match	208,032	87.9 ± 1.6	241,667	88.0 ± 1.7
Rubric Match Only	14,103	6.0 ± 1.2	15,298	5.6 ± 1.0
Type of Intervention Only	6,498	2.7 ± 0.6	9,455	3.4 ± 0.9
Part of Anatomy Only	6,453	2.7 ± 0.8	5,664	2.1 ± 0.7
Different	1,521	0.6 ± 0.2	2,588	0.9 ± 0.6

Source: CIHI 2005

Most discrepancies were cited by the reabstractor as non-compliance to the "standards/codebook/manual" (6.4% and 4.0%) or a differing interpretation in the "chart documentation" (5.6% and 7.9%), as illustrated in Table 10.2.2.

Table 10.2.2: Reasons for Discrepancies in Intervention Codes in the Primary Dataset, *Unadjusted for Coder Effect*

	DAD FY 2002–2003		DAD FY 2003–2004	
	Count	**%**	**Count**	**%**
Total Discrepancies	**28,574**	**12.1**	**33,005**	**12.0**
Standards/Codebook/Manual	15,221	6.4	10,998	4.0
Chart Documentation	13,333	5.6	21,757	7.9
Acceptable Difference	20	0.0	250	0.1

Source: CIHI 2005

Note: "Acceptable Difference" is the "Optional/Not Wrong" reason code.
Note: The denominator for percentages is 236,606 in FY 2002–2003 and 274,672 in FY 2003–2004.

Intervention codes have dates and attributes associated with them. This study reviewed all the intervention dates, and those attributes that are mandatory for DAD submission. The following highlights the findings for these related data elements[11].

- Discrepancy rates for intervention dates are less than 2% for each fiscal year.
- Most interventions do not have mandatory attributes (93.6% for FY 2002–2003 and 85.9% in FY 2003–2004). When mandatory attributes are present, a slight number of discrepancies for the location and extent attributes are statistically significant for FY 2003–2004 only.

[11] These findings for intervention dates and attributes are not adjusted for coder effect.

10.2.1 Coder Effect on Selection of Intervention Code

In FY 2003–2004 only, there was a statistically significant change in the match rate for intervention code selection after adjustments were made for coder effect. For that fiscal year, the match rate changed from an unadjusted 88% to an adjusted 91%. When comparing results between the two fiscal years, there is a significant increase in the agreement rate between fiscal years after the adjustments.

Table 10.2.1.1: Coder Effect on the Agreement Rates for Mandatory Intervention Codes in the Primary Dataset

	DAD FY 2002–2003		DAD FY 2003–2004	
	%	Confidence Interval	%	Confidence Interval
Mandatory in Original and Reabstracted				
Adjusted for Coder Effect	86.1	[84.1, 88.0]	91.4	[90.1, 92.5]
Unadjusted	87.9	[86.4, 89.5]	88.0	[86.3, 89.7]

Source: CIHI 2005

Note: The denominator for percentages is 236,606 in FY 2002–2003 and 274,672 in FY 2003–2004.

10.2.2 Results for Selected Medical Conditions

Of the 16 medical conditions that were analyzed separately from the entire primary dataset, one applied to interventions. These were surgical interventions. No significant difference in this subset of interventions was observed.

10.2.3 Coding Specific Observations

Chart documentation often does not meet the level of detail required in the CCI classification. Code differences are partly the consequence of incomplete or conflicting chart documentation, differences in interpretation, information missed or specificity of condition not supported. A separate issue is non-compliance with codebook standards that are embedded in CCI. The following are some specific observations.

Computerized tomography (CT) and magnetic resonance imaging (MRI)
In FY 2003–2004, anatomy sites of “CT head” were reabstracted to “CT brain” as well as “MRI head” reabstracted to “MRI brain”. Also noted was “CT lung” reabstracted to “thoracic cavity”. Variations in coding are due to hospital policies, availability of diagnostic imaging reports at time of original coding, and chart interpretation.

Fixation femur
“Fixation femur” was often reabstracted to “fixation hip joint”. Differences are a consequence of both chart interpretation and codebook standards not followed.

Insertion intravenous device (vascular access for hemodialysis)
Code discrepancies were observed for short-term peripheral access versus long-term infusion of drugs and fluids. Codes of 1.KX.53.^^ "implantation of internal device, vein" were reabstracted to 1.IS.53.^^ "implantation of internal device, vena cava". Differences are due to both chart interpretation and codebook standards not being followed. There was improvement in FY 2003–2004.

There were also discrepancies associated with vascular access interventions. Codes of "creation of a hemodialysis fistula" were reabstracted to "implantation of internal device, vena cava for short-term hemodialysis". The specificity in CCI with hemodialysis codes coupled with chart documentation issues makes this area more subject to coding variation than others.

Inspection stomach vs. inspection small intestine
Discrepancies were noted with endoscopy coding in both fiscal years. There were instances where "inspection, stomach" was reabstracted to "inspection small intestine". Differences are attributed to chart documentation and non-compliance to coding standards.

Inspection vs. biopsy
There were instances of "inspection" being reabstracted as "biopsy". Affected anatomy sites included bronchus and small intestine.

Pacemaker insertion
In FY 2002–2003, "pacemaker insertion" was occasionally reabstracted to "replacement of pacemaker/defibrillator leads" resulting in different rubrics being assigned. Differences are due to chart interpretation and standards not being followed.

Partial excision vs. repair
In FY 2003–2004, "partial excision abdominal aorta" was reabstracted as "repair". Discrepancies are due to chart interpretation and lack of knowledge of intervention definitions.

Tube insertion, stomach
There were a few instances when "insertion, drainage tube" was reabstracted to "insertion, permanent endoscopic gastrostomy [PEG] tube". Discrepancy rates improved in FY 2003–2004.

10.2.4 Improving the Quality of Intervention Code Selection

About half of the discrepancies in intervention codes are attributed to chart documentation, indicating incomplete or conflicting information, differences in interpretation, information missed or specificity of condition not supported. Improvements to chart documentation are a necessary requirement to facilitate coding accuracy. Physicians require continuing education on documentation requirements to meet the level of specificity in CCI.

Other issues identified were the result of non-compliance to specific coding standards. Health information professionals are encouraged to review the coding conventions in CCI, particularly include and exclude notes. Also, intervention definitions are located in Appendix A of CCI. Full definitions of intervention types are also available in the ICD-10-CA and CCI self-learning product titled "Introduction to ICD-10-CA and CCI".

CIHI has undertaken various initiatives to facilitate the quality of intervention codes.

- In FY 2005–2006, there is an edit that will not permit "dehydration" to be submitted as the most responsible diagnosis with the intervention "gastroenteritis".
- CIHI will be revising various standards in FY 2006–2007 for improved clarification including: diabetes mellitus, and post-procedural conditions and complications.

10.3 Selection of Diagnosis Code

Analysis of the selection of diagnosis code was performed on a subset of conditions identified as significant (i.e. type (M) (1) (2) (W) (X) (Y)) in both the original and reabstracted data. Diagnosis code comparisons can result in one of the following:

- **Exact Code Match**, where *exact code* is all 6 characters of the diagnosis code
- **Category Match Only**, where *category* is the first 3 characters of the diagnosis code
- **Block Match Only**, where *block* is a grouping of categories
- **Chapter Match Only**, where *chapter* relates to ICD-10-CA chapters
- **Different Chapter**, where *chapter* relates to ICD-10-CA chapters

There is no significant difference in the quality of code selection between fiscal years, as illustrated in Table 10.3.1. Exact agreement of diagnosis codes is about 85% for both fiscal years.

Table 10.3.1: Diagnosis Code Comparisons for Conditions Identified as Significant in the Original and Reabstracted Data, *Unadjusted for Coder Effect*

	DAD FY 2002–2003		DAD FY 2003–2004	
	Count	%	Count	%
Significant in Original and Reabstracted	**486,806**		**455,251**	
Exact Code Match	411,746	84.6 ± 1.3	387,092	85.0 ± 1.4
Category Match Only	35,889	7.4 ± 0.8	35,434	7.8 ± 1.0
Block Match Only	17,726	3.6 ± 0.6	14,025	3.1 ± 0.7
Chapter Match Only	12,888	2.6 ± 0.6	9,926	2.2 ± 0.5
Different Chapter	8,556	1.8 ± 0.5	8,775	1.9 ± 0.6

Source: CIHI 2005

Reabstractors identified a reason for a discrepancy in diagnosis code when one existed between the original and reabstracted data. Table 10.3.2 shows that most of the discrepancies are attributed to "chart documentation" (9.6% and 8.6%), whereas "standards/codebook/manual" account for 5.7% and 6.2%.

Table 10.3.2: Reasons for Discrepancies in the Selection of Diagnosis Code in the Primary Dataset, *Unadjusted for Coder Effect*

	DAD FY 2002–2003		DAD FY 2003–2004	
	Count	%	Count	%
Total Discrepancies	**75,060**	**15.4**	**68,159**	**15.0**
Standards/Codebook/Manual	27,527	5.7	28,139	6.2
Chart Documentation	46,920	9.6	39,182	8.6
Acceptable Difference	612	0.1	838	0.2

Source: CIHI 2005

Note: "Acceptable Difference" is the "Optional/Not Wrong" reason code.
Note: The denominator for percentages is 486,806 in FY 2002–2003 and 455,251 in FY 2003–2004.

10.3.1 Coder Effect on Selection of Diagnosis Code

There was no statistically significant difference in the match rate for diagnosis code selection after adjustments were made for coder effect. For both fiscal years, the match rates changed from an unadjusted 85% to an adjusted 86%.

Table 10.3.1.1: Coder Effect on the Agreement Rates of Diagnosis Code for Conditions Identified as Significant in the Original and Reabstracted Data

	DAD FY 2002–2003		DAD FY 2003–2004	
	%	Confidence Interval	%	Confidence Interval
Significant in Original and Reabstracted				
Adjusted for Coder Effect	85.8	[84.6, 86.9]	85.5	[84.3, 86.7]
Unadjusted	84.6	[83.3, 85.9]	85.0	[83.6, 86.4]

Source: CIHI 2005

Note: The denominator for percentages is 486,806 in FY 2002–2003 and 455,251 in FY 2003–2004.

10.3.2 Results for Selected Medical Conditions

Of the 16 selected medical conditions described on page 39, 15 are related to diagnoses and one is related to interventions. Discrepancies in diagnosis code selection for the 15 selected medical conditions are shown in Table 10.3.2.1. These results have not been adjusted for coder effect. ***Note that these conditions account for approximately two-thirds of all the conditions presented for the primary dataset.***

Table 10.3.2.1: Diagnosis Code Discrepancies for Selected Medical Conditions that are Significant in the Original and Reabstracted Data, *Unadjusted for Coder Effect*

Medical Condition	DAD FY 2002–2003			DAD FY 2003–2004		
	Total	Discrepancy		Total	Discrepancy	
		Count	%		Count	%
All Medical Conditions	**486,806**	**75,060**	**15.4 ± 1.3**	**455,251**	**68,159**	**15.0 ± 1.4**
Blood Disorders	11,871	3,790	31.9 ± 11.6	8,196	2,260	27.6 ± 8.5
Circulatory Conditions	61,261	7,913	12.9 ± 2.7	54,905	7,774	14.2 ± 3.4
Diabetes Mellitus	11,323	5,881	51.9 ± 9.0	8,791	4,044	46.0 ± 9.5
Digestive Conditions	6,076	1,692	27.8 ± 15.7	5,354	382	7.1 ± 4.7
Genitourinary Conditions	20,842	3,693	17.7 ± 10.0	16,245	2,617	16.1 ± 6.8
Injuries	4,673	584	12.5 ± 8.2	4,071	1,725	42.4 ± 21.5
Mental Health	18,175	3,156	17.4 ± 4.7	19,510	4,072	20.9 ± 8.0
Metabolic Disorders	9,885	425	4.3 ± 1.8	8,272	206	2.5 ± 1.3
Neonates	36,835	1,158	3.1 ± 2.0	38,911	1,429	3.7 ± 2.4
Neoplasms	9,631	1,539	16.0 ± 8.9	9,572	1,017	10.6 ± 6.5
Obstetrics	76,653	7,771	10.1 ± 3.3	75,648	6,295	8.3 ± 2.8
Other factors	11,010	459	4.2 ± 3.3	12,667	483	3.8 ± 1.8
Post-procedural Compl.	18,542	3,438	18.5 ± 5.6	18,904	3,954	20.9 ± 9.1
Respiratory Conditions	16,240	3,996	24.6 ± 5.1	17,128	3,989	23.3 ± 6.3
Symptoms	11,635	2,312	19.9 ± 13.8	13,795	837	6.1 ± 3.1

Source: CIHI 2005

Note: Denominators for percentages are all original conditions belonging to a medical condition, listed under “Total”.
Note: Refer to page 39 for the listing of diagnosis codes that belong to each medical condition.

Key findings for selection of diagnosis codes, by selected medical condition

- ***Metabolic disorders, neonates, obstetrics, other factors, digestive conditions (FY 2003–2004 only), and symptoms (FY 2003–2004 only)*** have significantly lower discrepancy rates than the 15.4% and 15.0% discrepancy rates observed for all conditions.
- ***Blood disorders, diabetes mellitus, respiratory conditions, and injuries (FY 2003–2004 only)*** have significantly higher discrepancy rates than the rates observed for all conditions.

10.3.3 Coding Specific Observations

Blood disorders

Discrepancies were noted for anemia (D50-D64). Many discrepancies were found for original codes of D50.0 “iron deficiency anemia secondary to blood loss (chronic)” and D64.9 “anemia, unspecified”. Another common discrepancy was between D63.0 “anemia in neoplastic disease” and D64.9 “anemia, unspecified”. When anemia is present in neoplastic disease, the codebook leads to D63.0. These discrepancies are due to difficulties with chart interpretation, and also not following codebook directives.

Diabetes mellitus
Discrepancies for diabetes mellitus (E10-E14) are mainly attributed to the sixth digit, indicating level of control. Correct assignment of the sixth digit is dependant on comprehensive physician documentation and laboratory results indicating blood glucose levels. Variation was also noted for diabetes mellitus with or without complication (e.g. E11.9 "type 2 diabetes mellitus without (mention of) complication", versus E11.2 "type 2 diabetes mellitus with renal complication").

Discrepancies are attributed to non-compliance to specific coding standards, and incomplete or conflicting information in the chart documentation.

Fractures
Discrepancies were noted with coding fractures. For instance, original codes of "fracture neck of femur (traumatic)" were reabstracted as "osteoporosis with pathological fracture". Variation in coding is due to interpretation of chart documentation and non-compliance to specific coding standards.

"Fractured neck of femur" was also reabstracted as "complications of internal orthopedic prosthetic devices". Coding standards classify complications of devices as mechanical complications when involving fractured prosthetic devices. The CIHI online Coding Query Database provides clarification on traumatic fractures versus mechanical complications of devices.

Post-procedural conditions and complications
Original codes for a specific condition were often reabstracted as a post-procedural code or complication. Some examples are:

- Renal failure vs. post-procedural renal failure
- Pulmonary embolism vs. post-operative embolism
- Arterial embolism vs. post-procedural/complication
- Post hemorrhagic anemia vs. post-operative hematoma

The code describing the condition should only be captured as a supplemental code for further specificity, as noted in the coding standards. Differences in code selection are due to a combination of chart documentation and not following coding standards or codebook look-up.

Respiratory conditions
Original codes of "pulmonary edema" were reabstracted to "congestive heart failure, pneumonia", "adult respiratory distress syndrome", or some other condition. These reabstracted conditions are likely the underlying cause of pulmonary edema. Discrepancies are due to the interpretation of chart documentation.

Symptoms vs. underlying conditions
A high proportion of original codes describing a symptom were reabstracted as the underlying condition. Some examples are:

- Chest pain vs. angina/acute myocardial infarction

- Convulsions vs. epilepsy
- Urinary retention vs. benign prostatic hypertrophy

Coding standards and education sessions have stated that symptoms should not be coded when the underlying condition is known, but may be coded as secondary condition at a facility's discretion. Discrepancies are due to a combination of interpretation of and incomplete chart documentation.

10.3.4 Improving the Quality of Diagnosis Code Selection

More than half of the discrepancies with diagnosis coding were attributed to issues related to chart documentation, indicating incomplete or conflicting information, differences in interpretation, information missed or specificity of condition not supported. Improvements to chart documentation are a necessary requirement to facilitate coding accuracy. Health information professionals, health records department managers and directors, and senior hospital staff must continually take the initiative to ensure physicians are educated about documentation issues at their facilities.

The remaining discrepancies with diagnosis coding were attributed to the non-compliance to "standards/codebook/manual". Awareness about coding standards is essential to ensure consistent coding practice. Health information professionals are encouraged to review the Canadian Coding Standards for ICD-10-CA/CCI and attend educational workshops whenever possible.

Since FY 2003–2004, CIHI has provided clarification on several topics related to coding standards. The following issues have been covered in the Canadian Coding Standards and Diagnosis Typing for DAD:

- Diabetes mellitus (selecting the appropriate 6^{th} digit, diabetes without mention of complication)
- Post-procedural conditions and complications
- Underlying symptoms or conditions

CIHI has undertaken various initiatives to facilitate coding quality. In the 2006 version of ICD-10-CA and CCI, the 6^{th} digit will be removed from diabetes mellitus codes. Coding standards will be completely revised to supplement the enhancements made to the classification and related new education sessions will be offered for further clarification.

10.4 Selection of Most Responsible Diagnosis

The most responsible diagnosis (MRDx) is the one condition that is accountable for the greatest portion of the length of stay or greatest use of resources. Each discharge must be assigned one and only one MRDx. Among other grouper outputs, the assignment of MRDx determines which major clinical category a discharge gets assigned to by the CIHI grouping methodology.

Unlike the analysis of diagnosis code selection in section 10.3, analysis of code selection for MRDx is done irrespective of whether the original and reabstracted MRDx are describing the same condition. This can be done because only one MRDx code is present in the original data for a discharge, and one MRDx is present in the reabstracted data for that same discharge.

There is no significant difference in the quality of code selection of the MRDx code between fiscal years, as illustrated in Table 10.4.1. The proportion of discharges where the original coder and reabstractor selected the exact same code as the most responsible diagnosis is about 73% for both fiscal years. Several discharges had MRDx code comparisons resulting in "Chapter Match Only" and "Different Chapter", illustrating that different conditions were assigned the most responsible diagnosis upon reabstraction.

Table 10.4.1: Code Comparisons for Most Responsible Diagnoses in the Original and Reabstracted Data, *Unadjusted for Coder Effect*

	DAD FY 2002–2003		DAD FY 2003–2004	
	Count	**%**	**Count**	**%**
Total Most Responsible Diagnoses	**266,790**		**265,372**	
Exact Code Match	196,461	73.6 ± 2.3	194,454	73.3 ± 2.5
Category Match Only	17,828	6.7 ± 1.1	20,555	7.7 ± 1.4
Block Match Only	16,611	6.2 ± 1.3	12,117	4.6 ± 1.1
Chapter Match Only	15,292	5.7 ± 1.2	12,939	4.9 ± 1.1
Different Chapter	20,598	7.7 ± 1.4	25,307	9.5 ± 1.8

Source: CIHI 2005

The 73% agreement rates in the above table show how often the reabstractor agreed in the code selection and typing for codes originally assigned the MRDx. These results are similar to the agreement rates of 77% for diagnosis code and type for *all significant conditions*, as shown in Table 10.4.2. The margins of error associated with these estimates indicate the two percentages are not significantly different.

Table 10.4.2: Code and Type Comparisons for Conditions that are Significant in Both the Original and Reabstracted Data, *Unadjusted for Coder Effect*

	DAD FY 2002–2003		DAD FY 2003–2004	
	Count	**%**	**Count**	**%**
Significant in Original and Reabstracted	**486,806**		**455,251**	
Code Match and Type Match	373,107	76.6 ± 1.6	351,262	77.2 ± 1.7
Code Match Only	38,639	7.9 ± 1.0	35,830	7.9 ± 1.2
Type Match Only	65,407	13.4 ± 1.2	62,251	13.7 ± 1.3
Neither	9,653	2.0 ± 0.4	5,908	1.3 ± 0.3

Source: CIHI 2005

10.4.1 Coder Effect on Selection of Most Responsible Diagnosis

There was no statistically significant change in the match rate for code selection of the most responsible diagnosis after adjustments were made for coder effect. These results are shown in Table 10.4.1.1. In FY 2002–2003, the match rate changed from an unadjusted 74% to an adjusted 76%. In FY 2003–2004, the match rate changed from an unadjusted 73% to an adjusted 74%.

Table 10.4.1.1: Coder Effect on the Agreement Rates for Code Selected as the Most Responsible Diagnosis

	DAD FY 2002–2003		DAD FY 2003–2004	
	%	Confidence Interval	%	Confidence Interval
Total Most Responsible Diagnoses				
Adjusted for Coder Effect	75.5	[72.9, 77.9]	73.6	[70.9, 76.1]
Unadjusted	73.6	[71.4, 75.9]	73.3	[70.8, 75.7]

Source: CIHI 2005

Note: The denominator for percentages is 266,790 in FY 2002–2003 and 265,372 in FY 2003–2004.

10.4.2 Results for Selected Medical Conditions

Original MRDx codes that belonged to one of the selected medical conditions were compared against the reabstracted code selected as the MRDx. The results for this analysis are presented in Table 10.4.2.1. These results have not been adjusted for coder effect. ***Note that discharges with original MRDx codes that are assigned to one of the selected medical conditions account for approximately two-thirds of all discharges.***

Table 10.4.2.1: Discrepancies in Code Selected as the Most Responsible Diagnosis for Selected Medical Conditions, *Unadjusted for Coder Effect*

Medical Condition	DAD FY 2002–2003			DAD FY 2003–2004		
	Total	Discrepancy		Total	Discrepancy	
		Count	%		Count	%
All Medical Conditions	**266,790**	**70,329**	**26.4 ± 2.3**	**265,372**	**70,918**	**26.7 ± 2.5**
Blood Disorders	880	459	52.1 ± 31.0	532	170	32.1 ± 37.6
Circulatory Conditions	30,336	8,134	26.8 ± 5.5	25,835	6,111	23.7 ± 5.6
Diabetes Mellitus	2,934	982	33.5 ± 21.4	1,180	451	38.2 ± 12.2
Digestive Conditions	5,078	1,905	37.5 ± 19.3	3,756	860	22.9 ± 17.9
Genitourinary Conditions	8,629	4,143	48.0 ± 19.3	6,322	2,230	35.3 ± 15.0
Injuries	4,075	496	12.2 ± 9.2	3,839	1,698	44.2 ± 22.4
Mental Health	10,779	2,550	23.7 ± 7.0	11,505	2,653	23.1 ± 9.9
Metabolic Disorders	2,279	1,016	44.6 ± 29.5	1,518	270	17.8 ± 16.9
Neonates	31,249	2,650	8.5 ± 5.6	30,985	4,836	15.6 ± 7.7
Neoplasms	4,507	1,645	36.5 ± 9.6	6,953	2,823	40.6 ± 13.3
Obstetrics	43,322	11,174	25.8 ± 6.3	46,072	11,125	24.1 ± 5.5
Other factors	3,642	993	27.3 ± 16.8	4,049	628	15.5 ± 12.9
Post-procedural Compl.	5,838	1,790	30.7 ± 12.3	9,548	5,055	52.9 ± 16.9
Respiratory Conditions	7,632	2,883	37.8 ± 10.1	8,948	2,270	25.4 ± 9.1
Symptoms	5,912	2,457	41.6 ± 24.8	9,511	1,168	12.3 ± 9.1

Source: CIHI 2005

Note: Denominators for percentages are all discharges with original MRDx codes belonging to a selected medical condition, listed under "Total"
Note: Refer to page 39 for the listing of diagnosis codes that belong to each medical condition.

Key findings for the code selection of most responsible diagnoses

- ***Injuries (FY 2002–2003), neonates, and symptoms (FY 2003–2004)*** have significantly lower discrepancy rates than what was observed for all conditions.
- ***Genitourinary conditions (FY 2002–2003) and post-procedural complications (FY 2003–2004)*** have significantly higher discrepancy rates than what was observed for all conditions.

10.4.3 Coding Specific Observations

Further analysis of the coding for the most responsible diagnosis was done by comparing the original MRDx code to the reabstracted MRDx code. The following are the highlights of this analysis.

Admission for chemotherapy and/or radiation therapy
Discharges with original MRDx codes of "malignant neoplasm" had reabstracted MRDx codes of "admission for chemotherapy" or "admission for radiation therapy". Coding standards state that Z51 is to be assigned the MRDx when the patient is specifically admitted for treatment of the malignancy. The neoplasm codes (active and historical malignancies) may be coded as a diagnosis type (3).

Angina

Discharges with original MRDx codes of "angina" had reabstracted MRDx codes of other cardiac conditions including "unstable angina", "chronic ischemic heart disease", and "acute myocardial infarction". Here, the original coder regarded angina as a symptom of the underlying condition such as coronary atherosclerosis. Coding standards state that if treatment was directed towards the atherosclerosis such as angioplasty or coronary artery bypass graft, this should be the MRDx. Angina is to be coded as an additional condition. Coding variation is attributed to issues related to chart documentation.

Cellulitis

There are two observations for discharges that had cellulites originally coded as the MRDx. First, reabstractors instead coded "complication infection" (T80-T88) as the MRDx. This is the result of different interpretation of chart information. Second, reabstractors instead coded "open wound" as the MRDx. Here, coding standards state "if the course of treatment only involves oral antibiotics, cellulitis is presumed to be a co-morbid condition while the wound considered to be the MRDx".

Chronic obstructive pulmonary disease (COPD)/ pneumonia

For discharges with original MRDx codes of "pneumonia", reabstractors instead selected COPD as the MRDx. Coding standards state that COPD must always be sequenced as the MRDx when present with pneumonia.

Congestive heart failure

Particularly in FY 2002–2003, discharges with original MRDx codes of chronic heart failure were reabstracted to other 'conditions' including post-procedural and complication codes. Discrepancies are a result of non-compliance to coding standards.

Diabetes mellitus with renal failure

For discharges with original MRDx codes of "renal failure" (N17, N18), reabstractors instead selected diabetes mellitus as the MRDx. Coding standards relating to diabetes when complications are present state that renal failure would be coded as a supplemental condition for additional specificity.

Diabetic ulcers

Discharges with original MRDx codes of "leg ulcer" and "decubitus ulcer" were reabstracted with MRDx codes of the respective diabetes mellitus code with complication. This is due to issues related to chart interpretation and non-compliance with codebook directives that require the diabetes code sequenced first, followed by ulcer as a supplemental code for specificity.

Gastroenteritis with dehydration

Discharges with original MRDx codes of "dehydration" had reabstracted MRDx codes of "gastroenteritis". The coding standard on gastroenteritis states "in admissions for treatment of gastroenteritis and dehydration, sequence gastroenteritis as the MRDx."

Gastrointestinal (GI) bleeding
Original MRDx codes of "GI bleeding" were reabstracted as a type (3), the underlying cause. Here, reabstracted MRDx conditions such as "ulcer with hemorrhage" were used. Discrepancies are due to non-compliance to the coding standards, the interpretation of chart documentation, or incomplete documentation at the time of original coding. The discrepancy rate improved in FY 2003–2004.

Injuries
Discrepancies were identified in code selection for intracranial injury and fractures. Refer to section 9.2.1.2 for coding specific observations.

Palliative care and/or pain management
Discrepancies were identified with coding palliative care and other medical care including pain control as the MRDx. Original MRDx codes for these conditions were reabstracted as a Z51 code. The opposite discrepancy also occurred, where the original MRDx was a Z51 code but reabstracted as palliative care and pain control.

Currently, national standards for coding palliative care and pain control do not exist. The online CIHI Coding Query Database has addressed palliative care and pain management on a case specific basis.

Post-procedural conditions and complications
Discharges with original MRDx codes of complications (T codes) were reabstracted as MRDx codes of post-procedural codes. Coding standards exist for coding post-procedural conditions, and were significantly revised in 2003 to provide further clarification.

Neoplasms
Discrepancies in the original assignment of malignant neoplasms as MRDx are associated with anatomy site, which sometimes led to differences at the category or block level following reabstraction. On other occasions, it was differences between the nature of the neoplasm such as malignant versus carcinoma insitu. These discrepancies could be attributed to interpretation of chart documentation.

Discrepancies were also identified with the sequencing of the primary versus the secondary neoplasm as MRDx. Additional discrepancies were identified when neoplasm codes were originally assigned but reabstracted to other conditions, possibly complications of neoplastic disease. Particularly in FY 2003–2004, C80 "Malignant neoplasm without specification of site" was originally assigned the MRDx but reabstracted to specific neoplasm sites. Variation is due to issues related to chart documentation and non-compliance to specific neoplasm coding standards.

Obstetrics
The majority of discrepancies were for discharges that had Z37 "outcome of delivery" originally coded as the MRDx but the reabstracted MRDx was a condition from the obstetrics chapter. According to the standard Delivery in a Completely Normal Case, "when a code from O10-O99 is applicable, the outcome of delivery

should be added as a diagnosis type (3)". Other discrepancies were due to issues related to chart documentation.

Respiratory failure
Particularly in FY 2002–2003, discharges with original MRDx codes of "respiratory failure" were reabstracted with MRDx codes belonging to other conditions including "adult respiratory distress syndrome (ARDS)" and "post-procedural respiratory disorders". Differences are attributed to non-compliance to coding standards.

Respiratory failure with Chronic obstructive pulmonary disease (COPD)
Discharges with original MRDx codes of "respiratory failure" were reabstracted with MRDx codes of COPD. While respiratory failure may be a condition, it can also be considered a symptom to an underlying condition. Variation in coding is the result of chart interpretation and non-compliance to coding standards.

Septicemia
Discharges with original MRDx codes of "septicemia" were reabstracted with MRDx codes describing different conditions including: complication infection codes (T80-T88), urinary tract infection, and other diagnoses. Differences are due to interpretation of chart documentation. Particularly in FY 2003–2004, discharges with original MRDx codes of septicemia were reabstracted with MRDx of complication codes (e.g. T81.4 "infection following a procedure"). Coding standards state that when certain conditions are present, that code should be sequenced first, followed by septicemia as a type (3). Beyond coding standards, there are also codebook directives that lead to specific complication infection codes, when applicable.

Symptoms versus underlying conditions
There were occasions where symptom codes were originally assigned the MRDx but the reabstracted MRDx was a definitive diagnoses. Findings include:

- "Fever" vs. "neutropenia"
- "Cardiogenic shock" vs. "acute myocardial infarction"
- "Malaise and fatigue" (R53) vs. specific diagnoses

Coding standards state, "when a patient presents with a symptom or condition, and during that episode of care the underlying disease or disorder is identified, then the underlying disease or disorder is assigned as the MRDx". Another contributing factor to the variation is interpretation of chart documentation.

10.4.4 Improving the Quality of Selection for Most Responsible Diagnosis

Improvements to chart documentation are a necessary requirement to facilitate coding accuracy. Health information professionals, health records department managers and directors, and senior hospital staff must continually take the initiative to ensure high quality chart documentation.

Other issues identified were the result of non-compliance to specific coding standards for the selection of most responsible diagnosis. Health information professionals can refer to Canadian Coding Standards for ICD-10-CA and CCI for details. Health information professionals can also refer to the following education materials: the Canadian Coding Standards and Diagnosis Typing for DAD and ICD-10-CA and CCI Refresher.

10.5 Major Clinical Category

Major clinical category (MCC) is a methodology to aggregate patients to a body system or specific type of clinical problem. The methodology was developed by CIHI and is based on the MRDx. That is, each MRDx is assigned to one of the 25 MCC categories.

Original data and reabstracted data were both processed through the 2003 CIHI grouping methodology and the resultant MCC categories were compared to each other. Overall, the MCC values changed upon reabstraction for an estimated 5.6 ± 0.9% (FY 2002–2003) and 6.4 ± 1.4% (FY 2003–2004) of the discharges. The observed change between fiscal years is not significant.

10.5.1 Coder Effect on the Assignment of Major Clinical Category

There was no statistically significant change in the agreement rate for major clinical category after adjustments were made for coder effect. These results are shown in Table 10.5.1.1. In FY 2002–2003, the match rate changed from an unadjusted 94% to an adjusted 96%. In FY 2003–2004, the match rate changed from an unadjusted 94% to an adjusted 95%.

Table 10.5.1.1: Coder Effect on the Agreement Rates for Major Clinical Category

	DAD FY 2002–2003		DAD FY 2003–2004	
	%	Confidence Interval	%	Confidence Interval
Total Discharges				
Adjusted for Coder Effect	95.8	[91.8, 97.9]	95.2	[90.2, 97.7]
Unadjusted	94.4	[93.5, 95.3]	93.6	[92.1, 95.0]

Source: CIHI 2005

Note: The denominator for percentages is 266,790 in FY 2002–2003 and 265,372 in FY 2003–2004.

10.5.2 Results for Specific Major Clinical Categories

Table 10.5.2.1 shows changes in assignment for each MCC category upon reabstraction for each fiscal year. These results have not been adjusted for coder effect. ***Note that these results have large margins of error associated with them.*** This table contains three columns: "same", "outflow", and "inflow".

- The ***"same"*** column presents discharges that were grouped to the same MCC category with the original and reabstracted MRDx. Percentages are calculated out of all instances that were *originally* assigned to that particular MCC.

- The ***"outflow"*** column presents discharges with a reabstracted MRDx which groups to a different MCC category than the original MRDx. The results in this column are complementary to the "same" column; meaning that if the reabstracted MRDx did not group to the same MCC category as the original MRDx, then it grouped to a different MCC category. Hence, percentages for these two figures will add to 100%.
- The ***"inflow"*** column presents discharges with a reabstracted MRDx which groups to a particular MCC category but the original MRDx groups to a different MCC category. Results for this analysis are presented as estimated counts only.

Table 10.5.2.1: Major Clinical Category Assignments upon Reabstraction, *Unadjusted for Coder Effect*

Major Clinical Category	DAD FY 2002–2003					DAD FY 2003–2004				
	Same		Out-flow		In-flow	Same		Out-flow		In-flow
	Count	%	Count	%	Count	Count	%	Count	%	Count
1 - Nervous System	9,145	95	510	5	653	9,143	86	1,471	14	576
2 - Eye Diseases	2,653	94	169	6	20	830	100	0	0	743
3 - Ear Nose & Throat	3,016	94	178	6	283	4,914	92	400	8	263
4 - Respiratory	15,316	91	1,467	9	1,618	17,663	94	1,056	6	1,983
5 - Card & Vasc Diseases Circ System	38,151	96	1,649	4	912	36,337	97	1,153	3	2,434
6 - Digestive	18,787	93	1,454	7	1,800	19,975	97	708	3	906
7 - Hepatobiliary & Pancreas	7,145	94	457	6	293	5,493	92	464	8	315
8 - Musculoskeletal & Connect	13,351	91	1,358	9	914	14,022	96	639	4	1,015
9 - Skin Subcut & Breast	3,202	87	478	13	801	4,075	79	1,095	21	54
10 - Endocrine Nutrit & Metabolism	6,741	85	1,218	15	799	4,316	78	1,232	22	353
11 - Kidney & Urinary Tract	14,338	95	774	5	713	10,284	88	1,355	12	518
12 - Male Reproductive	493	99	6	1	0	60	96	3	4	0
13 - Female Reproductive	9,822	95	506	5	55	5,757	80	1,424	20	81
14 - Pregnancy & Childbirth	42,915	99	407	1	6	46,072	100	0	0	493
15 - Newborns & Other Neonates	32,261	100	0	0	0	31,443	100	0	0	0
16 - Bld, BldForming Org & Immun	1,977	82	446	18	337	2,390	96	92	4	243
17 - Lymph/Leukem & Neoplasm Unspec	3,514	87	504	13	1,378	4,108	89	526	11	2,647
18 - Multisys/Unspec Site Infection	2,900	73	1,058	27	897	1,962	70	840	30	1,133
19 - Mental Disease & Disorders	10,722	97	324	3	150	11,363	96	469	4	136
21 - Injuries Poison & Toxic Effect	2,264	71	907	29	1,671	5,009	64	2,772	36	766
22 - Burns	6	100	0	0	8	125	100	0	0	0
23 - Other reasons for hospitalization	2,331	79	636	21	965	3,343	80	854	20	1,292
24 - HIV Infections (AIDS)	137	66	69	34	8	192	69	85	31	144
25 - Significant Trauma	10,678	97	350	3	337	9,402	98	219	2	979
99 - Ungroupable Data	0	*	0	*	309	0	0	236	100	19

Source: CIHI 2005

Note: Original and reabstracted data were grouped using the 2003 CIHI grouping methodology.
Note: Estimates based off a sample size of less than 20 are placed in grey font.
Note: The denominator for percentages is the estimated number of discharges originally grouped to each MCC category, which can be derived by adding the counts under the “Same” and “Outflow” columns.

Key findings for major clinical category

- ***Highly populated MCC categories generally have high agreement rates.*** The following have agreement rates of 96% or better for both fiscal years: MCC5, MCC14, and MCC15.
- ***Less populated MCC categories generally have lower agreement rates.*** The following are moderately populated MCC categories that have agreement rates less than 90% for both years: MCC9, MCC10, MCC17, MCC18, MCC21, and MCC23.
- ***MCC categories with low "outflow" counts tend to have low "inflow" counts.*** The following have noticeable improvements in "Inflow" counts for FY 2003–2004: MCC6, MCC9, and MCC21.

10.5.3 Coding Specific Observations

Changes in MCC assignment are due to the change in assignment of the most responsible diagnosis (MRDx) upon reabstraction. Therefore, a review of the subset of discharges where the original MRDx did not match the reabstracted MRDx was performed as summarized below. Further details on MRDx analysis can be found in section 10.4.3.

- ***MCC9 Diseases & Disorders of Skin, Subcutaneous Tissue & Breast:*** changes in the assignment of MCC category are due in part to MRDx discrepancies of: cellulitis, diabetic ulcers, primary vs. secondary neoplasm site.
- ***MCC10 Endocrine Nutritional & Metabolic Diseases & Disorders:*** changes in the assignment of MCC category are due in part to MRDx discrepancies of "gastroenteritis with dehydration" and symptoms versus underlying condition. There were also a number of instances where two completely different conditions were assigned as MRDx due to differing chart interpretations.
- ***MCC18 Multisystemic or Unspecified Site Infections:*** changes in the assignment of MCC category are due in part to MRDx discrepancies of septicemia and symptom versus underlying condition.
- ***MCC21 Injury, Poisoning & Toxic Effects of Drugs:*** changes in the assignment of MCC category are due in part to MRDx discrepancies of post-procedural conditions versus complications.
- ***MCC23 Other Reasons for Hospitalization:*** changes in the assignment of MCC category are due in part to MRDx discrepancies of: symptom versus underlying condition, palliative care/other medical care.
- ***MCC17A Lymphoma or Leukemia:*** changes in the assignment of MCC category are due in part to MRDx discrepancies of: neoplasms, palliative care/other medical care.

10.5.4 Improving the Quality of Major Clinical Category

Assignment of major clinical category is directly affected by the code selection for most responsible diagnosis. As such, previous discussion in section 10.4.4 for improving the code selection of the most responsible diagnosis is applicable here.

10.6 Case Mix Group

Case mix group (CMG) is a methodology that aggregates patients into clusters based on clinical diagnoses, procedures and resource utilization. Acute care inpatients are assigned to one of the 478 CMG groups based on clinical and administrative data collected through the DAD.

Original data and reabstracted data were both processed through the 2003 CIHI grouping methodology and the resultant CMG groups were compared to each other. Overall, the CMG values changed upon reabstraction for 16% of the discharges for both fiscal years. Specifically, the estimated results are 16.1 ± 1.7% and 15.9 ± 2.1% for the respective fiscal years.

10.6.1 Coder Effect on the Assignment of Case Mix Group

There was no statistically significant change in the agreement rate for case mix group after adjustments were made for coder effect. These results are shown in Table 10.6.1.1. In FY 2002–2003, the match rate changed from an unadjusted 84% to an adjusted 86%. In FY 2003–2004, the match rate was 84% before and after adjustments.

Table 10.6.1.1: Coder Effect on the Agreement Rates for Case Mix Group

	DAD FY 2002–2003		DAD FY 2003–2004	
	%	Confidence Interval	%	Confidence Interval
Total Discharges				
Adjusted for Coder Effect	86.1	[83.8, 88.1]	84.2	[81.7, 86.5]
Unadjusted	83.9	[82.2, 85.6]	84.1	[82.1, 86.2]

Source: CIHI 2005

Note: The denominator for percentages is 266,790 in FY 2002–2003 and 265,372 in FY 2003–2004.

10.6.2 Results for Selected Case Mix Groups

Table 10.6.2.1 shows changes in assignment for a selection of CMG groups upon reabstraction for each fiscal year. The 18 CMG groups shown in this table represent the most commonly used CMG groups for both fiscal years of data[12]. These results have not been adjusted for coder effect. ***Note that these results have large margins of error associated with them.*** This table contains three columns: “same”, “outflow”, and “inflow”.

[12] That is, the most common case mix groups in the Case Costing dataset.

- The ***"same"*** column presents discharges that were grouped to the same CMG group with the original and reabstracted data. Percentages are calculated out of all instances that were *originally* assigned to that particular CMG group.
- The ***"outflow"*** column presents discharges with reabstracted data which groups to a different CMG group than the original data. The results in this column are complementary to the "same" column; meaning that if the reabstracted data did not group to the same CMG group as the original data, then it grouped to a different CMG group. Hence, percentages for these two figures will add to 100%.
- The ***"inflow"*** column presents discharges with reabstracted data which groups to a particular CMG group but the original data groups to a different CMG group. Results for this analysis are presented as estimated counts only.

Table 10.6.2.1: Selected Case Mix Group Assignments upon Reabstraction, *Unadjusted for Coder Effect*

Case Mix Group	DAD FY 2002–2003					DAD FY 2003–2004				
	Same		Out-flow		In-flow	Same		Out-flow		In-flow
	Count	%	Count	%	Count	Count	%	Count	%	Count
648 - Neonates Weight > 2500 gm (Normal Ne	16,999	100	0	0	503	16,904	100	0	0	1,754
611 - Vaginal Delivery	16,709	90	1,778	10	753	16,396	91	1,689	9	414
646 - Neonates Weight > 2500 gm w Caesarea	9,533	100	0	0	0	4,555	100	0	0	0
579 - Major Uterine & Adnexal Procedures w	6,128	99	55	1	13	3,466	100	4	0	125
609 - Vaginal Delivery w Complicating Diag	5,381	88	711	12	1,286	7,266	92	640	8	1,605
189 - PTCA w/o Cardiac Conditions	4,521	98	95	2	1,181	3,353	98	57	2	1,676
354 - Knee Replacement	3,476	99	47	1	21	2,523	100	0	0	36
294 - Esophagitis, Gastroenteritis & Misce	3,443	91	355	9	1,562	6,161	97	195	3	728
179 - Coronary Bypass w Heart Pump w/o Car	3,224	97	111	3	111	2,552	93	178	7	182
143 - Simple Pneumonia & Pleurisy	3,205	73	1,163	27	1,123	4,083	86	663	14	344
604 - Caesarean Delivery	3,103	84	609	16	1,621	3,092	77	918	23	874
253 - Major Intestinal & Rectal Procs	2,628	93	210	7	153	2,306	86	376	14	323
013 - Cerebrovascular except Transient Isc	2,600	95	151	5	282	2,591	97	78	3	220
242 - Chest Pain	2,538	66	1,290	34	46	5,109	91	476	9	208
222 - Heart Failure	2,524	69	1,114	31	369	3,120	94	209	6	454
603 - Repeat Caesarean Delivery	2,391	89	285	11	75	3,724	95	193	5	146
352 - Hip Replacement	2,321	94	152	6	105	2,333	99	13	1	19
602 - Caesarean Delivery w Complicating Di	2,019	56	1,571	44	578	3,141	79	857	21	918

Source: CIHI 2005

Note: The original and reabstracted data were grouped using the 2003 CIHI grouping methodology.
Note: Estimates based off a sample size of less than 20 are placed in grey font.
Note: The denominator for percentages is the estimated number of discharges originally grouped to each CMG group, which can be derived by adding the counts under the "Same" and "Outflow" columns.

Key findings for case mix group

- CMG groups representing caesarean deliveries (602, 604) have lower agreement upon reabstraction than the other common CMG groups. CMG group 602 improves in FY 2003–2004.
- CMG 222 "heart failure" has a notable improvement in agreement rates in FY 2003–2004.
- CMG 143 "simple pneumonia and pleurisy" has a lower agreement rate than other common CMG groups for both fiscal years.
- CMG 611 "vaginal delivery" has a 90% agreement in CMG upon reabstraction.

10.6.3 Improving the Quality of Case Mix Group

Diagnoses with a type of (M) (1) (2) (0) (W) (X) (Y) are used to determine the case mix group to which a discharge gets assigned. As such, discussion on how to improve diagnosis codes and types are needed to improve this grouper output variable.

10.7 Selection of Diagnosis Type

The assignment of a diagnosis type to a condition signifies the impact that the condition had on the patient's care. All conditions identified on the DAD must be assigned one of the following diagnosis types:

- **(M)**, the most responsible diagnosis
- **(1)**, a pre-admit co-morbidity
- **(2)**, a post-admit co-morbidity
- **(W)(X)(Y)**, a service transfer diagnosis
- **(3)**, a secondary diagnosis
- **(0)**, an optional diagnosis
- **(9)**, an external cause/place of occurrence/activity

Refer to Appendices D and E for full definitions of diagnosis types and the rules for determining significance.

Findings for diagnosis typing for FY 2002–2003 are presented in Table 10.7.1. Inset in the table is a box marked with double lines, which contains conditions that are present in the original and reabstracted data. Conditions that match on diagnosis type upon reabstraction are represented along the diagonal.

Table 10.7.1: Diagnosis Types Before and After Reabstraction, FY 2002–2003, *Unadjusted for Coder Effect*

Original Diagnosis Type		Reabstracted Diagnosis Type							
		Significant				Secondary			
		M	1	2	W,X,Y	3	0	9	Not Reabstracted
Significant	M	230,705 90%	16,382 6%	230 0%	63 0%	9,189 4%	173 0%	0 0%	10,047
	1	20,899 7%	160,852 52%	4,133 1%	602 0%	121,999 40%	41 0%	35 0%	60,338
	2	417 1%	5,022 7%	40,669 55%	11 0%	27,618 37%	0 0%	0 0%	25,793
	W,X,Y	162 2%	222 3%	62 1%	6,375 93%	27 0%	0 0%	0 0%	2,200
Secondary	3	2,845 6%	3,252 7%	773 2%	2 0%	42,135 86%	25 0%	0 0%	0
	0	1,765 24%	133 2%	38 1%	0 0%	17 0%	5,449 74%	0 0%	0
	9	0 0%	0 0%	0 0%	0 0%	0 0%	0 0%	46,722 100%	0
	Not present in Original DAD	9,997	33,152	11,704	566	8,726	561	9,485	

Source: CIHI 2005

Note: Original secondary conditions that were not reabstracted are excluded.
Note: Conditions that are "Not present in Original DAD" are either due to the original coder considering the diagnosis as secondary and not coding it, as it was optional to code. Or, the original coder may not have coded the diagnosis because they did not consider the condition present at all.

There are 230,705 original diagnoses assigned a type (M) that were also reabstracted as a type (M). Conditions that differ in diagnosis type upon reabstraction are not along the diagonal. There are a total of 159,082 original diagnoses assigned a significant type that were reabstracted as secondary. This is 21% of the original significant diagnoses.

Outside the double-lined box are the conditions present in the original data only, or the reabstracted data only. Conditions in the original data only are listed in the last *column* labeled "Not Reabstracted". There are 98,378 original conditions assigned a significant diagnosis type that were not reabstracted at all. This is 13% of the original significant diagnoses. Conditions in the reabstracted data only are listed in the last *row* labeled "Not present in Original DAD"[13]. There are 55,419 reabstracted conditions assigned a significant diagnosis type that are not present in the original data. This is 10% of the reabstracted significant diagnoses.

Findings for FY 2003–2004 are presented in Table 10.7.2. As observed in the previous fiscal year, a large number of original diagnoses assigned a significant type were reabstracted as secondary, though the figure drops to 113,183. This is 18% of the original significant diagnoses.

[13] See footnote to Table 10.7.1 for interpretation of "Not present in Original DAD".

Table 10.7.2: Diagnosis Types Before and After Reabstraction, FY 2003–2004, *Unadjusted for Coder Effect*

Original Diagnosis Type		Reabstracted Diagnosis Type							
		Significant				Secondary			
		M	1	2	W,X,Y	3	0	9	Not Reabstracted
Significant	M	230,595 90%	13,677 5%	1,550 1%	48 0%	8,974 4%	658 0%	0 0%	9,870
Significant	1	16,783 7%	140,260 57%	4,666 2%	133 0%	85,027 34%	280 0%	4 0%	30,528
Significant	2	547 1%	3,358 6%	35,112 61%	0 0%	18,175 32%	0 0%	0 0%	11,307
Significant	W,X,Y	249 3%	605 7%	20 0%	7,648 89%	66 1%	0 0%	0 0%	1,919
Secondary	3	3,042 5%	8,629 13%	691 1%	2 0%	53,128 81%	0 0%	0 0%	0
Secondary	0	3,235 25%	474 4%	13 0%	0 0%	232 2%	8,969 69%	0 0%	0
Secondary	9	0 0%	0 0%	4 0%	0 0%	0 0%	0 0%	44,522 100%	0
	Not Present in Original DAD	10,920	41,902	9,908	403	5,007	327	12,546	

Source: CIHI 2005

Note: Original secondary conditions that were not reabstracted are excluded.
Note: Conditions that are "Not present in Original DAD" are either due to the original coder considering the diagnosis as secondary and not coding it, as it was optional to code. Or, the original coder may not have coded the diagnosis because they did not consider the condition present at all.

Looking to the conditions that were present in the original data only, or the reabstracted data only, similar results are observed in FY 2003–2004. The number of original conditions assigned significant diagnosis types that were not reabstracted decreases to 53,624. This is 9% of the original significant diagnoses. Reabstracted conditions assigned significant diagnosis types that are not present in the original data increases to 63,132. This is 12% of the reabstracted significant diagnoses.

These aggregate figures illustrate this concern. The match rates on pre-admit co-morbid conditions (type (1)) are low at 52% and 57% for the respective study years. For post-admit conditions (type (2)), the match rate is not much better at 55% and 61% respectively. The discrepancies in co-morbidities (type (1) and (2)) are largely related to assignment to secondary conditions upon reabstraction. This type of discrepancy accounts for 40% and 37% of the original co-morbidities in FY 2002–2003 with a slight improvement of 34% and 32% in FY 2003–2004.

Another concern is related to discrepancies regarding the ***presence*** of significant conditions. There were 98,378 original significant conditions that reabstractors were unable to identify on the discharge in FY 2002–2003. This is 13% of all the original significant conditions. Also for this fiscal year, reabstractors identified

55,419 significant diagnoses that were not present in the original data. This is 10% of all reabstracted significant conditions. The equivalent figures are 53,624 (8%) and 63,132 (12%) in FY 2003–2004.

Other observations between fiscal years are that there are less significant diagnoses in the original data. In FY 2002–2003, there are an estimated 744,267 significant diagnoses. This decreases to 622,059 in FY 2003–2004. Also, the number of co-morbid conditions (types (1) and (2)) that were only present in the original data decreases between fiscal years.

Highlights of statistically significant observations:

- ***The decrease between fiscal years in type (1) conditions present only in the original data.*** In FY 2002–2003, this value was 16.4 ± 1.5% of the original type (1) conditions. In FY 2003–2004, this value is 11.0 ± 1.5%.
- ***The decrease between fiscal years in type (2) conditions present only in the original data.*** In FY 2002–2003, this value was 25.9 ± 1.8% of the original type (2) conditions. In FY 2003–2004, this value is 16.5 ± 1.7%.
- ***The increase between fiscal years in type (1) conditions present only in the reabstracted data.*** In FY 2002–2003, this value was 15.1 ± 1.8%. In FY 2003–2004, this value is 20.1 ± 2.2%.
- ***The increase between fiscal years in the match rate of type (1) conditions.*** In FY 2002–2003, this value was 52.1 ± 2.3% of the original type (1) conditions which were identified by the reabstractor. In FY 2003–2004, this value is 56.8 ± 2.3%.
- ***The increase between fiscal years in the match rate of type (2) conditions.*** In FY 2002–2003, this value was 55.2 ± 2.4% of the original type (2) conditions which were identified by the reabstractor. In FY 2003–2004, this value is 61.4 ± 2.6%.

The key concern with diagnosis typing is defining what constitutes a secondary condition compared to a significant condition. This issue is present for both fiscal years of data, though to a lesser extent in FY 2003–2004.

Tables 10.7.3 and 10.7.4 show that the majority of discrepancies related to the typing of conditions present in the original and reabstracted data is for the reason of “significance”, meaning that the reabstractor did not agree with the degree of significance assigned originally. This is in line with the observations presented earlier. Reasons of “chart documentation” and “standards/codebook/manual” are commonly used for original type (M) diagnoses.

Table 10.7.3: Reasons for Diagnosis Typing Discrepancies in FY 2002–2003, *Unadjusted for Coder Effect*

Original Type	Total		Reasons							
			Standards/ Codebook/Manual		Significance		Chart Documentation		Acceptable Difference or OOS	
	Count	%	Count	%	Count	%	Count	%	Count	%
M	26,038	10.1	4,859	1.9	7,745	3.0	13,356	5.2	77	0.0
1	147,709	47.9	8,929	2.9	117,717	38.2	20,114	6.5	948	0.3
2	33,068	44.8	2,302	3.1	25,926	35.2	4,644	6.3	195	0.3
W,X,Y	474	6.9	32	0.5	27	0.4	295	4.3	120	1.8
3	6,897	14.1	1,610	3.3	3,322	6.8	1,896	3.9	69	0.1
0	1,953	26.4	643	8.7	642	8.7	644	8.7	24	0.3
9	0	0.0	0	0.0	0	0.0	0	0.0	0	0.0

Source: CIHI 2005

Note: Excludes diagnoses present in the original data only or the reabstracted data only.
Note: The denominator used to calculate each percentage is the sum of original diagnoses that were assigned a particular diagnosis type, and for which the reabstractor agreed the condition existed.

Table 10.7.4: Reasons for Diagnosis Typing Discrepancies in FY 2003–2004, *Unadjusted for Coder Effect*

Original Type	Total		Reasons							
			Standards/ Codebook/Manual		Significance		Chart Documentation		Acceptable Difference or OOS	
	Count	%	Count	%	Count	%	Count	%	Count	%
M	24,906	9.7	5,508	2.2	7,698	3.0	11,188	4.4	511	0.2
1	106,893	43.2	6,612	2.7	82,277	33.3	16,260	6.6	1,745	0.7
2	22,080	38.6	731	1.3	18,024	31.5	2,881	5.0	444	0.8
W,X,Y	940	10.9	87	1.0	61	0.7	392	4.6	401	4.7
3	12,364	18.9	1,424	2.2	8,686	13.3	2,169	3.3	85	0.1
0	3,954	30.6	899	7.0	1,412	10.9	1,392	10.8	252	1.9
9	4	0.0	4	0.0	0	0.0	0	0.0	0	0.0

Source: CIHI 2005

Note: Excludes diagnoses present in the original data only or the reabstracted data only.
Note: The denominator used to calculate each percentage is the sum of original diagnoses that were assigned a particular diagnosis type, and for which the reabstractor agreed the condition existed.

10.7.1 Analysis of the Assignment of Significance

The analysis presented in Table 10.7.1.1 compares the presence of significant conditions in the original and reabstracted data. There are three subgroups in which the data are characterized.

- **Significant in original and reabstracted** indicates that the reabstractor agreed that the original condition was both present and significant.
- **Significant in original only** indicates one of two things. Either the reabstractor identified that the original condition was present but did not determine the condition was significant. Or, the reabstractor did not agree that the original significant condition was present.

- **Significant in reabstracted only** indicates one of two things. Either the reabstractor entered a significant condition that was originally secondary, or the reabstractor entered a significant condition that was not present in the original data.

Table 10.7.1.1: Assignment of Significance to All Conditions

	DAD FY 2002–2003		DAD FY 2003–2004	
	Count	%	Count	%
Significant Case Costing Conditions	**808,493**		**701,281**	
in Original and Reabstracted	486,806	60.2 ± 1.3	455,251	64.9 ± 1.4
in Original only	257,461	31.8 ± 1.1	166,807	23.8 ± 1.1
in Reabstracted only	64,227	7.9 ± 0.7	79,222	11.3 ± 1.0

Source: CIHI 2005

Note: The figures presented in this table can be derived from those in Tables 10.7.1 and 10.7.2.

There is an increase in the proportion of conditions where the original coder and reabstractor agreed on the assignment of significance and the presence of the condition. The estimated increase in FY 2003–2004 is 5%.

10.7.2 Coder Effect on the Assignment of Diagnosis Type

For those conditions present in the original and reabstracted data, the assignment of diagnosis types upon reabstraction was analyzed for coder effect. This was performed on the typing of the most responsible diagnosis (type (M)) and co-morbid conditions (types (1) and (2)).

As shown in Table 10.7.2.1, there was no statistically significant change in the match rate in the typing of the most responsible diagnosis after adjustments were made for coder effect. The agreement rates slightly decreased after the adjustments, but the change was not statistically significant.

There was a coder effect on the typing of co-morbid conditions. In both fiscal years, the coder effect adjustments significantly decreased these match rates.

- Diagnosis type (1): in FY 2002–2003, the match rate changed from an unadjusted 52% to an adjusted 44%. In FY 2003–2004, the match rate changed from an unadjusted 57% to an adjusted 52%.
- Diagnosis type (2): in FY 2002–2003, the match rate changed from an unadjusted 55% to an adjusted 40%. In FY 2003–2004, the match rate changed from an unadjusted 61% to an adjusted 53%.

Despite the coder effect observed, the trend between fiscal years was maintained, with higher agreement rates observed in FY 2003–2004.

Table 10.7.2.1: Coder Effect on the Agreement Rate for the Assignment of Diagnosis Type to Conditions in the Original and Reabstracted Data

Original Diagnosis Type	DAD FY 2002–2003		DAD FY 2003–2004	
	%	Confidence Interval	%	Confidence Interval
Adjusted for Coder Effect				
M	87.2	[84.8, 89.2]	88.5	[86.4, 90.4]
1	43.9	[42.2, 45.6]	51.6	[49.6, 53.5]
2	40.2	[37.6, 42.9]	53.2	[49.7, 56.7]
Unadjusted				
M	89.9	[88.4, 91.3]	90.3	[88.7, 91.8]
1	52.1	[49.8, 54.4]	56.8	[54.4, 59.1]
2	55.2	[52.8, 57.5]	61.4	[58.8, 64.0]

Source: CIHI 2005

Note: The denominator used to calculate each percentage is the sum of original diagnoses that were assigned a particular diagnosis type, and for which the reabstractor agreed the condition existed.

10.7.3 Coder Effect on the Assignment of Significance

If considering *all* conditions assigned a significant diagnosis type by the original coder and/or the reabstractor (refer to Table 10.7.1.1), the rate of agreement in assigning significance does not change substantially when adjusting for coder effect. As shown in Table 10.7.3.2, agreement rates change from an unadjusted 60% to an adjusted 62% in FY 2002–2003. In FY 2003–2004, the rates change from 65% to 67%. ***The 5% increase in the agreement rates between fiscal years remains significant after the coder effect adjustments.***

Table 10.7.3.2: Coder Effect on the Agreement Rate for the Assignment of Significance to Conditions in the Original and Reabstracted Data

	DAD FY 2002–2003		DAD FY 2003–2004	
	%	Confidence Interval	%	Confidence Interval
Significant in Original and Reabstracted				
Adjusted for Coder Effect	62.2	[60.9, 63.5]	67.4	[66.2, 68.6]
Unadjusted	60.2	[58.9, 61.5]	64.9	[63.5, 66.3]

Source: CIHI 2005

Note: The denominator for percentages is 808,493 in FY 2002–2003 and 701,281 in FY 2003–2004.

10.7.4 Results for Selected Medical Conditions

Tables 10.7.4.1 and 10.7.4.2 present the diagnosis type discrepancy rates for the selected medical conditions, for each fiscal year. These tables include results for diagnosis types (M), (1), and (2) only and have not been adjusted for coder effect. Other diagnosis types are not analyzed in this section due to the large margins of error associated with them.

Table 10.7.4.1: Diagnosis Type Discrepancy Rate for Selected Medical Conditions for FY 2002–2003, *Unadjusted for Coder Effect*

Medical Condition	Original Diagnosis Type		
	M	1	2
All Medical Conditions	**10.1 ± 1.5**	**47.9 ± 2.3**	**44.8 ± 2.4**
Blood Disorders	48.5 ± 31.2	47.2 ± 10.3	54.6 ± 5.6
Circulatory Conditions	14.8 ± 4.8	56.0 ± 4.2	43.9 ± 5.9
Diabetes Mellitus	2.1 ± 1.9	50.0 ± 7.1	95.8 ± 6.6
Digestive Conditions	4.8 ± 4.6	50.0 ± 13.7	57.9 ± 35.4
Genitourinary Conditions	23.1 ± 12.9	36.9 ± 8.1	34.9 ± 6.5
Injuries	3.1 ± 3.3	38.2 ± 35.7	6.9 ± 16.5
Mental Health	4.3 ± 4.4	44.6 ± 9.0	31.2 ± 15.5
Metabolic Disorders	34.0 ± 31.0	57.1 ± 7.0	68.9 ± 7.2
Neonates	4.8 ± 3.8	21.2 ± 9.0	74.9 ± 29.8
Neoplasms	21.5 ± 8.5	55.7 ± 10.4	100.0 ± 0.0
Obstetrics	12.1 ± 4.7	24.3 ± 7.2	56.6 ± 24.4
Other factors	18.6 ± 17.8	52.3 ± 11.9	100.0 ± 0.0
Post-procedural Compl.	3.8 ± 3.1	58.6 ± 19.7	17.1 ± 4.3
Respiratory Conditions	13.4 ± 6.7	48.5 ± 7.6	52.1 ± 7.7
Symptoms	8.4 ± 8.3	64.7 ± 5.9	67.3 ± 5.8

Source: CIHI 2005

Note: Excludes diagnoses present in the original data only or the reabstracted data only.
Note: The denominator used to calculate each percentage is the sum of original diagnoses that were assigned a particular diagnosis type, and for which the reabstractor agreed the condition existed.
Note: Estimates based off a sample size of less than 20 are in grey font.
Note: Refer to page 39 for the listing of diagnosis codes that belong to each medical condition.

Key findings for selection of diagnosis type for FY 2002–2003:

The following medical conditions have statistically significant ***lower discrepancy rates*** for diagnosis typing than the results for all medical conditions:

- ***Type (M)***: ***diabetes mellitus, injuries, neonates, post-procedural complications***
- ***Type (1)***: ***genitourinary conditions, neonates, obstetrics***
- ***Type (2)***: ***genitourinary conditions, post-procedural complications***

The following medical conditions have statistically significant ***higher discrepancy rates*** for diagnosis typing than the results for all medical conditions:

- ***Type (M)***: ***blood disorders, neoplasms***
- ***Type (1)***: ***circulatory conditions, symptoms***
- ***Type (2)***: ***blood disorders, metabolic disorders, symptoms***

Table 10.7.4.2: Diagnosis Type Discrepancy Rate for Selected Medical Conditions for FY 2003–2004, *Unadjusted for Coder Effect*

Medical Condition	Original Diagnosis Type		
	M	1	2
All Medical Conditions	**9.7 ± 1.5**	**43.2 ± 2.3**	**38.6 ± 2.6**
Blood Disorders	12.8 ± 24.7	44.4 ± 10.3	43.0 ± 8.9
Circulatory Conditions	9.9 ± 3.7	48.0 ± 4.8	43.1 ± 9.6
Diabetes Mellitus	9.2 ± 5.5	36.5 ± 8.8	72.1 ± 35.9
Digestive Conditions	10.2 ± 12.3	29.3 ± 20.0	13.3 ± 13.5
Genitourinary Conditions	14.8 ± 10.3	38.8 ± 8.6	35.5 ± 9.0
Injuries	1.3 ± 1.3	42.7 ± 28.3	65.6 ± 47.0
Mental Health	6.0 ± 5.6	32.0 ± 9.2	28.9 ± 24.7
Metabolic Disorders	15.7 ± 16.0	40.9 ± 9.5	67.8 ± 6.1
Neonates	10.6 ± 5.9	16.2 ± 7.1	54.0 ± 38.6
Neoplasms	23.5 ± 10.5	50.7 ± 10.4	100.0 ± 0.0
Obstetrics	13.1 ± 4.5	22.8 ± 7.5	49.5 ± 29.4
Other factors	9.3 ± 11.8	61.8 ± 13.0	100.0 ± 0.0
Post-procedural Compl.	17.9 ± 19.4	31.9 ± 13.5	14.0 ± 4.1
Respiratory Conditions	7.1 ± 3.7	37.8 ± 8.0	42.2 ± 8.8
Symptoms	2.7 ± 2.4	71.9 ± 8.8	61.2 ± 7.4

Source: CIHI 2005

Note: Excludes diagnoses present only in the original or reabstracted data
Note: The denominator used to calculate each percentage is the sum of original diagnoses that were assigned a particular diagnosis type, and for which the reabstractor agreed the condition existed.
Note: Estimates based off a sample size of less than 20 are in grey font.
Note: Refer to page 39 for the listing of diagnosis codes that belong to each medical condition.

Key findings for selection of diagnosis type for FY 2003–2004:

The following medical conditions have statistically significant ***lower discrepancy rates*** for diagnosis typing than the results for all medical conditions:

- ***Type (M)***: ***injuries, symptoms***
- ***Type (1)***: ***neonates, obstetrics***
- ***Type (2)***: ***digestive conditions, post-procedural complications***

The following medical conditions have statistically significant ***higher discrepancy rates*** for diagnosis typing than the results for all medical conditions:

- ***Type (M)***: ***neoplasms***
- ***Type (1)***: ***other factors, symptoms***
- ***Type (2)***: ***metabolic disorders, symptoms***

10.7.5 Coding Specific Observations

Many of the issues associated with the typing of conditions are explained in the paragraphs below. Reasons for discrepancies are due to different interpretation of chart documentation and non-compliance to specific coding standards.

Asterisk codes
Asterisk codes, denoted by an asterisk (*) symbol, describe the manifestation of a condition. They are sequenced after a dagger (†) code representing the underlying condition[14]. In FY 2002–2003, discrepancies in asterisk codes were found when originally assigned a type (1). Asterisks codes are to be assigned a diagnosis type (3), though not explicitly stated in the coding standards. Fewer discrepancies were found in FY 2003–2004.

Circulatory conditions
Particularly in FY 2002–2003, circulatory conditions originally assigned a type (1) were often reabstracted as secondary because they did not satisfy the requirements for significant conditions. Discrepancy rates were particularly high for hypertension, hypotension, angina and atrial fibrillation. Specific coding standards on angina state "angina may only be coded as a significant diagnosis when there is a documented episode of angina at admission or at any given time during the hospital stay." Discrepancies are due to issues related to chart documentation.

Diabetes mellitus
Most discrepancies were associated with E11 "type 2 diabetes mellitus". Diabetes codes originally assigned type (2) were reabstracted as secondary. Coding standards state "diabetes mellitus should never be assigned a diagnosis type (2)." There are a few exceptions when diabetes may be assigned type (2), such as in the case of drug-induced hypoglycemia. Generally, diabetes is regarded as a chronic condition. If significant, it should be coded either as a type (1) or (M).

Genitourinary conditions
Renal failures (N17, N18, N19) originally assigned types (M) or (1) were reabstracted as secondary. Most often the reabstractor coded renal failure as a supplemental condition for additional specificity to diabetes mellitus. Diabetes associated with complications requires an additional code from other chapters to fully describe the condition and is generally assigned as type (3). This is demonstrated in examples in the coding standards. Fewer discrepancies were found in FY 2003–2004. When renal failure is not associated with diabetes, it was reabstracted as a type (3) because it did not satisfy the criteria for significance.

Malignant neoplasm site unspecified
Code C80 "malignant neoplasm without specification of site" is a vague category that is usually a supplemental code when the primary/metastatic sites are not indicated. C80 codes that were originally typed as significant were reabstracted as

[14] For example, M90.7*, "Fracture of bone in neoplastic disease", is an asterisk code that would be sequenced after a code from the neoplasm chapter (C00-D48†).

type (3). C80 is to be coded as significant only if a diagnosis of "carcinomatosis" was documented and C80 met the criteria of a significant diagnosis. New education sessions will be offered in FY 2005/2006 to clarify existing neoplasm standards.

Metabolic disorders
For both fiscal years, discrepancies were noted for code E87 "other disorders of fluid, electrolyte and acid-base balance". Conditions such as hyponatremia, hypernatremia, and metabolic acidosis are included in this category. These discrepancies indicate different interpretation of chart documentation and diagnosis typing definitions.

Neoplasms
Neoplasms originally assigned type (M) were often reabstracted as secondary. The diagnosis type should reflect the impact on the hospital stay and whether treatment was directed towards it. In several instances the reabstractor assigned the MRDx as chemotherapy or radiation therapy, and the neoplasm code was assigned as secondary, as per coding standards. In addition to variance from coding standards, chart documentation is another contributing factor to the discrepancies.

Obstetrics (mother's chart)
Discrepancies were found for code Z37 "outcome of delivery" originally assigned a type (M) but reabstracted as a type (3). According to coding standards, when a code from the obstetric chapter is applicable, this should be sequenced first, followed by Z37 as a diagnosis type (3). Discrepancy rates were lower in FY 2003–2004.

Other factors influencing health status
The greatest proportion of discrepancies was related to codes from Z51 "other medical care". This category includes admission for palliative care and other medical care such as pain management. In most instances, codes from Z51 were assigned as type (1) but reabstracted as secondary. Currently, coding standards that address diagnosis typing for palliative care and other medical care do not exist. The online Coding Query database has addressed queries on a case-specific basis.

Post-procedural conditions/complications
Though the discrepancy rate for post-procedural conditions is low, trends were identified, particularly in FY 2002–2003. Conditions from the circulatory and respiratory chapters were originally coded as type (2) and reabstracted as type (3). Reabstracted type (3) codes suggest these are supplemental codes, adding further specificity to the post-procedural code.

Problems related to medical facilities and other health care
Codes from the Z75 category were noted to have discrepancies with diagnosis typing. This includes codes related to awaiting admission to another facility or awaiting treatment. Variation was largely attributed to chart documentation.

Symptoms
Symptom codes assigned a type (1) or (2) were reabstracted as secondary. Discrepancies are due to non-compliance to coding standards. The coding standard on underlying symptoms or conditions state "when a patient presents with a symptom and during that episode of care the underlying disease is identified, the underlying disease is assigned and the symptom may be coded as a type (3), based on a facility's data needs". The coding standard for post-procedural signs and symptoms state the criteria that the symptom must qualify in order to be coded as a type (2). If symptoms do not meet the criteria, they are assigned a type (3). Discrepancies were noted for R41.0 "disorientation, unspecified" and codes from R13 "dysphagia".

10.7.5.1 Coding issues related to significant conditions in the original data only

The majority of discrepancies where the reabstractor did not agree on the presence of an original diagnosis were assigned a reason of "chart documentation" (68% and 53% for the respective study years), meaning the reabstractor disagreed with the original interpretation of the chart information. The other common reason assigned was "standards/codebook/manual" (22% and 34%) meaning there was non-compliance to the coding standards. For instance, if the original condition was noted in X-ray results and not substantiated by physician documentation, this would not confirm the presence of the condition. According to coding standards, investigation reports are to be used for additional specificity to physician documentation. The following code categories were originally coded as significant and not reabstracted:

- D62 "Acute posthemorrhagic anemia"
- E87 "Other disorders of fluid, electrolyte and acid-base balance"
- J90 "Pleural effusion"
- T81 "Complications of procedures"
- D68 "Other coagulation defects"
- J18 "Pneumonia, organism unspecified"
- E11 "Type 2 diabetes mellitus"
- I50 "Heart failure"
- D64 "Other anemias" (FY 2002–2003 only)
- R26 "Abnormalities of gait and mobility (FY 2002–2003 only)
- I10 "Essential (primary) hypertension" (FY 2002–2003 only)
- Z75 "Problems related to medical facilities and other health care (awaiting admission elsewhere, awaiting investigation or treatment)" (FY 2003–2004 only)
- A41 "Other septicemia" (FY 2003–2004 only)

10.7.5.2 Coding issues related to significant conditions in the reabstracted data only

When analyzing the significant conditions that were present in the reabstracted data only, several conditions listed above were identified again as the source of the difference. Discrepancy reasons of "chart documentation" (76% and 82% for the respective fiscal years) indicate that most conditions were not in the original data due to issues related to chart documentation. Other reasons of "significance" (11% and 6%) and "standards/codebook/manual" (11% and 11%) were used for the remaining discrepancies. The following code categories were reabstracted as significant but not present in the original data:

- T81 "Complications of procedures"
- I97 "Post-procedural disorders of circulatory system" *
- J95 "Post-procedural respiratory disorders" *
- E11 "Type 2 Diabetes mellitus"
- E86 "Volume depletion"
- Z51 "Other medical care"
- E87 "Other disorders of fluid, electrolyte and acid-base balance"
- D64 "Other anaemias"
- I50 "Heart failure"
- J18 "Pneumonia, organism unspecified"
- I10 "Essential (primary) hypertension" (FY 2002–2003 only)
- Z75 "Problems related to medical facilities and other health care (awaiting admission elsewhere, awaiting investigation or treatment)" (FY 2003–2004 only)
- A41 "Other septicemia" (FY 2003–2004 only)
- K91 "Post-procedural disorders of digestive system" * (FY 2003–2004 only)

* Post-procedural codes captured by the reabstractor but not present in original data suggests non-compliance to coding standards. Typically, post-procedural codes are sequenced first followed by a secondary code to fully describe the condition.

10.7.6 Improving the Quality of Diagnosis Type Selection

The relatively low agreement rate on the assignment of significance to conditions indicates that significance is either not well-defined, not well understood by coders, or difficult to interpret from the chart documentation.

Some discrepancies identified in diagnosis type selection are related to non-compliance to existing coding standards. Health information professionals can refer to the following Canadian Coding Standards for ICD-10-CA and CCI:

- Diagnosis typing definitions
- Neoplasms (all coding standards)
- Diagnosis typing of diabetes mellitus
- Underlying symptoms or conditions
- Post-procedural conditions and complications
- Angina
- Delivery in a completely normal case
- Using diagnostic test results in coding, impending or threatened conditions

CIHI education sessions have provided additional reference material in participant workbooks that address several coding standards outlined. Health information professionals can refer to the following workbook materials for further clarification:

- The Canadian Coding Standards and Diagnosis Typing for DAD
- Obstetrical and Newborn Coding

CIHI has undertaken the following initiatives to address diagnosis typing in FY 2005–2006:

- 25 classification edits were implemented in the DAD, some to directly improve the quality of diagnosis typing.
- Diagnosis typing definitions were updated to include type (6), known as a proxy MRDx. There are also 194 codes that are only valid as diagnosis type (3).

Effective in FY 2006–2007, a number of existing standards will be reformatted to emphasize the coding directives.

10.8 Complexity Level Assignment

Complexity assignment reflects the interaction of multiple diagnoses on length of stay or resources within each case mix group. Complexity level definitions are:

1 = No complexity
2 = Complexity related to chronic condition(s)
3 = Complexity related to serious/important condition(s)
4 = Complexity related to potentially life-threatening condition(s)
9 = Complexity not applied

Original data and reabstracted data were both processed through the 2003 CIHI grouping methodology and the resultant complexity level assignments were compared to each other. Overall, the complexity levels changed upon reabstraction for 13.7 ± 0.9% (FY 2002–2003) and 12.5 ± 1.4% (FY 2003–2004) of the discharges. The observed change between fiscal years is not significant. These percentages include discharges originally assigned a complexity level 9. Complexity level 9 indicates that a complexity overlay was not applied, and is assigned to obstetrical, neonate, and mental health discharges. If excluding these discharges, complexity levels changed upon reabstraction for 20.0 ± 1.3% (FY 2002–2003) and 18.4 ± 2.1% (FY 2003–2004) of the discharges. The difference between fiscal years is not statistically significant.

10.8.1 Coder Effect on the Assignment of Complexity Level

The increase in the agreement rate of complexity assignment between fiscal years (from 79.0% to 84.6%) was significant after coder effect and case mix effect adjustments.

There was no statistically significant change in the agreement rate for complexity assignment after adjustments were made for coder effect. These results are shown in Table 10.8.1.1. In FY 2002–2003, the match rate when including discharges assigned to complexity level 9 changed from an unadjusted 86% to an adjusted 87%. In FY 2003–2004, the match rate changed from an unadjusted 88% to an adjusted 90%. There was also no statistically significant difference between the adjusted and unadjusted results when discharges originally assigned to complexity level 9 were excluded.

Table 10.8.1.1: Coder Effect on the Agreement Rates for Complexity Assignment

	DAD FY 2002–2003		DAD FY 2003–2004	
	%	Confidence Interval	%	Confidence Interval
Including Plx Level 9				
Adjusted for Coder Effect	87.2	[84.3, 89.6]	90.0	[87.6, 92.0]
Unadjusted	86.3	[85.5, 87.2]	87.5	[86.1, 88.9]
Excluding Plx Level 9				
Adjusted for Coder Effect	79.0	[75.8, 81.9]	84.6	[81.9, 87.0]
Unadjusted	80.0	[78.8, 81.3]	81.6	[79.5, 83.7]

Source: CIHI 2005

Note: For "Including Plx level 9", the denominator for percentages is 266,790 in FY 2002–2003 and 265,372 in FY 2003–2004.
Note: For "Excluding Plx level 9", the denominator for percentages is 179,802 in FY 2002–2003 and 175,003 in FY 2003–2004.

10.8.2 Results for Specific Complexity Levels

Tables 10.8.2.1 and 10.8.2.2 show the comparisons of complexity assignment upon reabstraction for each fiscal year. These tables contain estimated counts and row percentages, and have not been adjusted for coder effect.

Matches in complexity assignment before and after reabstraction are represented along the diagonal. For instance, 118,963 discharges in FY 2002–2003 were grouped to a complexity level of "1" before and after reabstraction. Below this count is the row percentage of 95%, which refers to the match rate of all the discharges that were *originally* assigned complexity level "1".

Changes in complexity assignment upon reabstraction are represented in the cells that do not fall on the diagonal. Cells below the diagonal represent discharges that were grouped to a *lower* complexity level upon reabstraction (with the exception of complexity 9). Those above the diagonal represent discharges that were grouped to a *higher* complexity level upon reabstraction. For both fiscal years, reabstracted data tend to be grouped to a lower complexity level. This can be seen by the larger counts in the cells below the diagonal. For example, in FY 2002–2003, 11,564 discharges decrease in complexity from an original value "2" to a reabstracted value of "1". Only 3,571 discharges satisfy the opposite trend, from an original value "1" increasing to a reabstracted value of "2".

Complexity levels generated using the original data generally match the complexity level using the reabstracted data. However, the key observation for those discharges that do not match is that the original complexity levels tend to be higher than the reabstracted complexity levels.

Table 10.8.2.1: Complexity Level Assignment Before and After Reabstraction for FY 2002–2003, *Unadjusted for Coder Effect*

Original Complexity Level	Reabstracted Complexity Level					
	1	2	3	4	9	Total Original
1	118,963 95%	3,571 3%	1,770 1%	286 0%	16 0%	**124,606** 100%
2	11,564 48%	10,424 44%	1,385 6%	383 2%	143 1%	**23,898** 100%
3	5,922 38%	3,297 21%	5,245 34%	964 6%	80 1%	**15,509** 100%
4	1,940 12%	1,723 11%	2,823 18%	9,254 59%	49 0%	**15,789** 100%
9	380 0%	119 0%	31 0%	15 0%	86,444 99%	**86,988** 100%
Total Reabstracted	**138,768**	**19,134**	**11,254**	**10,901**	**86,732**	

Source: CIHI 2005

Note: The original and reabstracted data were grouped using the 2003 CIHI grouping methodology

Table 10.8.2.2: Complexity Level Assignment Before and After Reabstraction for FY 2003–2004, *Unadjusted for Coder Effect*

Original Complexity Level	Reabstracted Complexity Level					
	1	2	3	4	9	Total Original
1	119,973 93%	4,103 3%	4,503 3%	179 0%	620 0%	**129,377** 100%
2	8,370 42%	9,825 50%	1,173 6%	349 2%	102 1%	**19,819** 100%
3	4,048 31%	2,218 17%	5,379 42%	1,190 9%	74 1%	**12,909** 100%
4	1,951 15%	1,203 9%	2,134 17%	7,589 59%	21 0%	**12,898** 100%
9	648 1%	100 0%	73 0%	72 0%	89,476 99%	**90,369** 100%
Total Reabstracted	**134,990**	**17,448**	**13,262**	**9,380**	**90,292**	

Source: CIHI 2005

Note: The original and reabstracted data were grouped using the 2003 CIHI grouping methodology

10.8.3 Results by Major Clinical Category

Further analysis of complexity assignment was done by MCC category to determine if certain categories were more prone to changes. For this analysis, the percent change in complexity was calculated for the subset of discharges where the original data were grouped to a particular MCC category. This calculation is done irrespective of the MCC category in which the reabstracted data were grouped. Note that these results have not been adjusted for coder effect.

Table 10.8.3.1: Change in Complexity Level upon Reabstraction by Major Clinical Category, *Unadjusted for Coder Effect*

Major Clinical Category	DAD FY 2002–2003			DAD FY 2003–2004		
	% of Records			% of Records		
Overall	**13.7**	±	**0.9**	**12.5**	±	**1.4**
1 - Nervous System	33.8	±	8.5	24.4	±	6.7
2 - Eye Diseases	13.8	±	27.0	0.0	±	0.0
3 - Ear Nose & Throat	12.3	±	7.3	4.6	±	4.2
4 - Respiratory	23.5	±	4.9	18.6	±	4.7
5 - Card & Vasc Diseases Circ System	25.1	±	4.0	20.8	±	4.4
6 - Digestive	18.6	±	4.0	19.1	±	8.1
7 - Hepatobiliary & Pancreas	20.7	±	8.1	29.7	±	13.9
8 - Musculoskeletal & Connect	14.8	±	4.1	13.0	±	3.9
9 - Skin Subcut & Breast	10.3	±	7.2	15.9	±	13.6
10 - Endocrine Nutrit & Metabolism	14.3	±	6.7	21.1	±	11.4
11 - Kidney & Urinary Tract	14.1	±	5.3	13.2	±	4.2
12 - Male Reproductive	5.6	±	76.5	74.8	±	55.7
13 - Female Reproductive	4.7	±	3.0	9.3	±	11.2
14 - Pregnancy & Childbirth	0.3	±	0.7	0.0	±	0.0
15 - Newborns & Other Neonates	0.0	±	0.0	0.0	±	0.0
16 - Bld, BldForming Org & Immun	21.9	±	11.7	8.9	±	8.5
17 - Lymph/Leukem & Neoplasm Unspec	15.5	±	7.0	16.9	±	7.7
18 - Multisys/Unspec Site Infection	28.5	±	13.6	30.0	±	19.5
19 - Mental Disease & Disorders	2.9	±	1.9	4.0	±	3.8
21 - Injuries Poison & Toxic Effect	24.7	±	10.6	23.1	±	21.2
22 - Burns	0.0	±	0.0	4.8	±	43.9
23 - Other reasons for hospitalization	45.2	±	12.9	25.9	±	12.5
24 - HIV Infections (AIDS)	33.5	±	43.4	30.5	±	36.5
25 - Significant Trauma	16.7	±	4.8	16.9	±	5.9
99 - Ungroupable Data	0.0	±	0.0	100.0	±	0.0

Source: CIHI 2005

Note: The original and reabstracted data were grouped using the 2003 CIHI grouping methodology.
Note: The denominators for the percentages are the weighted sum of discharges assigned to each particular MCC category.
Note: Estimates based off a sample size of less than 20 are placed in grey font.
Note: These percentages were calculated for all discharges, including those originally grouped to "no complexity" (level 9).

Key findings for complexity level assignment, by major clinical category

- A significantly lower proportion of discharges originally assigned to ***MCC3 (FY 2003–2004), MCC13 (FY 2002–2003), MCC14 (FY 2002–2003), and MCC19*** changed complexity level upon reabstraction when compared to the overall discrepancy rate of 13.7% in FY 2002–2003 and 12.5% in FY 2003–2004 for all medical conditions.
- A significantly higher proportion of discharges originally assigned to ***MCC1, MCC4, MCC5, MCC6 (FY 2002–2003), MCC7 (FY 2003–2004), MCC18 (FY 2002–2003), and MCC23 (FY 2002–2003)*** changed complexity level upon reabstraction when compared to the overall discrepancy rate of 13.7% and 12.5% for all medical conditions.
- No discharges assigned to ***MCC14 (FY 2003–2004) and MCC15*** changed complexity level upon reabstraction.

10.8.4 Improving the Quality of Complexity Assignment

Complexity assignment is affected by the typing of conditions as significant compared to secondary, or vice versa. It also is affected when conditions are coded in error due to interpretation of chart documentation, or not coded when they should be due to incomplete chart documentation. As such, previous discussion on improving the assignment of significance is applicable here, as is the need for timely and high quality chart documentation.

10.9 Expected Length of Stay

Expected length of stay (ELOS) is the projected duration of a patient's visit, in days, for a typical acute care case in a case mix group. ELOS is a national average length of stay (ALOS) estimate that accounts for differences in age and complexity when these factors are found to be predictive of length of stay.

Original data and reabstracted data were both processed through the 2003 CIHI grouping methodology and the resultant ELOS values were compared to each other. The percent net change in the ELOS values upon reabstraction was -7.3 ± 1.2% in FY 2002–2003 and –4.5 ± 1.2% in FY 2003–2004. ***The decrease in the net change of ELOS observed in FY 2003–2004 is statistically significant. Also, for both fiscal years, reabstracted data consistently resulted in lower ELOS values than that calculated using the original data.***

10.9.1 Coder Effect on Expected Length of Stay

The model developed to adjust results for coder effect could not be applied to the analysis of expected length of stay. For this data element, the measure of interest is the percent net change in value upon reabstraction.

10.9.2 Results by Major Clinical Category

Further analysis of ELOS values was done for each of the 25 MCC categories to identify if changes in ELOS values were related to these categories. For this analysis, the percent net change in ELOS was calculated for the subset of discharges where the original data were grouped to a particular MCC category. This calculation is done irrespective the MCC category to which the reabstracted data were grouped. These results are shown in Table 10.9.2.1.

Table 10.9.2.1: Net Change in Expected Length of Stay, by Major Clinical Category, *Unadjusted for Coder Effect*

Major Clinical Category	DAD FY 2002–2003 % Net Change	DAD FY 2003–2004 % Net Change
Overall	**-7.3 ± 1.2**	**-4.5 ± 1.2**
1 - Nervous System	-10.1 ± 6.5	-13.3 ± 4.9
2 - Eye Diseases	36.7 ± 59.5	0.0 ± 0.0
3 - Ear Nose & Throat	-8.0 ± 13.5	-2.7 ± 15.6
4 - Respiratory	-7.8 ± 5.4	-3.1 ± 3.6
5 - Card & Vasc Diseases Circ System	-7.7 ± 2.2	-3.5 ± 2.7
6 - Digestive	-8.4 ± 3.5	-1.6 ± 6.2
7 - Hepatobiliary & Pancreas	-11.5 ± 5.8	-4.6 ± 9.1
8 - Musculoskeletal & Connect	-10.0 ± 4.3	-5.8 ± 3.5
9 - Skin Subcut & Breast	-13.2 ± 7.1	7.1 ± 19.8
10 - Endocrine Nutrit & Metabolism	-5.6 ± 6.4	-16.3 ± 5.6
11 - Kidney & Urinary Tract	-6.0 ± 5.3	-7.9 ± 5.8
12 - Male Reproductive	-8.6 ± 53.2	-17.5 ± 61.5
13 - Female Reproductive	-6.2 ± 4.6	-5.6 ± 5.3
14 - Pregnancy & Childbirth	0.3 ± 1.4	0.5 ± 1.4
15 - Newborns & Other Neonates	-0.7 ± 0.6	-0.6 ± 1.2
16 - Bld, BldForming Org & Immun	-13.1 ± 11.4	-4.9 ± 5.0
17 - Lymph/Leukem & Neoplasm Unspec	-14.4 ± 5.8	-15.4 ± 6.3
18 - Multisys/Unspec Site Infection	-13.0 ± 8.2	-20.9 ± 10.0
19 - Mental Disease & Disorders	-7.8 ± 3.6	-1.5 ± 2.9
21 - Injuries Poison & Toxic Effect	-14.7 ± 9.6	11.7 ± 18.0
22 - Burns	0.0 ± 0.0	-39.0 ± 52.6
23 - Other reasons for hospitalization	-15.1 ± 27.7	-19.0 ± 10.0
24 - HIV Infections (AIDS)	-19.2 ± 36.4	8.1 ± 36.1
25 - Significant Trauma	-4.6 ± 4.2	-7.1 ± 3.2
99 - Ungroupable Data	0.0 ± 0.0	92.1 ± 0.0

Source: CIHI 2005

Note: The original and reabstracted data were grouped using the 2003 CIHI grouping methodology.
Note: Denominators for percentages are the weighted sum of the original ELOS values for each MCC category.
Note: Estimates based off a sample size of less than 20 are placed in grey font.
Note: The percent net change was calculated using the original and reabstracted ELOS values for all discharges, including those grouped to "no complexity" (level 9).

Key findings for expected length of stay, by major clinical category

- Discharges originally assigned to ***MCC14 and MCC15*** have percent net changes in ELOS values significantly closer to zero than the −7.3% and −4.5% net change observed for all discharges combined. That is, discharges belonging to these MCC categories have the original ELOS value significantly closer in value to the reabstracted ELOS value.
- Discharges originally assigned to ***MCC1 (FY 2003–2004), MCC10 (FY 2003–2004), MCC17, MCC18 (FY 2003–2004), and MCC23 (FY 2003–2004)*** have net changes in ELOS values that are significantly greater from the −7.3% and −4.5% net change for all discharges combined. Discharges belonging to these MCC categories have original ELOS values significantly greater in value than the reabstracted ELOS values.

10.9.3 Improving the Quality of Expected Length of Stay

Calculations for expected length of stay are driven by CMG assignment, complexity level, and age. The ELOS calculation also factors in the effect of age at different complexity levels, if there is such an effect. Age is calculated using the non-medical data elements “Birth Date” and “Admission Date”, which were found to have very good agreement upon reabstraction. Hence, to improve the quality of the ELOS values, improvements are needed in the data that feed into the calculation of CMG and complexity.

10.10 Resource Intensity Weight

Resource intensity weight (RIW) provides users with a tool to estimate expected resource use and relationships of costs between patient types. It indicates the relative value of treating a patient compared to the average patient that has an RIW value of 1.0000.

Original data and reabstracted data were both processed through the 2003 CIHI grouping methodology and the resultant RIW values were compared to each other. The percent net change in RIW upon reabstraction was $-4.3 \pm 1.1\%$ in FY 2002–2003 and $-2.8 \pm 1.2\%$ in FY 2003–2004. The observed difference between fiscal years is not statistically significant. ***For both fiscal years, reabstracted data consistently resulted in lower RIW values than that calculated using the original data.***

10.10.1 Coder Effect on Resource Intensity Weight

The model developed to adjust results for coder effect could not be applied to the analysis of resource intensity weight. For this data element, the measure of interest is the percent net change in value upon reabstraction.

10.10.2 Results by Major Clinical Category

Further analysis of RIW values was done for each of the 25 MCC categories to identify if changes in RIW values were related to these categories. For this analysis, the percent net change in RIW value was calculated for the subset of discharges where the original data were grouped to a particular MCC category. This calculation is done irrespective of the MCC value to which the reabstracted data were grouped. These results are shown in Table 10.10.2.1.

Table 10.10.2.1: Net Change in Resource Intensity Weight, by Major Clinical Category, *Unadjusted for Coder Effect*

Major Clinical Category	DAD FY 2002–2003 % Net Change	DAD FY 2003–2004 % Net Change
Overall	**-4.3 ± 1.1**	**-2.8 ± 1.2**
1 - Nervous System	-6.2 ± 3.9	-8.4 ± 3.5
2 - Eye Diseases	25.5 ± 45.8	0.0 ± 0.0
3 - Ear Nose & Throat	-8.7 ± 10.0	10.3 ± 30.1
4 - Respiratory	-6.2 ± 5.9	-5.3 ± 1.9
5 - Card & Vasc Diseases Circ System	-3.7 ± 2.0	-2.0 ± 1.9
6 - Digestive	-4.1 ± 3.6	-0.9 ± 4.6
7 - Hepatobiliary & Pancreas	-11.1 ± 5.2	-3.4 ± 7.5
8 - Musculoskeletal & Connect	-2.8 ± 3.9	-4.6 ± 3.0
9 - Skin Subcut & Breast	-0.2 ± 9.7	3.7 ± 12.6
10 - Endocrine Nutrit & Metabolism	-2.1 ± 6.4	-11.4 ± 7.9
11 - Kidney & Urinary Tract	-4.8 ± 4.9	-2.8 ± 4.5
12 - Male Reproductive	-13.2 ± 56.6	-1.5 ± 21.5
13 - Female Reproductive	-3.3 ± 3.4	-1.3 ± 6.8
14 - Pregnancy & Childbirth	-0.1 ± 1.6	0.8 ± 1.3
15 - Newborns & Other Neonates	-0.2 ± 2.1	-1.3 ± 2.2
16 - Bld, BldForming Org & Immun	-14.4 ± 11.8	-5.4 ± 5.6
17 - Lymph/Leukem & Neoplasm Unspec	-5.5 ± 8.1	-13.2 ± 5.3
18 - Multisys/Unspec Site Infection	-15.1 ± 7.3	-13.5 ± 7.1
19 - Mental Disease & Disorders	-0.9 ± 5.9	-2.0 ± 2.6
21 - Injuries Poison & Toxic Effect	-15.0 ± 8.0	16.8 ± 29.1
22 - Burns	0.0 ± 0.0	-36.5 ± 49.6
23 - Other reasons for hospitalization	-9.9 ± 6.3	-4.5 ± 5.3
24 - HIV Infections (AIDS)	-9.6 ± 22.3	8.9 ± 32.1
25 - Significant Trauma	-3.5 ± 3.8	-2.3 ± 3.5
99 - Ungroupable Data	0.0 ± 0.0	138.1 ± 0.0

Source: CIHI 2005

Note: The original and reabstracted data were grouped using the 2003 CIHI grouping methodology.
Note: Denominators for percentages are the weighted sum of the original RIW values for each MCC category.
Note: Estimates based off a sample size of less than 20 are placed in grey font.
Note: The percent net change was calculated using the original and reabstracted RIW values for all discharges, including those grouped to "no complexity" (level 9).

Key findings for resource intensity weight, by major clinical category

- Discharges originally assigned to ***MCC14 (FY 2002–2003) and MCC15 (FY 2002–2003)*** have percent net changes in RIW values that are significantly closer to zero than the results calculated for all MCC categories combined. That is, original RIW values were closer to the reabstracted RIW values.
- Discharges originally assigned to ***MCC14 (FY 2003–2004)*** have percent net changes in RIW values that are significantly different from the −2.8% net change for all discharges. The net change for this MCC category, however, is not significantly closer to zero.
- Discharges originally assigned to ***MCC1 (FY 2003–2004), MCC7 (FY 2002–2003), MCC17 (FY 2003–2004), MCC18, MCC21 (FY 2002–2003)*** have percent net changes in RIW values that are significantly greater than −4.3% (FY 2002–2003) and −2.8% (FY 2003–2004) net change calculated for all discharges. For these MCC categories, original RIW values were much greater than the reabstracted RIW values.

11 Facility Specific Results

Findings in this section have been adjusted for case mix effect. When possible, adjustments for coder effect were also applied.

This section analyzes the facility specific findings from the first two years of implementation of ICD-10-CA and CCI for Ontario's case-costing facilities.

Some facilities treat a high volumes of obstetrical and neonate patients, while others treat few or no patients of these types. The codes associated with obstetrical and neonate patients were reabstracted with high agreement, as shown in section 10. Differences between facilities are partly due to the differing proportions of diagnoses (or discharges) associated with these types of patients.

When analyzing facility specific results, ***the population of reference was redefined in the analysis to reduce the case mix effect***. That is, discrepancy rates were recalculated for subsets of the study data not related to obstetrics or neonates so that facility comparisons were made on similar types of patients. This was done in two ways:

(1) When analyzing diagnoses (diagnosis code, diagnosis type, and most responsible diagnosis), original codes describing obstetric or neonate conditions were excluded.

(2) When analyzing the grouper output variables (major clinical category, case mix group, complexity assignment, expected length of stay, resource intensity weight), discharges originally assigned to complexity level "9" were excluded. Complexity level "9" is assigned to mental health, obstetric, and neonate patients.

The facility results have also been adjusted to account for the coder effect, when possible. This eliminated the bias due to the variability in discrepancy rates with the original data found between reabstractors. Note that coder effect adjustments were not applied to the findings presented for major clinical category, expected length of stay, or resource intensity weight.

The graphs presented in this section have inset boxes that contain information for "all facilities combined". Here, the combined results have undergone the same coder effect and case mix effect adjustments as the facility results. The facility specific results are compared against this combined estimate when statements are made regarding how the facilities compare to each other. Refer to Appendix F for information on the facility names associated with the facility letters used in this section.

Refer to section 7.4.1.1 and section 8.5 for more information on coder effect.
Refer to section 7.4.1.2 and section 8.6 for more information on case mix effect.

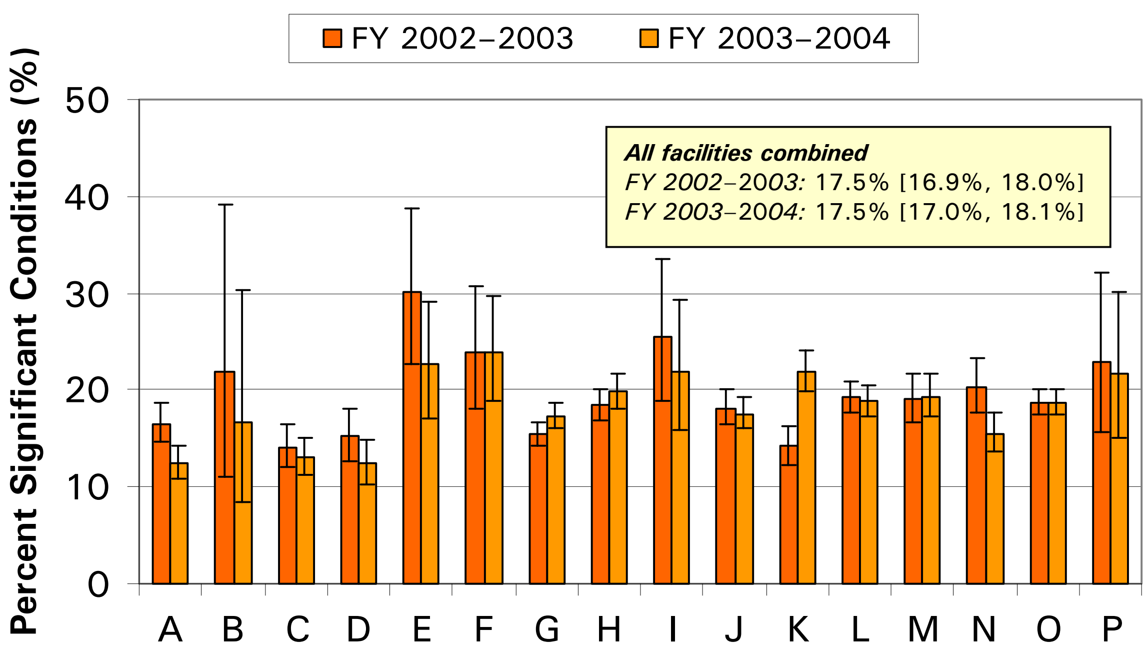
11.1 Selection of Diagnosis Code

Analysis of the selection of diagnosis code was performed for each facility on the subset of conditions identified as significant (i.e. type (M) (1) (2) (W) (X) (Y)) in both the original and reabstracted data. These findings, after adjusting for coder effect and case mix effect, are presented in Figure 11.1.1.

Key findings for selection of diagnosis code, by facility

- In FY 2002–2003, three facilities (E, F, I) had statistically significant higher discrepancy rates than the 17.5% discrepancy rate observed for the entire primary dataset. Three facilities (C, G, K) had statistically significant lower discrepancy rates.
- In FY 2003–2004, three facilities (F, H, K) had statistically significant higher discrepancy rates than the 17.5% discrepancy rate observed for the entire primary dataset. Three facilities (A, C, D) had statistically significant lower discrepancy rates.
- One facility (K) had a significant increase in the discrepancy rate in FY 2003–2004.
- One facility (A) had a significant decrease in the discrepancy rate in FY 2003–2004.

Figure 11.1.1: Diagnosis Code Discrepancy Rates for Significant Conditions, Facility Results after *Adjusting for Coder Effect and Case Mix Effect*



Source: CIHI 2005

Note: The denominator is the weighted sum of all conditions (excluding obstetrical and neonate conditions) that were coded as significant in both the original and reabstracted data.

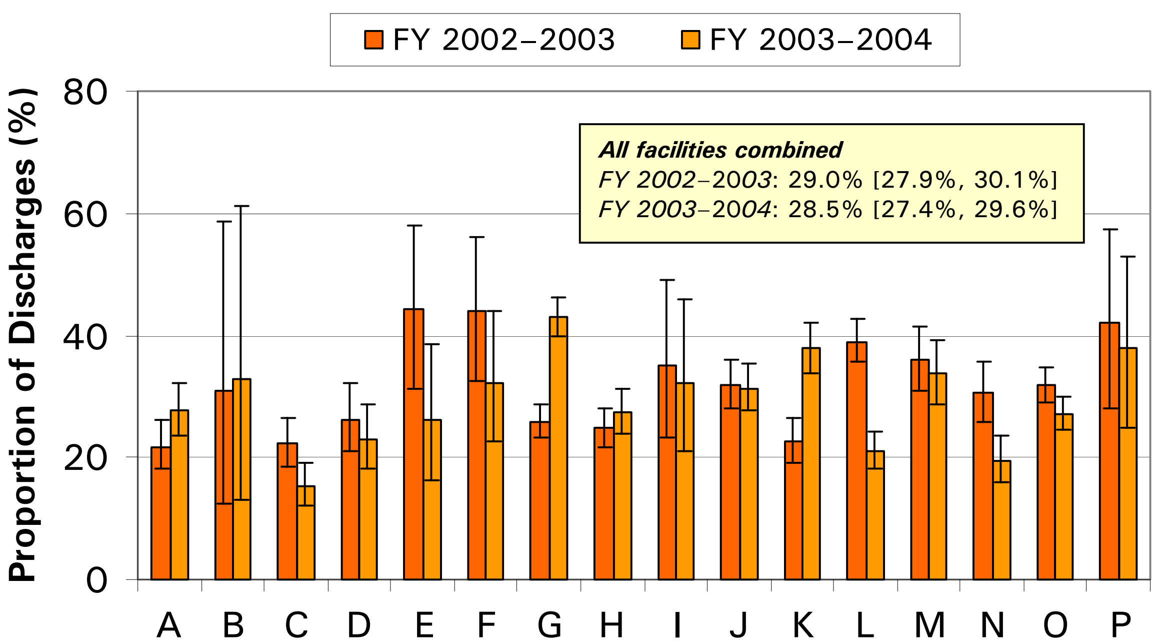
11.2 Selection of Most Responsible Diagnosis

Analysis of code selected as the most responsible diagnosis (MRDx) is done irrespective of whether the original and reabstracted MRDx are describing the same condition. This can be done because only one MRDx code is present in the original data for a discharge, and one MRDx is present in the reabstracted data for that same discharge. Figure 11.2.1 illustrates the proportion of discharges for each facility that had different codes selected as the MRDx upon reabstraction. These results have been adjusted for coder effect and case mix effect.

Key findings for the code selection of the most responsible diagnosis, by facility

- In FY 2002–2003, four facilities (E, F, L, M) had statistically significant higher discrepancy rates than the 29.0% discrepancy rate observed for the entire primary dataset. Three facilities (A, C, K) had statistically significant lower discrepancy rates.
- In FY 2003–2004, two facilities (G, K) had statistically significant higher discrepancy rates than the 28.5% discrepancy rate observed for the entire primary dataset. Three facilities (C, L, N) had statistically significant lower discrepancy rates.
- Two facilities (G, K) had significant increases in the discrepancy rates in FY 2003–2004. Two facilities (L, N) had significant decreases in the discrepancy rates in FY 2003–2004.

Figure 11.2.1: Proportion of Discharges Assigned Different Codes as the Most Responsible Diagnosis upon Reabstraction, Facility Results after *Adjusting for Coder Effect and Case Mix Effect*



Source: CIHI 2005

Note: The denominator for percentages is the weighted sum of all discharges (excluding those with original MRDx codes of obstetrical and neonate conditions) assigned to each facility.

11.3 Major Clinical Category

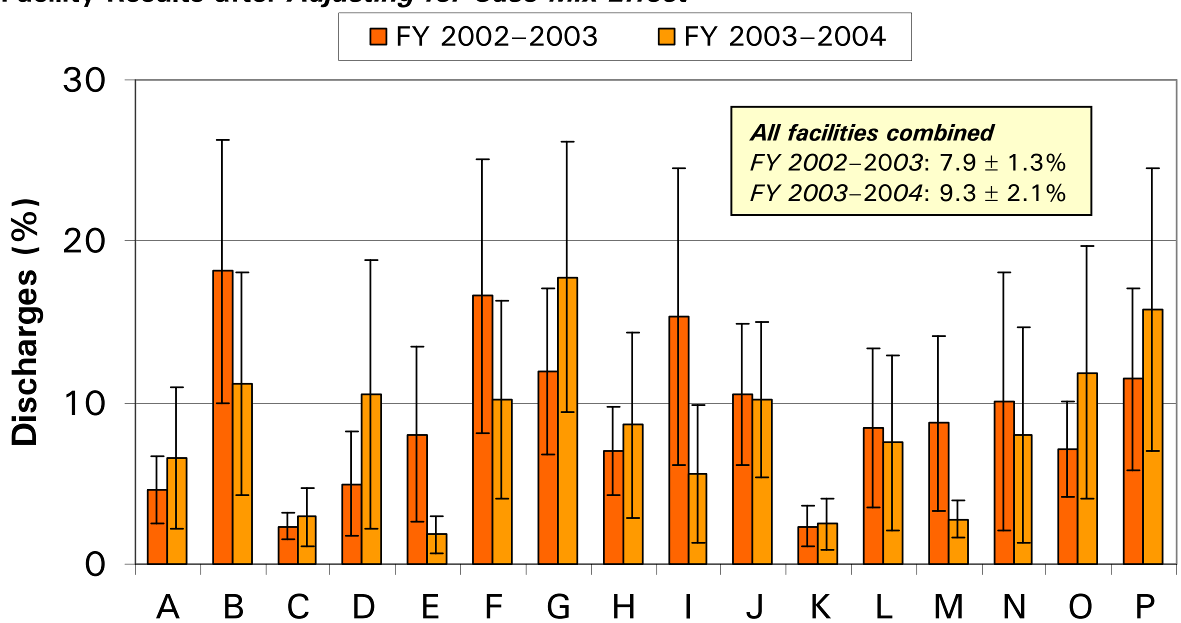
The percent change in major clinical category upon reabstraction was calculated and plotted for each case-costing facility as shown below.

Due to the high agreement rate observed across all facilities for major clinical category, the logistic regression model was unstable at the facility level due to the insufficient sample of records containing discrepancies. For this reason, the coder effect at the facility level could not adequately be assessed for this data element. *The results presented in Figure 11.3.1 have been adjusted for case mix effect only.*

Key findings for the assignment of major clinical category, by facility

- In FY 2002–2003, one facility (B) had a statistically significant higher disagreement rate than the 7.9% rate observed for the entire primary dataset. Three facilities (A, C, K) had statistically significant lower disagreement rates.
- In FY 2003–2004, no facility had a statistically significant higher disagreement rate than the 9.3% observed for the entire primary dataset. Four facilities (C, E, K, M) had significantly lower disagreement rates.
- No facility showed a significant change in the percent change in major clinical category between fiscal years.

Figure 11.3.1: Percent Change in Major Clinical Category upon Reabstraction, Facility Results after *Adjusting for Case Mix Effect*



Source: CIHI 2005

Note: The original and reabstracted data were grouped using the 2003 CIHI grouping methodology
Note: Denominators for percentages are the weighted number of discharges (excluding those originally assigned complexity level 9) for each facility.
Note: These estimates have not been adjusted to account for coder effect.

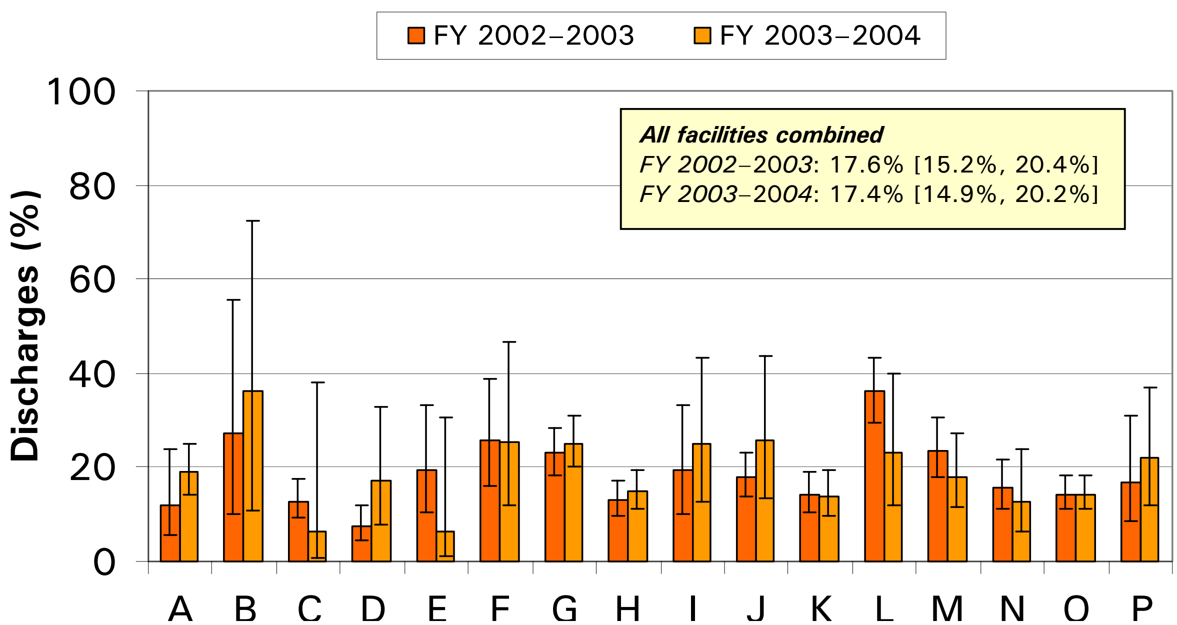
11.4 Case Mix Group

The percent change in case mix group upon reabstraction was calculated and plotted for each case-costing facility as shown below. These results have been adjusted for coder effect and case mix effect.

Key findings for case mix group, by facility

- In FY 2002–2003, one facility (L) had a statistically significant higher disagreement rate for case mix group than the 17.6% disagreement rate observed for the entire primary dataset. One facility (D) had a statistically significant lower disagreement rate for case mix group.
- In FY 2003–2004, no facility had a disagreement rate in major clinical category that was significantly different from the 17.4% disagreement rate observed for the entire primary dataset.
- No facility showed a significant change in the percent change in case mix group between fiscal years.

Figure 11.4.1: Percent Change in Case Mix Group upon Reabstraction, Facility Results after *Adjusting for Coder Effect and Case Mix Effect*



Source: CIHI 2005

Note: The original and reabstracted data were grouped using the 2003 CIHI grouping methodology
Note: Denominators for percentages are the weighted number of discharges (excluding those originally assigned complexity level 9) for each facility.

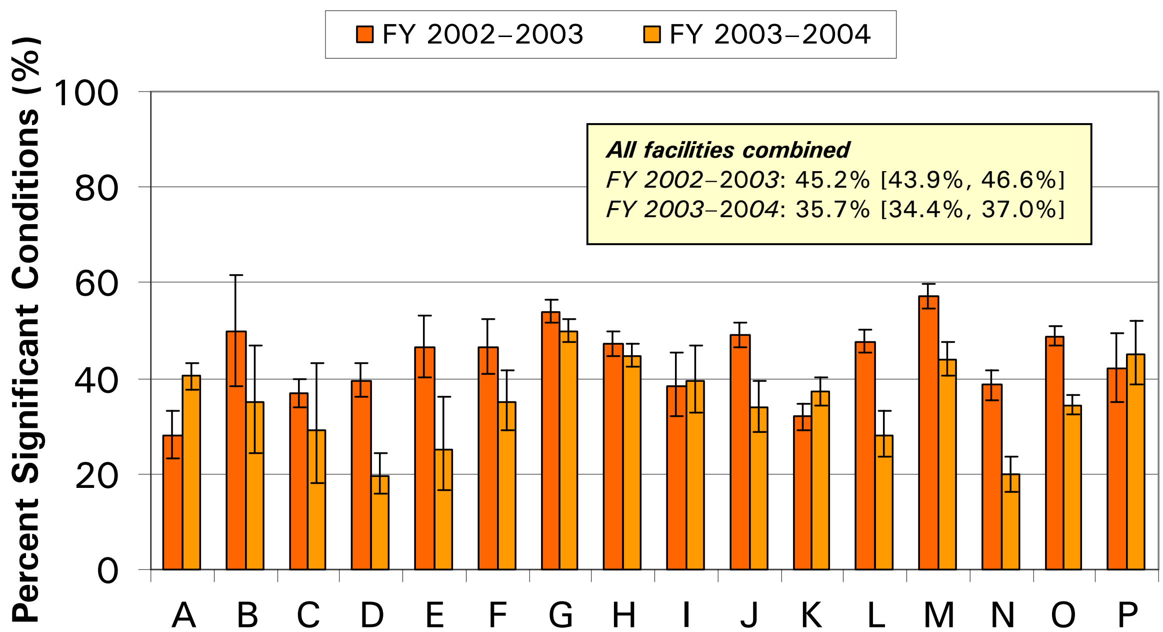
11.5 Assignment of Significance (Diagnosis Type)

Analysis of the assignment of significance to a condition was performed by facility, as presented in Figure 11.5.1. This analysis considers all conditions that were identified as being significant by either the original coder or the reabstractor. The graph plots discrepancy rates, which can comprise either disagreement in the presence of the condition, or disagreement in the assignment of significance. These results have been adjusted for coder effect and case mix effect.

Key findings for the Case Costing significant conditions, by facility

- In FY 2002–2003, four facilities (G, J, M, O) had significantly higher discrepancy rates than the 45.2% observed for all facilities combined. Five facilities (A, C, D, K, N) had significantly lower discrepancy rates.
- In FY 2003–2004, five facilities (A, G, H, M, P) had significantly higher discrepancy rates than the 35.7% observed for all facilities combined. Three facilities (D, L, N) had lower discrepancy rates.
- Discrepancy rates decreased for most facilities in FY 2003–2004, of which many are statistically significant (facilities D, E, J, L, M, N, O).
- The discrepancy rate for one facility (A) significantly increased in FY 2003–2004.

Figure 11.5.1: Discrepancies in the Assignment of Significance to Conditions between the Original and Reabstracted Data, Facility Results after *Adjusting for Coder Effect and Case Mix Effect*



Source: CIHI 2005

Note: The denominator is the weighted sum of all significant conditions (excluding obstetrical and neonate conditions) in the original data and/or the reabstracted data.

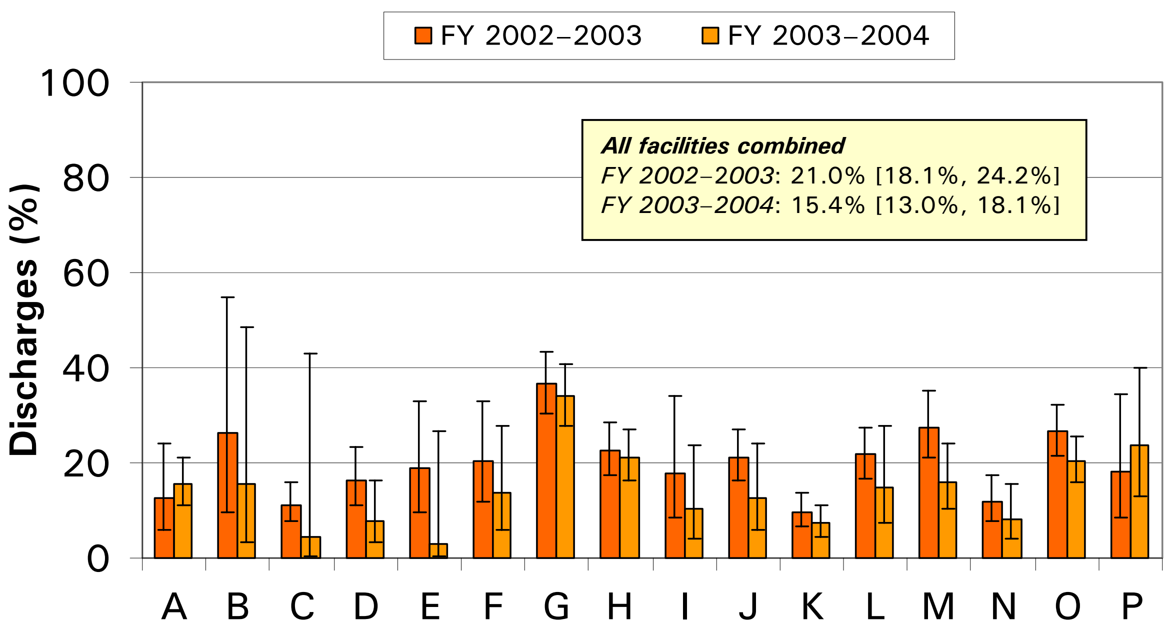
11.6 Complexity Level Assignment

The percent change in complexity assignment upon reabstraction was calculated and plotted for each case-costing facility as shown below. These results have been adjusted for coder effect and case mix effect.

Key findings for complexity level assignment, by facility

- In FY 2002–2003, one facility (G) had a statistically significant higher disagreement rate for complexity level than the 21.0% change observed for the entire primary dataset. Three facilities (C, K, N) have statistically significant lower disagreement rates for complexity level.
- In FY 2003–2004, one facility (G) had a statistically significant higher disagreement rate for complexity level than the 15.4% change observed for the entire primary dataset. One facility (K) had a statistically significant lower disagreement rate for complexity level.
- Though the overall tendency was for a lower discrepancy rate in FY 2003–2004, no facility showed a statistically significant change between fiscal years.

Figure 11.6.1: Percent Change in Complexity Level upon Reabstraction, Facility Results after *Adjusting for Coder Effect and Case Mix Effect*



Source: CIHI 2005

Note: The original and reabstracted data were grouped using the 2003 CIHI grouping methodology.
Note: Denominators for percentages are the weighted number of discharges (excluding those originally assigned complexity level 9) for each facility.

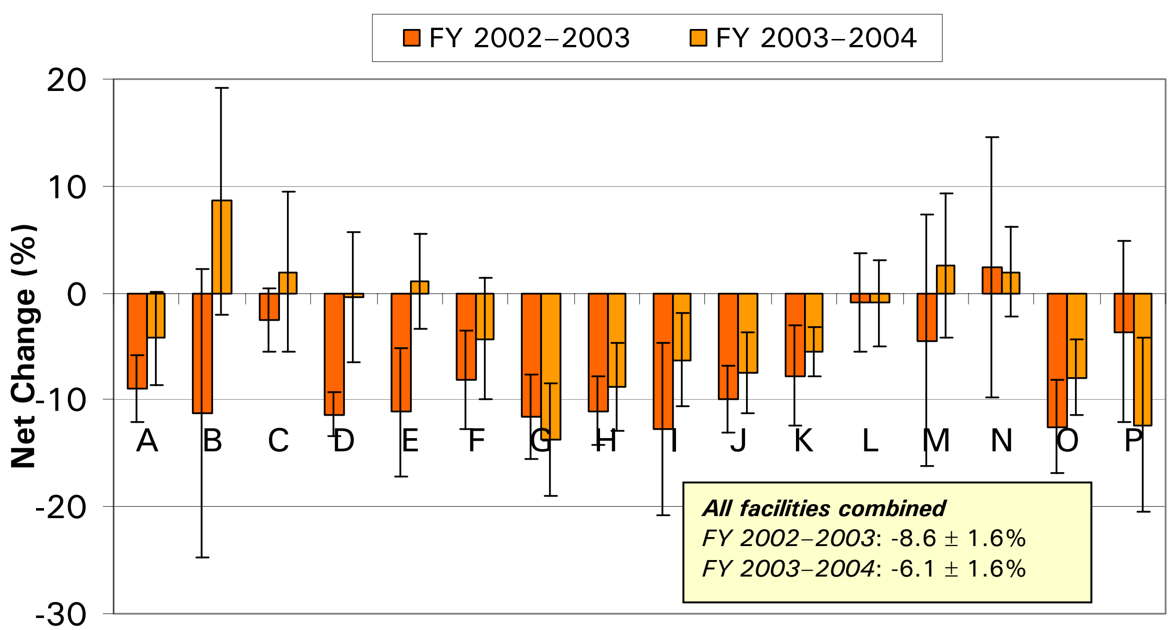
11.7 Expected Length of Stay

The percent net change in expected length of stay (ELOS) upon reabstraction was calculated and plotted for each case-costing facility as shown below. *These findings have been adjusted for case mix effect only.*

Key findings for expected length of stay, by facility

- In FY 2002–2003, ten facilities had percent net changes in ELOS that were significantly different from zero. None of these facilities had a net change in ELOS that was significantly greater than the overall net change of –8.6%. Two facilities (C, L) had a percent net change in ELOS significantly closer to zero than the combined results.
- In FY 2003–2004, seven facilities had percent net changes in ELOS that were significantly different from zero. One facility (G) had a statistically significant greater difference in ELOS than the –6.1% observed for the combined results. Four facilities (B, E, M, N) that had net changes in ELOS that were significantly different that the combined results. However, the net changes in ELOS for these facilities were not significantly closer to zero.
- Two facilities (D, E) had a significant difference in the net change in ELOS between fiscal years, though the results have not accounted for coder effect.

Figure 11.7.1: Percent Net Change in Expected Length of Stay upon Reabstraction, Facility Results *after Adjusting for Case Mix Effect*



Source: CIHI 2005

Note: The original and reabstracted data were grouped using the 2003 CIHI grouping methodology.
Note: Denominators for percentages are the weighted sum of the original ELOS values (excluding discharges not originally assigned to complexity level 9) for each facility.
Note: These estimates have not been adjusted to account for coder effect.

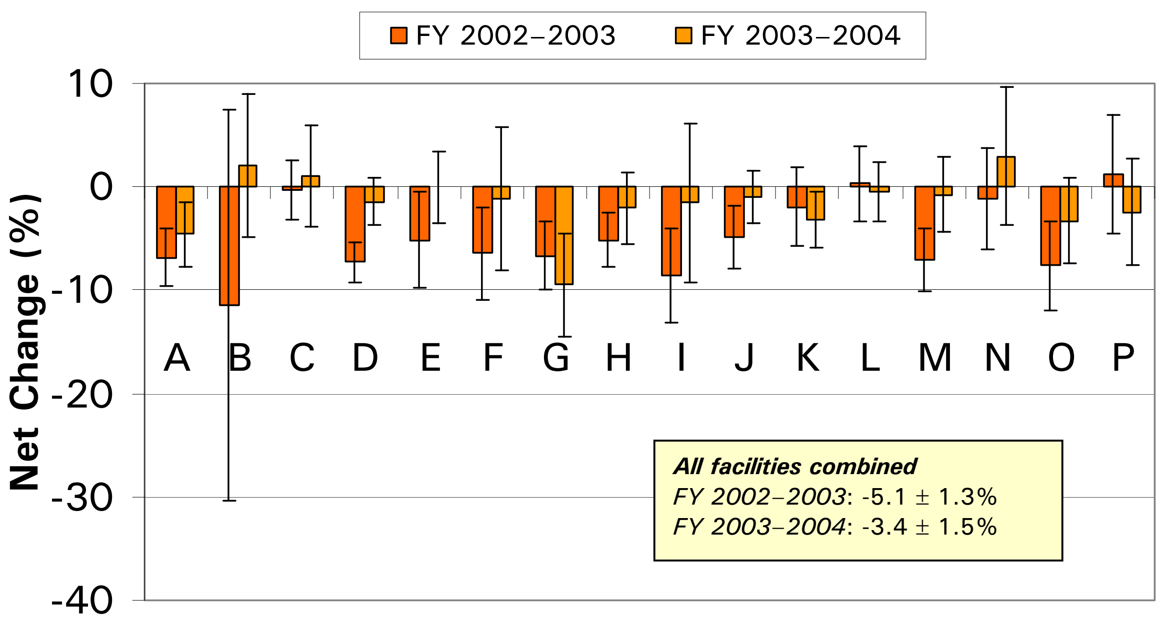
11.8 Resource Intensity Weight

The percent net change in resource intensity weight upon reabstraction was calculated and plotted for each case-costing facility as shown below. *These findings have been adjusted for case mix effect only.*

Key findings for resource intensity weight, by facility

- In FY 2002–2003, ten facilities had percent net changes in RIW values that were significantly different from zero. One facility (C) had a percent net changes in RIW that was significantly closer to zero than the –5.1% net change observed for all facilities combined. Another facility (L) had a significantly different net change in RIW, though the net change was not closer to zero.
- In FY 2003–2004, three facilities (A, G, K) had net changes in RIW values that were significantly different from zero. No facility was significantly different from the –3.4% net change observed for all facilities combined.
- One facility (D) had a significant difference in the net change in RIW between fiscal years, though the results have not accounted for coder effect.

Figure 11.8.1: Percent Net Change in Resource Intensity Weight upon Reabstraction, Facility Results after *Adjusting for Case Mix Effect*



Source: CIHI 2005

Note: The original and reabstracted data were grouped using the 2003 CIHI grouping methodology.
Note: Denominators for percentages are the weighted sum of the original RIW values (excluding discharges not originally assigned to complexity level 9) for each facility.
Note: These estimates have not been adjusted to account for coder effect.

12 Inter-rater Dataset Results

Results in this section are aggregated to all the case-costing facilities. The sample selected for the inter-rater portion of this study was not intended to provide facility specific analysis.

The presentation of the inter-rater analysis consists of two parts:

(1) ***A detailed review of its findings, in a similar manner as the review performed on the primary dataset in section 10.*** Analysis of the data elements in the inter-rater dataset starts with a presentation of the results *without adjustment* for coder effect or case mix effect.

(2) ***A comparison of the inter-rater findings to the primary dataset findings, which includes coder effect and case mix effect, when applicable.***

For the diagnostic and grouper output data elements, the population of reference was redefined to account for case mix effect. The diagnoses associated with obstetric or neonate patients were reabstracted with high agreement in both the inter-rater and primary datasets. If included, these high agreement records can conceal true differences between the datasets.

There is a difference in how the case mix effect was accounted for in each dataset. In the primary dataset, the point of reference is the original data. Here, records *originally* pertaining to obstetric or neonate patients were excluded. In the inter-rater dataset, there is no point of reference. Here, if *either* reabstracted diagnosis code (or discharge) pertained to an obstetric or neonate patient, the record was excluded.

Also, comparisons between the inter-rater and primary datasets accounted for coder effect, when the coder effect was found to be significant.

Refer to section 7.4.1.1 and section 8.5 for more information on coder effect.
Refer to section 7.4.1.2 and section 8.6 for more information on case mix effect.

12.1 Non-medical Data Elements

Discrepancy rates for non-medical data elements between the reabstractors are minimal, as illustrated in Table 12.1.1. One discrepancy occurred for “Health Care Number” on a discharge with a large sample weight, which produced an estimated discrepancy rate of 5.5% for FY 2002–2003. The margin of error associated with the estimate shows that it is not significantly different from 0%.

Table 12.1.1: Discrepancy Counts and Rates for Non-medical Data Elements in the Inter-rater Dataset, *Unadjusted for Coder Effect or Case Mix Effect*

Data Element	DAD FY 2002–2003		DAD FY 2003–2004	
	Count	%	Count	%
Health Care Number	745	5.5 ± 7.7	171	1.1 ± 2.3
Gender	0	0.0 ± 0.0	0	0.0 ± 0.0
Birth Date	0	0.0 ± 0.0	0	0.0 ± 0.0
Birth Date is Estimated	0	0.0 ± 0.0	0	0.0 ± 0.0
Admission Category	137	1.0 ± 0.9	314	2.1 ± 3.4
Admission Date	50	0.4 ± 0.5	18	0.1 ± 0.2
Discharge Disposition	407	3.0 ± 1.9	212	1.4 ± 1.7
Discharge Date	0	0.0 ± 0.0	0	0.0 ± 0.0
Alternate Level of Care Days	204	1.5 ± 0.9	54	0.4 ± 0.3
Total Length of Stay	50	0.4 ± 0.5	18	0.1 ± 0.2
Acute Length of Stay	254	1.9 ± 1.1	72	0.5 ± 0.4
Institution From	152	1.1 ± 0.8	66	0.4 ± 0.3
Institution From Type	104	0.8 ± 0.6	18	0.1 ± 0.1
Institution To	260	1.9 ± 1.8	152	1.0 ± 2.1
Institution To Type	204	1.5 ± 1.5	48	0.3 ± 0.3
Weight	0	0.0 ± 0.0	21	0.1 ± 0.2

Source: CIHI 2005

Note: The denominator for percentages is 13,643 in FY 2002–2003 and 15,295 in FY 2003–2004.
Note: Variations in the version code of Health Care Number are not flagged as a discrepancy.
Note: These results have not been adjusted for coder effect or case mix effect.

12.1.1 Comparison to the Primary Dataset

The following analysis of the non-medical data elements has not been adjusted for coder effect or case mix effect.

There are no statistically significant differences between the primary and inter-rater datasets[15], indicating that the discrepancies are likely a result of unclear or unavailable documentation, or lack of clarity in the DAD Abstracting Manual or the coding standards.

Analysis of discrepancies between reabstractors identified similar issues to the primary dataset results, particularly for "Discharge Disposition".

12.2 Selection of Intervention Code

There is no significant difference between the agreement rates for intervention codes present in both reabstractors data during the two fiscal years of the study. As shown in Table 12.2.1, the agreement rate up to and including the rubric level was over 91% for both years of the study.

[15] See Table 10.1.1 for the equivalent results in the primary dataset.

Table 12.2.1: Comparison of Codes for Mandatory Interventions in the Inter-rater Dataset, *Unadjusted for Coder Effect*

	DAD FY 2002–2003		DAD FY 2003–2004	
	Count	**%**	**Count**	**%**
Mandatory Interventions Present in Both Reabstractors Data	**12,275**		**14,623**	
Exact Match	10,419	84.9 ± 5.4	12,847	87.9 ± 5.8
Rubric Match Only	1,205	9.8 ± 4.6	582	4.0 ± 2.0
Type of Intervention Only	138	1.1 ± 0.8	303	2.1 ± 2.2
Part of Anatomy Only	256	2.1 ± 1.6	710	4.9 ± 4.3
Different	257	2.1 ± 1.8	182	1.2 ± 1.8

Source: CIHI 2005

Reasons for discrepancies between reabstractors in the inter-rater dataset were derived from the reasons assigned to the discrepancies through the primary dataset. That is, if one reabstractor disagreed with an original intervention code due to "chart documentation" but the second reabstractor agreed with the original code, then it was inferred that the two reabstractors disagreed with each other due to chart documentation. Reasons for intervention code discrepancies in the inter-rater dataset are illustrated in Table 12.2.2. Reasons of "chart documentation" and "standards/codebook/manual" were used approximately the same number of times for both fiscal years.

Table 12.2.2: Reasons for Discrepancies in Intervention Codes in the Inter-rater Dataset, *Unadjusted for Coder Effect*

	DAD FY 2002–2003		DAD FY 2003–2004	
	Count	**%**	**Count**	**%**
Total Discrepancies	**1,856**	**15.1**	**1,776**	**12.1**
Standards/Codebook/Manual	873	7.1	632	4.3
Chart Documentation	853	6.9	878	6.0
Acceptable Difference	9	0.1	6	0.0
Mismatched Reasons	89	0.7	241	1.6
Cannot Infer	32	0.3	19	0.1

Source: CIHI 2005

Note: "Acceptable Difference" is the "Optional/Not Wrong" reason code.
Note: "Mismatched Reasons" refers to instances where the two reabstractors assigned different reason codes.
Note: "Cannot Infer" occurs for the instances when the two reabstractors codes were not linked through the original data, and no reason for the discrepancy can be deduced.
Note: The denominator for percentages is 12,275 in FY 2002–2003 and 14,623 in FY 2003–2004.

12.2.1 Comparison to the Primary Dataset

The following analysis of intervention code has been adjusted for coder effect in the primary dataset only. Coder effect was significant for this data element in the primary dataset but not in the inter-rater dataset. Case mix effect does not apply to this data element.

There is no significant difference when comparing results between the primary dataset and the inter-rater dataset. The agreement of code selection in the inter-rater data occurred for 84.9% (FY 2002–2003) and 87.9% (FY 2003–2004) of the interventions. These rates are similar to the results for the primary dataset where agreement rates were 86.1% and 91.4% for the respective fiscal years.

Though the results for the primary dataset found that code selection improved significantly between fiscal years, this was not the case in the inter-rater dataset.

Table 12.2.1.1: Agreement Rates for Intervention Codes in the Inter-rater Dataset and Primary Dataset

	DAD FY 2002–2003		DAD FY 2003–2004	
	%	Confidence Interval	%	Confidence Interval
Primary Dataset *	86.1	[84.1, 88.0]	91.4	[90.1, 92.5]
Inter-rater Dataset	84.9	[79.5, 90.3]	87.9	[82.0, 93.7]

Source: CIHI 2005

Note: *Coder effect adjustments were applied to the primary dataset and not the inter-rater dataset.
Note: The denominator for percentages for the primary dataset is 236,606 in FY 2002–2003 and 274,672 in FY 2003–2004.
Note: The denominator for percentages for the inter-rater dataset is 12,275 in FY 2002–2003 and 14,623 in FY 2003–2004.

The discrepancies observed in both datasets indicate that there is some lack of clarity around the selection of codes used to describe interventions. This is a result of chart documentation that is missing, unclear or contradictory, as well as unclear standards or different interpretation of standards.

12.3 Selection of Diagnosis Code

For the inter-rater dataset, diagnosis codes were compared when both reabstractors identified a condition as significant. Reabstractors agreed on code selection in 80.5% of cases in FY 2002–2003 and 86.4% of the cases in FY 2003–2004, without adjusting for coder effect or case mix effect. The difference between years is not statistically significant.

Table 12.3.1: Diagnosis Code Comparisons for Conditions Identified as Significant by both Reabstractors, *Unadjusted for Coder Effect or Case Mix Effect*

	DAD FY 2002–2003		DAD FY 2003–2004	
	Count	%	Count	%
Significant in Both Reabstractors	**23,651**		**27,141**	
Exact Code Match	19,033	80.5 ± 5.4	23,439	86.4 ± 3.9
Category Match Only	2,438	10.3 ± 3.8	2,144	7.9 ± 3.2
Block Match Only	1,269	5.4 ± 4.0	550	2.0 ± 1.3
Chapter Match Only	442	1.9 ± 1.3	465	1.7 ± 1.1
Chapter Different	469	2.0 ± 0.9	543	2.0 ± 1.5

Source: CIHI 2005

Reasons for discrepancies in diagnosis code selection in the inter-rater dataset are illustrated in Table 12.3.2. Most discrepancies are due to "chart documentation" for both fiscal years, though to a lesser degree in FY 2003–2004.

Table 12.3.2: Reasons for Discrepancies in the Selection of Diagnosis Code in the Inter-rater Dataset, *Unadjusted for Coder Effect or Case Mix Effect*

	DAD FY 2002–2003		DAD FY 2003–2004	
	Count	%	Count	%
Total Discrepancies	**4,617**	**19.5**	**3,702**	**13.6**
Standards/Codebook/Manual	914	3.9	1,225	4.5
Chart Documentation	2,840	12.0	2,158	8.0
Acceptable Difference	6	0.0	38	0.1
Mismatched Reasons	390	1.6	147	0.5
Cannot Infer	467	2.0	133	0.5

Source: CIHI 2005

Note: "Acceptable Difference" is the "Optional/Not Wrong" reason code.
Note: "Mismatched Reasons" refers to instances where the two reabstractors assigned different reason codes.
Note: "Cannot Infer" occurs for the instances when the two reabstractors codes were not linked through the original data, and no reason for the discrepancy can be deduced.
Note: The denominator for percentages is 23,651 in FY 2002–2003 and 27,141 in FY 2003–2004.

12.3.1 Comparison to the Primary Dataset

The following analysis of diagnosis code has been adjusted for coder effect in the inter-rater dataset only. Coder effect was significant for this data element in the inter-rater dataset but not in the primary dataset. Case mix effect has been accounted for by excluding diagnoses pertaining to obstetrical or neonate conditions.

The agreement of code selection in the inter-rater data occurred for 79.0% (FY 2002–2003) and 85.8% (FY 2003–2004) of the diagnoses. These rates are similar to the results for the primary dataset where agreement rates were 82.3% for both fiscal years, after adjusting for case mix effect.

There is no significant difference when comparing the results between the primary dataset and the inter-rater dataset.

Though the results for the primary dataset found that code selection did not change between fiscal years, this was not the case in the inter-rater dataset. ***The 6% increase of the agreement rate in the inter-rater dataset in FY 2003–2004 is statistically significant.***

Table 12.3.1.1: Agreement Rates for Diagnosis Codes in the Inter-rater Dataset and Primary Dataset, *After Adjusting for Case Mix Effect*

	DAD FY 2002–2003		DAD FY 2003–2004	
	%	Confidence Interval	%	Confidence Interval
Primary Dataset	82.3	[80.8, 83.8]	82.3	[80.7, 83.8]
Inter-rater Dataset *	79.0	[76.4, 81.3]	85.8	[83.7, 87.8]

Source: CIHI 2005

Note: *Coder effect adjustments were applied to the inter-rater dataset and not the primary dataset.
Note: The denominators for percentages in the primary dataset are 373,318 in FY 2002–2003 and 340,692 in FY 2003–2004.
Note: The denominators for percentages in the inter-rater dataset are 19,114 in FY 2002–2003 and 20,014 in FY 2003–2004.

The observed discrepancy rates for code selection in both datasets indicate that there is lack of clarity around the selection of codes to identify conditions. This can be attributed to unclear or unavailable documentation, or differing interpretations of the coding standards.

Analysis of the inter-rater data indicated similar coding issues outlined in the primary dataset analysis for code selection.

12.4 Selection of Most Responsible Diagnosis

The inter-rater dataset has a relatively high agreement rate for most responsible diagnoses. In over 85% of the cases, the same code was used to describe the most responsible diagnosis, and the codes matched at least at the category level in over 90% of the cases for both fiscal years, without adjusting for coder effect or case mix effect.

Table 12.4.1: Diagnosis Code Comparisons for Most Responsible Diagnoses Between Reabstractors, *Unadjusted for Coder Effect or Case Mix Effect*

	DAD FY 2002–2003		DAD FY 2003–2004	
	Count	%	Count	%
Total Most Responsible Diagnoses	**13,643**		**15,295**	
Exact Code Match	11,644	85.4 ± 7.0	13,614	89.0 ± 5.4
Category Match Only	1,011	7.4 ± 5.6	467	3.1 ± 2.6
Block Match Only	218	1.6 ± 1.6	349	2.3 ± 3.2
Chapter Match Only	423	3.1 ± 3.5	226	1.5 ± 1.7
Chapter Different	346	2.5 ± 1.6	641	4.2 ± 3.2

Source: CIHI 2005

12.4.1 Comparison to the Primary Dataset

The following analysis of the most responsible diagnosis (MRDx) has not been adjusted for coder effect because it was not significant for this data element. Case mix effect has been accounted for by excluding discharges assigned MRDx codes of obstetrical or neonate conditions.

Agreement of code selection of the MRDx in the inter-rater data occurred for 83.0% (FY 2002–2003) and 89.1% (FY 2003–2004) of the discharges. These rates are higher than the results for the primary dataset, where agreement rates were 70.6% (FY 2002–2003) and 70.8% (FY 2003–2004) when adjusting for case mix.

The agreement rates for the code selected as the most responsible diagnosis are significantly higher in the inter-rater dataset than in the primary dataset. The differences between the two datasets are statistically significant, with an estimated difference of 12% in FY 2002–2003 and 18% in FY 2003–2004.

Table 12.4.1.1: Agreement Rates for the Most Responsible Diagnosis in the Inter-rater Dataset and Primary Dataset, *After Adjusting for Case Mix Effect*

	DAD FY 2002–2003	DAD FY 2003–2004
	%	%
Primary Dataset	70.6 ± 2.6	70.8 ± 2.8
Inter-rater Dataset	83.0 ± 8.5	89.1 ± 5.5

Source: CIHI 2005

Note: These results have not been adjusted for coder effect.
Note: The denominators for percentages in the primary dataset are 192,218 in FY 2002–2003 and 188,315 in FY 2003–2004.
Note: The denominators for percentages in the inter-rater dataset are 10,504 in FY 2002–2003 and 11,645 in FY 2003–2004.

Variation in the codes selected as the most responsible diagnosis between reabstractors indicated similar issues to those in the primary dataset. In particular, this is noted for: symptom versus underling condition; palliative care; and neoplasm versus other condition (possibly complication of neoplastic disease).

Discrepancies that resulted in most responsible diagnosis codes from different chapters are mostly due to different interpretation of chart documentation, or incomplete or conflicting information in the chart.

12.5 Major Clinical Category

Discharges in the inter-rater dataset were grouped to the same major clinical category with both reabstractors' data in 93.7% (FY 2002–2003) and 95.3% (FY 2003–2004) of the cases, without adjusting for coder effect or case mix effect.

12.5.1 Comparison to the Primary Dataset

The following analysis of MCC values has not been adjusted for coder effect because it was not significant for this data element. Case mix effect has been applied by excluding discharges that were assigned complexity level 9.

The agreement of MCC assignment in the inter-rater data occurred for 91.6% and 93.3% of the cases. This is in line with the results from the primary dataset where the agreement rates were 92.1% (FY 2002–2003) and 90.7% (FY 2003–2004), after adjusting for case mix effect.

There were no significant differences between the results of the inter-rater and primary datasets.

Table 12.5.1.1: Agreement Rates for Major Clinical Category in the Inter-rater Dataset and Primary Dataset, *After Adjusting for Case Mix Effect*

	DAD FY 2002–2003	DAD FY 2003–2004
	% of Records	% of Records
Primary Dataset	92.1 ± 1.3	90.7 ± 2.1
Inter-rater Dataset	91.6 ± 5.2	93.3 ± 4.6

Source: CIHI 2005

Note: These results have not been adjusted for coder effect.
Note: The denominators for percentages in the primary dataset are 179,802 in FY 2002–2003 and 175,003 in FY 2003–2004.
Note: The denominators for percentages in the inter-rater dataset are 13,643 in FY 2002–2003 and 15,295 in FY 2003–2004.

12.6 Case Mix Group

Discharges in the inter-rater dataset were assigned to different case mix groups upon reabstraction in 15.0% of the cases in FY 2002–2003 and 11.7% of the cases in FY 2003–2004, without adjusting for coder effect or case mix effect.

12.6.1 Comparison to the Primary Dataset

The following analysis of CMG assignment has not been adjusted for coder effect because it was not significant for this data element. Case mix effect has been applied by excluding discharges that were assigned complexity level 9.

The agreement of CMG group in the inter-rater data occurred for 86.5% (FY 2002–2003) and 87.5% (FY 2003–2004) of the cases. This is in line with the results from the primary dataset where the agreement rates were 81.4% (FY 2002–2003) and 81.7% (FY 2003–2004), after adjusting for case mix effect.

Though agreement rates are higher in the inter-rater dataset, the results are not significantly different from the rates observed in the primary dataset.

Table 12.6.1.1: Agreement Rates for Case Mix Group in the Inter-rater Dataset and Primary Dataset, *After Adjusting for Case Mix Effect*

	DAD FY 2002–2003	DAD FY 2003–2004
	% of Records	% of Records
Primary Dataset	81.4 ± 2.2	81.7 ± 2.6
Inter-rater Dataset	86.5 ± 6.0	87.5 ± 6.2

Source: CIHI 2005

Note: These results have not been adjusted for coder effect.
Note: The denominators for percentages in the primary dataset are 179,802 in FY 2002–2003 and 175,003 in FY 2003–2004.
Note: The denominators for percentages in the inter-rater dataset are 13,643 in FY 2002–2003 and 15,295 in FY 2003–2004.

12.7 Selection of Diagnosis Type

In order to facilitate the inter-rater analysis of the diagnosis typing, the following tables were prepared in a manner similar to tables presented for the primary dataset. Diagnosis types assigned by the two reabstractors are presented as a diagnosis pair. This means that the discrepancies appear twice in this table: one time above the diagonal, and one time below the diagonal.

The match rate on the assignment of a condition as the MRDx in the inter-rater dataset is 83% in FY 2002–2003 and 86% in FY 2003–2004 whereas it is 90% in the primary dataset for both fiscal years. Similar rates were seen in the assignment of co-morbidities (types (1) and (2)) between both datasets: 55% and 54% (inter-rater) compared to 52% and 55% (primary) for FY 2002–2003; with a slight improvement in FY 2003–2004: 61% and 57% (inter-rater) and 57% and 61% (primary). These figures are not adjusted for coder effect or case mix effect. As was demonstrated in the primary dataset, many of the co-morbidities identified in the inter-rater dataset show a large discrepancy where one reabstractor assigned the condition as significant and the other as a secondary condition.

The last column in the table represents conditions that were reabstracted by only one of the two reabstractors. The second reabstractor either did not code the condition because it was felt the condition did not meet the requirements of significance, or it was not identified on the chart at all.

Table 12.7.1: Diagnosis Type Comparison Between Reabstractors in FY 2002–2003, *Unadjusted for Coder Effect or Case Mix Effect*

Diagnosis Type		Diagnosis Type							
		Significant				Secondary			
		M	1	2	W,X,Y	3	0	9	Not Reabstracted
Significant	M	11,894 83%	2,008 14%	38 0%	31 0%	360 3%	0 0%	0 0%	1,060
	1	2,008 15%	7,240 55%	379 3%	16 0%	3,409 26%	0 0%	0 0%	4,504
	2	38 1%	379 12%	1,733 54%	9 0%	1,053 33%	0 0%	0 0%	2,458
	W,X,Y	31 8%	16 4%	9 2%	302 83%	6 2%	0 0%	0 0%	138
Secondary	3	360 3%	3,409 33%	1,053 10%	6 0%	5,581 54%	0 0%	9 0%	4,941
	0	0 0%	0 0%	0 0%	0 0%	0 0%	291 100%	0 0%	10
	9	0 0%	0 0%	0 0%	0 0%	9 0%	0 0%	2,647 100%	1,254

Source: CIHI 2005

Note: Discrepancies appear twice in this table, once above (in grey) and once below the diagonal.
Note: "Not Reabstracted" indicates that one of the two reabstractors did not code the condition because it did not meet the requirements of significance, or because the condition was not identified.

Table 12.7.2: Diagnosis Type Comparison Between Reabstractors in FY 2003–2004, *Unadjusted for Coder Effect or Case Mix Effect*

Diagnosis Type		Diagnosis Type							
		Significant				Secondary			
		M	1	2	W,X,Y	3	0	9	Not Reabstracted
Significant	M	13,856 86%	1,869 12%	71 0%	15 0%	340 2%	0 0%	0 0%	584
Significant	1	1,869 13%	8,665 61%	481 3%	27 0%	3,104 22%	0 0%	0 0%	4,082
Significant	2	71 2%	481 14%	1,895 57%	0 0%	884 27%	0 0%	0 0%	1,703
Significant	W,X,Y	15 5%	27 8%	0 0%	262 83%	14 4%	0 0%	0 0%	306
Secondary	3	340 3%	3,104 28%	884 8%	14 0%	6,589 60%	0 0%	0 0%	4,516
Secondary	0	0 0%	0 0%	0 0%	0 0%	0 0%	491 100%	0 0%	80
Secondary	9	0 0%	0 0%	0 0%	0 0%	0 0%	0 0%	2,627 100%	1,176

Source: CIHI 2005

Note: Discrepancies appear twice in this table, once above (in grey) and once below the diagonal.
Note: "Not Reabstracted" indicates that one of the two reabstractors did not code the condition because it did not meet the requirements of significance, or because the condition was not identified.

Table 12.7.3 illustrates that most discrepancies in diagnosis type selection in the inter-rater dataset are due to the application of "significance" (10.9% and 9.5%). The second most common reason is "chart documentation" (4.2% and 3.2%).

Table 12.7.3: Reasons for Diagnosis Typing Discrepancies in the Inter-rater Dataset, *Unadjusted for Coder Effect or Case Mix Effect*

	DAD FY 2002–2003		DAD FY 2003–2004	
	Count	%	Count	%
Total Discrepancies	**7,326**	**19.8**	**6,804**	**16.5**
Standards/Codebook/Manual	716	1.9	1,053	2.6
Significance	4,047	10.9	3,928	9.5
Chart Documentation	1,551	4.2	1,312	3.2
Acceptable Difference	36	0.1	104	0.3
Mismatched Reasons	486	1.3	296	0.7
Cannot Infer	491	1.3	111	0.3

Source: CIHI 2005

Note: "Acceptable Difference" is the "Optional/Not Wrong" reason code.
Note: "Mismatched Reasons" refers to instances where the two reabstractors assigned different reason codes.
Note: "Cannot Infer" occurs for the instances when the two reabstractors codes were not linked through the original data, and no reason for the discrepancy can be deduced.
Note: The denominator for percentages is 37,008 in FY 2002–2003 and 41,190 in FY 2003–2004.

The assignment of significance to a condition was not clearly understood by the reabstractors. This can be the result of unclear chart documentation, differing interpretation of chart documentation, and lack of clarity about the interpretation of significance.

12.7.1 Comparison to the Primary Dataset

The following analysis of the assignment of significance to a condition has not been adjusted for coder effect because it was not significant for this data element. Case mix effect has been accounted for by excluding diagnoses pertaining to obstetrical or neonate conditions.

The agreement in the assignment of significance to a condition in the inter-rater data occurred for 60.7% and 67.6% of the conditions. These results are higher than the results from the primary dataset where the agreement rates were 55.8% (FY 2002–2003) and 60.5% (FY 2003–2004), when adjusting for case mix effect.

The rates of agreement for significance are higher in the inter-rater dataset than in the primary dataset. The higher rate observed in the inter-rater dataset is statistically significant in FY 2003–2004, with an estimated difference of 7%.

Table 12.7.1.1: Agreement Rates for the Assignment of Significance to Conditions in the Inter-rater Dataset and Primary Dataset, *After Adjusting for Case Mix Effect*

	DAD FY 2002–2003	DAD FY 2003–2004
	%	%
Primary Dataset	55.8 ± 1.2	60.5 ± 1.4
Inter-rater Dataset	60.7 ± 5.7	67.6 ± 5.2

Source: CIHI 2005

Note: These results have not been adjusted for coder effect.
Note: The denominators for percentages in the primary dataset are 669,474 in FY 2002–2003 and 562,689 in FY 2003–2004.
Note: The denominators for percentages in the inter-rater dataset are 31,483 in FY 2002–2003 and 29,622 in FY 2003–2004.

Analysis of inter-rater data revealed similar diagnosis typing discrepancies outlined in the primary dataset findings for:

- Post-procedural conditions and complications
- Metabolic disorders
- Genitourinary disorders (renal failure)

Additionally, inter-rater typing discrepancies were noted for other genitourinary conditions including urinary tract infection. On several occasions, urinary tract infection was coded as significant by one reabstractor and secondary by the other reabstractor. Other typing discrepancies included anemia. There were instances of different anemia codes assigned including D63.0* "anemia in neoplastic disease" as a type (3) and coded by the second reabstractor to D64.9 "anemia, unspecified" as a type (1). Lastly, symptoms were also identified as having typing discrepancies between reabstractors. The symptom having the greatest discrepancy was R41.0 "disorientation, unspecified".

Analysis of inter-rater data in which a condition was reabstracted as significant by one coder and not present by the other reabstractor revealed similar findings to the primary dataset. The following codes had the highest variation for diagnosis typing:

- Z51.5 "Palliative care"
- Z75.1 "Person awaiting admission to adequate facility elsewhere"
- D50-D64 "Anemia"

12.8 Complexity Level Assignment

The rate of disagreement on complexity level in the inter-rater dataset was 13.0 ± 4.7% in FY 2002–2003 and 8.9 ± 3.9% in FY 2003–2004. These percentages include discharges originally assigned complexity level 9. If excluding these discharges, complexity levels changed upon reabstraction for 17.8 ± 6.2% (FY 2002–2003) and 12.8 ± 6.0% (FY 2003–2004) of the discharges.

12.8.1 Comparison to the Primary Dataset

The following analysis of complexity assignment has not been adjusted for coder effect because it was not significant for this data element. Case mix effect has been applied by excluding discharges that were assigned complexity level 9.

The agreement of complexity assignment in the inter-rater data is 82.2% and 87.2%. This is in line with the results from the primary dataset where the agreement rates were 80.0% (FY 2002–2003) and 81.6% (FY 2003–2004), after adjusting for case mix effect[16].

Though agreement rates are higher in the inter-rater dataset, the results are not significantly different from the rates observed in the primary dataset.

Table 12.8.1.1: Agreement Rates for Complexity Assignment in the Inter-rater Dataset and Primary Dataset, *After Adjusting for Case Mix Effect*

	DAD FY 2002–2003	DAD FY 2003–2004
	% of Records	% of Records
Primary Dataset	80.0 ± 1.3	81.6 ± 2.1
Inter-rater Dataset	82.2 ± 6.2	87.2 ± 6.0

Source: CIHI 2005

Note: These results have not been adjusted for coder effect.
Note: The denominators for percentages in the primary dataset are 179,802 in FY 2002–2003 and 175,003 in FY 2003–2004.
Note: The denominators for percentages in the inter-rater dataset are 13,643 in FY 2002–2003 and 15,295 in FY 2003–2004.

[16] When adjusting the primary dataset results for both case mix effect and coder effect, there is a statistically significant increase in the agreement rates between fiscal years, as illustrated in section 10.8.1. These values are also not significantly different from the findings in the inter-rater dataset.

12.9 Expected Length of Stay

Analysis of change in expected length of stay for the inter-rater dataset was limited due to there being no clear point of reference (i.e. no "original" value). Calculations of the percent absolute difference were generated for both the primary and inter-rater datasets. No statistically significant difference was found between fiscal years or between the primary and inter-rater datasets.

12.10 Resource Intensity Weight

Analysis of change in resource intensity weight for the inter-rater dataset was limited due to there being no clear point of reference (i.e. no "original" value). Calculations of the percent absolute difference were generated for both the primary and inter-rater datasets. No statistically significant difference was found between fiscal years or between the primary and inter-rater datasets.

13 Summary of Key Findings

This section presents the figures for the main dataset that have been adjusted for ***coder effect.*** These adjusted estimates are presented for all data elements with the exceptions of non-medical, Expected Length of Stay, and Resource Intensity Weight. Some detailed analysis (e.g. selected medical conditions, inter-rater) could not be assessed for coder effect due to an insufficient sample of discrepancies. If figures are presented in this section that have not been adjusted for coder effect, it is explicitly stated.

The inter-rater and facility specific results have also been modified to account for the ***case mix effect***.

13.1 Non-medical Data Elements

Very few discrepancies were found in the non-medical data elements in both the primary and inter-rater datasets, with noted improvements in FY 2003–2004.

Admission Category, Discharge Disposition, Institution To/Type and Institution From/Type are the non-medical data elements that require the greatest attention on coding. Issues related to chart documentation were found to be the reason for most of these discrepancies.

13.2 Selection of Intervention Code

The agreement rate for the intervention codes was 86% in FY 2002–2003 and 91% in FY 2003–2004. The 5% increase in the agreement rate is statistically significant. About half of the discrepancies in intervention coding are attributed to chart documentation, indicating incomplete or conflicting information, differences in interpretation, information missed or specificity of condition not supported. Improvements to chart documentation are a necessary requirement to facilitate coding accuracy. Physicians require continuing education on documentation requirements to meet the level of specificity in CCI.

Analysis of the inter-rater showed similar findings as those for the main study dataset. However, the difference in agreement rates between fiscal years in the inter-rater dataset was not significant.

13.3 Selection of Diagnosis Code

Analysis of the significant conditions in the primary dataset resulted in about 86% codes matching exactly after reabstraction. This was found in both fiscal years studied. Additional conditions were found that matched at the category level in both fiscal years.

More than half of the discrepancies in code selection were attributed to chart documentation, indicating incomplete or conflicting information, differences in interpretation, information missed or specificity of condition not supported. Improvements to chart documentation are necessary to facilitate changes in coding accuracy.

Detailed analysis of the fifteen medical conditions studied was not adjusted for coder effect. This review found that blood disorders, diabetes mellitus, and respiratory conditions have a higher discrepancy rate in both study years than the combined results. This was true for injuries in FY 2003–2004 only.

In contrast, metabolic disorders, neonates, obstetrics and other factors influencing health status have lower discrepancy rates in both study years. Digestive conditions and symptoms had lower rates in FY 2003–2004 only.

These findings highlight specific conditions where the standards need to be reviewed by the health information professionals at the hospitals to ensure the standards are clearly understood. Health information professionals are encouraged to contact CIHI through their client support representatives, submit questions to CIHI via the coding query database, and attend educational workshops when possible.

The analysis of the data by case-costing hospital illustrated that three facilities (E, F, I) had percent agreements on code selection that were significantly lower than the agreement rate observed for the entire primary dataset in FY 2002–2003. In FY 2003–2004, three facilities (F, H, K) were found to have statistically significant lower agreement rates than the rate observed for the entire primary dataset.

Analysis of the inter-rater dataset showed that there was a significant increase in the agreement rate for diagnosis code selection between fiscal years. Though the primary dataset showed similar agreement rates to the inter-rater dataset, this improvement between fiscal years was not observed in the primary dataset.

CIHI has recently undertaken various initiatives to improve the accuracy of clinical coding. The existing standards are being reviewed and revised with the expectation that a new version will be available in FY 2006/2007. New educational workshops will be developed to offer further clarification on the existing and the revised standards. Also, in version 2006 of ICD-10-CA/CCI, the 6th digit will be removed from diabetes mellitus codes as a means to improve the coding of this condition.

13.4 Selection of the Most Responsible Diagnosis

A review of the Case Costing data found that for about 76% (FY 2002–2003) and 74% (FY 2003–2004) of the discharges, the reabstractor selected the exact same code to describe the most responsible diagnosis (MRDx) as the original coder. A notable proportion of discharges had original and reabstracted MRDx codes belonging to different ICD-10-CA chapters, which most likely are describing different conditions.

The analysis of the data by case-costing hospital found that four facilities (E, F, L, M) in FY 2002–2003 had agreement rates on most responsible diagnosis that were significantly lower than the rate observed for the entire primary dataset. In FY 2003–2004, two facilities (G, K) had agreement rates that were significantly lower.

The assignment for the most responsible diagnosis in the inter-rater dataset revealed that reabstractors have a higher agreement rate between each other for exact code match of the MRDx codes than the rate found in the primary dataset. The difference in agreement rates was 12% in FY 2002–2003 and 18% in FY 2003–2004.

13.5 Major Clinical Category

When applying the 2003 CIHI grouping methodology to both fiscal years of data, the major clinical category (MCC) changed upon reabstraction for about 5% of the discharges in both fiscal years.

A review of the changes showed that certain medical conditions had high agreement rates after reabstraction. Analysis of specific major clinical categories was performed on the unadjusted estimates. High agreement rates were observed for: cardio and vascular diseases of the circulatory system, pregnancy and child birth, and newborns and neonates. In contrast, the following MCC assignments have lower agreement rates (less than 90%) for both fiscal years: skin and subcutaneous and breast; endocrine, nutrition and metabolic disorders; lymph, leukemia and neoplasms unspecified; multi system/unspecified site infection; injuries, poison and toxic effects; and other reasons for hospitalization.

The analysis of the data by case-costing hospital could not be adjusted for coder effect due to an insufficient sample of records containing discrepancies. Little variation in the agreement rates was observed in the assignment of MCC by case-costing facility. One facility (B) in FY 2002–2003 had a statistically significant lower agreement rate upon reabstraction. No facility had a significantly lower agreement rate in FY 2003–2004.

There were no significant differences in the agreement rates observed between fiscal years or between the results of the inter-rater and primary datasets.

13.6 Case Mix Group

When applying the 2003 CIHI grouping methodology to both fiscal years of data, the case mix group (CMG) changed upon reabstraction for about 15% of the discharges in both fiscal years.

A review of the changes in 18 of the 478 CMG groups showed that certain medical conditions had lower agreement rates after reabstraction, particularly the CMG groups representing caesarean births. Many of these agreement rates improve in FY 2003–2004. Other CMG groups showed a notable improvement in agreement rates between the two fiscal years. The agreement rate for "heart failure" increased by over 20%. These observations have not accounted for coder effect.

Little variation in the agreement rates was observed in the assignment of CMG by case-costing facility. One facility (L) in FY 2002–2003 had a statistically significant lower agreement rate upon reabstraction. No facility had a significantly lower agreement rate in FY 2003–2004.

There were no significant differences in the agreement rates observed between fiscal years or between the results of the inter-rater and primary datasets.

13.7 Selection of Diagnosis Type

Analysis of the Case Costing data found results similar to previous studies: diagnoses originally typed as a significant condition affecting the patient's length of stay and use of hospital resources, were reabstracted as a secondary condition. Of the conditions typed as significant by the original coder and/or reabstractor, 62% were deemed to be significant by both in FY 2002–2003. This significantly increased to 67% of the conditions in FY 2003–2004. Most of the discrepancies associated with the assignment of significance are attributed to co-morbid conditions both pre and post-admit being reabstracted as secondary conditions.

A decrease in the number of original diagnoses that were not identified by the reabstractor was observed in FY 2003–2004. Without adjusting for coder effect, the diagnoses originally typed as a pre-admit co-morbidity had a significant decrease of 5% of conditions that were not reabstracted. A decrease of 9% was observed for post-admit co-morbidities.

Detailed analysis of the diagnosis typing for the selected medical conditions also did not account for coder effect. This analysis revealed that for both fiscal years, the assignment of the most responsible diagnosis for injuries was typed with a significantly higher agreement than was the case for other conditions. Also, neonates and obstetrics were typed well for pre-admit co-morbidities, and post-procedural complications were coded well for post-admit co-morbidities.

Several conditions were typed with a significantly lower agreement rate than the results for all medical conditions across both fiscal years. The most responsible diagnosis was poorly assigned for neoplasms; pre-admit co-morbidities were typed poorly for symptoms, and post-admit co-morbidities were typed poorly for metabolic disorders and symptoms.

The analysis of the data by case-costing hospital found that four facilities (G, J, M, O) had percent agreements on the assignment of significance to a condition that were significantly lower than the agreement rate observed for the entire primary dataset. In FY 2003–2004, five facilities (A, G, H, M, P) had percent agreements that were significantly lower than the agreement rate observed for the entire primary dataset.

The observations on the coding issues relating to the mistyping of diagnoses clearly identified two main reasons for these coding inaccuracies. One relates to standards that are specific about when a condition is deemed to be significant. These are not being followed well for some conditions (e.g. neoplasms, symptoms, circulatory conditions, metabolic disorders). The second reason is related to chart documentation that is being interpreted differently by the reabstractor. Clearer descriptions of the patient's condition would permit the coders to apply the standards more consistently.

Coding issues related to significant conditions present only in the original data and not reabstracted were mainly attributed to issues related to chart documentation. Also, when analyzing the significant conditions that were present in the reabstracted data only, chart documentation was again cited most often for the discrepancies. Reabstractors indicated that the information in the chart was either incomplete or conflicting, or that there were differences with its interpretation, or that information was missed.

Low agreement rates in the assignment of significance to a condition in the inter-rater dataset confirm that the identification of the significance of a condition is not clearly understood by the coders, that it is difficult to interpret from the chart documentation, and that it may itself not be well defined. Despite the low agreement rates, the inter-rater agreement rate was significantly higher than that observed for the main study dataset in FY 2003–2004, with an estimated difference of 7%. This analysis only accounted for case mix effect, and not coder effect.

CIHI undertook several initiatives to address the diagnosis typing issues in FY 2005/2006: 25 classification edits were implemented in the DAD, in some cases to directly improve the quality of the diagnosis typing; the definition of the proxy MRDx, type (6), was also introduced; and finally, validity edits were put in place on 194 codes which can only be used as secondary diagnoses.

In FY 2006/2007, a number of existing standards will be reformatted to be more user-friendly and more easily understood in order to emphasize the coding directives.

13.8 Complexity Level Assignment

Applying the 2003 CIHI grouping methodology to both fiscal years of data, complexity levels changed upon reabstraction for 13% of the discharges in FY 2002–2003 and 10% of the discharges in FY 2003–2004. The difference between fiscal years is not statistically significant. Where there are discharges that do not match on complexity level, reabstracted complexity levels generally are lower than those assigned originally.

These percentages include discharges originally assigned a complexity level 9. Complexity level 9 indicates that a complexity overlay was not applied, and is assigned to obstetrical, neonate, and mental health discharges. If excluding these discharges, complexity levels differ from one another 21% of the time in FY 2002–2003 and 15% in FY 2003–2004. The smaller proportion of discharges that changed complexity in FY 2003–2004 than in the previous year was statistically significant.

Further analysis of the discharges was performed by facility. Here, only one facility (G) was found to have a lower agreement for complexity upon reabstraction. This was the case for both fiscal years. Three facilities (C, K, N) had statistically significant higher agreement in complexity levels upon reabstraction in FY 2002–2003, with only facility (K) continuing to have a higher agreement rate in the subsequent fiscal year.

There were no statistically significant differences between the fiscal years in the inter-rater dataset. Nor were there any significant differences between the inter-rater dataset and the primary dataset.

13.9 Expected Length of Stay

The percent net change in the expected length of stay (ELOS) upon reabstraction was −7.3% in FY 2002–2003 and −4.5% in FY 2003–2004. The improvement in ELOS values observed in FY 2003–2004 is statistically significant. However, for both fiscal years, reabstracted data consistently resulted in a lower ELOS value than that calculated using the original data. Note that coder effect was not assessed for this data element.

The analysis of the data by case-costing hospital accounted only for case mix effect. Still, the analysis of ELOS values by case-costing facility revealed that ten facilities in FY 2002–2003 had percent net changes in ELOS values that were significantly different from zero. Two facilities (C, L) had net changes that were significantly closer to zero than the rate observed for the combined results. In FY 2003–2004, seven facilities had percent net changes in ELOS values that were significantly different from zero. Only one of these seven facilities (G) had a statistically significant greater difference in ELOS values than the combined results.

13.10 Resource Intensity Weight

The percent net change in resource intensity weight (RIW) upon reabstraction was −4.3% in FY 2002–2003 and −2.8% in FY 2003–2004. The observed difference between fiscal years is not statistically significant. For both fiscal years, reabstracted data consistently resulted in a lower RIW value than that calculated using the original data. Note that coder effect was not assessed for this data element.

The analysis of the data by case-costing hospital accounted only for case mix effect. Still, the analysis of RIW values by case-costing facility revealed that ten facilities in FY 2002–2003 had percent net changes in RIW values that were significantly different from zero. In FY 2003–2004, three facilities (A, G, K) had percent net changes in RIW values that were significantly different from zero. None of these facilities had a net change in RIW value that was statistically significant greater difference than the combined results.

13.11 Coder Effect

The influence the reabstractors have on the study results is called the coder effect. The Case Costing reabstractors, as mentioned earlier, are subject to the same influences as other health information professionals in terms of how they view and code the source data.

The process used to select the reabstractors for the study attempted to address these influences as much as possible by requiring the candidates to have several years of coding experience, experience coding in ICD-10-CA and CCI in particular, experience coding at a tertiary care centre, and attendance at specific CIHI educational workshops. The reabstractors in the study were also required to attend a one-week training session and to receive a passing score on the inter-rater test.

Despite these attempts to ensure similar coding practices between the reabstractors in the study, differences still exist.

- Facility specific results were most prone to coder effect. As a result, all the facility specific analysis presented has been adjusted for coder effect.

Aggregate findings presented when combining all the facility results were less prone to coder effect, with some exceptions.

- After adjusting for coder effect, intervention coding saw an increase in the agreement rate in FY 2003–2004.
- The typing of co-morbidities (both fiscal years) had a decrease in the agreement rate after adjustment. For pre-admit co-morbidities (type (1)) the unadjusted match rates of 52.1% (FY 2002–2003) and 56.8% (FY 2003–2004) significantly decreased to the adjusted rates of 43.9% (FY 2002–2003) and 51.6% (FY 2003–2004). Similarly for post-admit co-morbidities (type (2)), the unadjusted match rates of 55.2% (FY 2002–2003) and 61.4% (FY 2003–2004) significantly decreased to the adjusted rates of 40.2% (FY 2002–2003) and 53.2% (FY 2003–2004).

Some trends between fiscal years change when adjusting for coder effect.

- There was a significant increase in the agreement for intervention coding after adjusting for coder effect in FY 2003–2004 that was not there in the unadjusted data.
- There was a significant increase in the agreement for complexity assignment (excluding complexity level 9) after adjusting for coder effect in FY 2003–2004 that was not there in the unadjusted data.

14 Recommendations

The following recommendations are being made following a review of the study findings to improve the quality of the clinical data in the Discharge Abstract Database and in particular, that of the case-costing facilities:

1. *Review the current concept of diagnosis typing with a view to improving the consistency of implementation.*

   Strategy: A consultative process should be conducted with stakeholders such as CIHI, CHIMA, MOHLTC, hospital staff, and national advisory committees to assess the current concept of diagnosis typing and how it is being applied. A recommended strategy for change would be provided. (Lead: CIHI)

2. *Conduct further analysis on the FY 2004/2005 clinical data to assess the extent to which initiatives launched in 2003–2004 have had an impact on reducing the discrepancies noted in the study.*

   Strategy: A smaller scale reabstraction study could be conducted to assess improvements in the quality of 2004/2005 data to assess the impact of educational workshops, other training, enhancements to standards, and other initiatives conducted during 2003–2004. A critical component of this study would be the inclusion of charts for inter-rater reliability assessment. (Lead: CIHI)

3. *Establish Local Data Management Partnerships.*

   Strategy: Establish Local Data Management Partnerships that will be aligned with the Local Health Integration Networks, and will facilitate collaboration between health care providers to consolidate, coordinate and standardize local data management functions through best practices, policies, standards and tools. (Lead: Ministry)

4. *Establish a Physician Documentation Expert Panel to engage physicians in addressing chart documentation issues.*

   Strategy: Establish a panel of 'documentation champions' from across the province to promote timely, accurate and complete documentation by physicians through the development of guidelines and tools, including such things as:
   - Physician education package
   - Chart completion policy
   - Recommendations to College of Physicians and Surgeons of Ontario (CPSO)
   - Guidelines for including chart documentation in medical school curricula

   (Lead: Ministry)

5. *Conduct detailed analyses of the discrepancy rates within the case-costing hospitals to determine the specific factors contributing to the observed results.*

   Strategy: Led by the MOHLTC, this review would involve CIHI, MOHLTC and the hospitals in the review of coding practices, the coding environment, and drivers of discrepancy rates. An assessment of these will help to identify further strategies for data quality improvement. (Lead: Ministry)

Appendices

Appendix A: Glossary of Terms

Attributes provide extra detail about an intervention. Attributes are extraneous to but related to the intervention code and include: status, location, extent and mode of delivery.

Case mix effect accounts for variation in the types of conditions presented in the patient population between facilities, and the influence this has on the findings. Refer to section 7.4.1.2 for details.

Case mix group (CMG) groups patients into clusters based on clinical diagnoses, procedures and resource utilization. It is a methodology that provides a way to describe the mix of patients treated in a hospital or jurisdiction (i.e. case mix).

Category refers to a concept in ICD-10-CA by which diagnoses are assigned to a single disease entity or to a group of similar entities. In relation to an entire diagnosis code, the category is its first three characters.

Canadian Classification of Health Interventions (CCI) was developed at CIHI as a national classification for coding health interventions. CCI has an expanded scope to encompass a broad spectrum of interventions to meet the needs across the continuum of health services, in Canada.

Chart refers to a patient's medical record, which is uniquely identified and may contain more than one episode of care (i.e. discharge).

Coder effect accounts for variation in the coding between reabstractors, and the influence this has on the findings. Refer to section 7.4.1.1 for details.

Co-morbidity refers to a significant condition that either coexists at the time of admission, or develops while in hospital. Selection of a condition as type (1), (2), (W), (X) or (Y) depends on whether it satisfies the requirements of significant, according to diagnosis typing definitions.

Confidence interval is the range of values that is likely to include the true population value. A 95% confidence interval means that if all possible samples were drawn from the population and the estimates were computed, 95% of the confidence intervals would contain the true population value.

Complexity (Plx) reflects the interaction of multiple diagnoses on length of stay (LOS) or resources within each Case Mix Group. Complexity overlay identifies those acute inpatients with additional diagnoses (other than the MRDx) for which a prolonged LOS and/or more costly treatment might be reasonably expected.

Complications refer to post-procedural conditions not specific to a body system and more general in nature such as postoperative hemorrhage. They are located in Chapter XIX "Injury, poisoning and certain other consequences of external causes".

Diagnosis type is applied to all conditions for DAD submission and is meant to signify the impact that the condition had on the patient's care.

Discharge refers to an episode of care in hospital. One chart may contain more than one discharge in that facility.

Discharge Abstract Database (DAD) contains data on inpatient hospital discharges across Canada. The DAD contains demographic, administrative and clinical data for: hospital discharges (inpatient acute, chronic, rehabilitation) and day surgeries.

Discrepancy refers to any difference between the original and reabstracted data.

Expected length of stay (ELOS) is the duration of a typical acute care visit in a Case Mix Group, measured in days. ELOS is a national average length of stay (ALOS) estimate that accounts for differences in age and complexity when these factors are found to be predictive of length of stay.

Extent attribute is extraneous to the intervention code and identifies quantitative measures (e.g. length of laceration, size of calculus, number of anatomical structures involved).

ICD-10-CA stands for the International Statistical Classification of Diseases and Health Related Problems Tenth Revision – Canada. ICD-10-CA is soon to be the single set of national standards for diagnosis coding, once it is implemented in all provinces.

Inflow refers to the assignment of a discharge into a Grouper output variable (e.g. CMG, MCC, complexity level) after the CIHI grouping methodology is applied to the reabstracted data, when the original Grouper output variable was something different.

Inter-rater reliability is the degree to which multiple evaluators obtain the same result and is an important measure of consistency.

Location attribute is extraneous to the intervention code and identifies anatomic detail and laterality (e.g. left, right, bilateral).

Length of stay (LOS) is the duration of time from the date of admission to the date of discharge in an episode of care, measured in days.

Major clinical category (MCC) aggregates patients more broadly than Case Mix Group. They generally describe a body system or specific type of clinical problem, and are determined by the Most Responsible Diagnoses.

Mandatory attribute relates to intervention attributes that are required for DAD submission. They provide extra detail about an intervention and ensure parity of data collection from the previous classifications. Some mandatory attributes affect Case Mix Group assignment.

Mandatory interventions refer to those interventions identified in the Case Costing study as mandatory for the purposes of reabstraction. Certain interventions were reabstracted because they impact the CIHI grouping methodology outputs such as resource intensity weight and expected length of stay.

Margin of error is a relative figure that may be expressed as a percentage and is calculated using the sampling error of an estimate. It is used to build a confidence interval for that estimate.

Most responsible diagnosis (MRDx) is the one diagnosis or condition that is accountable for the greatest portion of the length of stay or greatest use of resources.

Outflow refers to the assignment of a discharge out of a Grouper derived variable (e.g. CMG, MCC, complexity level) once the CIHI grouping methodology is applied to the reabstracted data, when the original Grouper output variable was grouped to that particular value.

Post-procedural conditions refer to conditions specific to a particular body system, such as N99.0 "post procedural renal failure". In ICD-10-CA, a special category has been created near the end of each of the body system chapters for post-procedural conditions.

Resource intensity weight (RIW) provides users with a tool to estimate expected resource use and relationships of costs between patient types. It indicates the relative value of treating a patient compared with treating the average patient whose RIW is 1.0000.

Rubric refers to the base of an intervention code. It is the first five characters of the intervention code that describes "what" was performed.

Secondary condition refers to conditions assigned a diagnosis type of (3) (0) or (9) which is mostly optional coding, and at a facility's discretion. They do not satisfy the requirements for determining co-morbidity.

Significant condition refers to conditions assigned a diagnosis type of (M) (1) (2) (W) (X) or (Y). They are conditions that influence the patient's stay in hospital.

Status attribute is extraneous to the intervention code and identifies revisions, abandoned interventions, converted interventions, or staged interventions.

Stratified sampling is a sampling procedure in which the population is divided into homogeneous subgroups or strata and the selection of samples is done independently in each stratum.

Appendix B: Data Elements Reabstracted in the Case Costing Study

Table B.1: Non-Medical Data Elements Reabstracted in the Case Costing Study[17]

Field Name	Description	Valid Values
Health care number	The patient's medical (insurance) number as assigned by the provincial/territorial government of the patient's home residence	Up to 12 digits
Gender	Patient's gender (sex)	M, F, U, O
Birth date	Age of Patient at the time of admission	Dates in format YYYYMMDD
Birth date is estimated	Used when birth date is unknown or only a partial birth date is known	Y, blank
Admission category	Differentiates between admission types: elective, urgent/emergent, newborn, stillborn, cadaveric donor	U, L, R, N, S
Admission date	The date that the patient officially registered as an inpatient	Dates in format YYYYMMDD
Discharge disposition	Status of the patient upon leaving the facility	01-09
Discharge date	The date that the patient was formally discharged from the facility	Dates in format YYYYMMDD
Alternate level of care days (ALC)	The days a patient has finished the acute care phase of treatment but remains in the acute care bed.	0-99999
Institution from	Facility or level of care which the patient was transferred from	5-digit provincially assigned institution number
Institution from type	Institution type the patient was transferred from	0-9, A, E, blank
Institution to	Facility or level of care the patient was transferred to	5-digit provincially assigned institution number
Institution to type	Institution type the patient was transferred to	0-9, A, E, blank
Weight (in grams)	Captured for newborns or neonates less than or equal to 28 days of age	0001-9000

[17] Definitions and valid values taken from the DAD 2002 Abstract Manual

Table B.2: Medical Data Elements Reabstracted in the Case Costing Study

Field Name	Description	Valid Values
Diagnosis prefix	Assigned to further distinguish ICD-10-CA diagnoses for study purposes	Q
Diagnosis code	ICD-10-CA codes recorded to describe the diagnoses/conditions of the patient while in hospital	ICD-10-CA codes
Diagnosis type	A one-digit code used to indicate the relationship of the diagnosis to the patient's stay in hospital	M, 1, 2, 3, 9, 0, W, X, Y
Intervention date	The date that the intervention was performed on the patient	Dates in format YYYYMMDD
Intervention code	CCI codes recorded to describe the operative and non-operative interventions performed during the patient's hospital stay	CCI codes
Status attribute	Used to identify interventions which are revisions, abandoned after onset, part of a staged process etc.	CCI codes
Location attribute	Used to identify additional anatomical detail or information on laterality	CCI codes
Extent attribute	Used to indicate a quantitative measure related to the intervention	CCI codes

Appendix C: Discrepancy Reasons Assigned in the Case Costing Study

There are four codes that reabstractors could choose from to assign a reason to the discrepancies they identified upon reabstraction.

"Chart Documentation" is a reason applied to discrepancies caused from the information provided in the chart. Examples of when this reason code would be used are:

- Different interpretation of documentation
- Incomplete documentation available at time of original abstraction
- Inconsistent or conflicting documentation on paper chart
- Information on chart missed
- Chart documentation supports more or less specific code selection

"Significance" is a reason assigned to discrepancies with the diagnoses typing only. Examples of when this reason code would be used are:

- Diagnosis coded did not have significant impact on treatment and/or LOS
- Diagnosis had significant impact on treatment and/or LOS
- Originally coded as diagnosis type (M) (1) (2) (W) (X) (Y), reabstracted as a (3) (or type (0) for a newborn chart).
- Originally coded as a type (3), reabstracted as an (M) (1) (2) (W) (X) (Y)

"Optional/Not wrong" is a reason assigned when the data reviewed is optional and/or not wrong to code. Examples of when this reason code would be used are:

- Either code correct
- Not wrong/necessary to code for DAD submission
- Reabstractor unable to access required information
- Original type (3) (0) (9) not reabstracted because the diagnosis was not included in the study.

"Standards/Codebook/Manual" is a reason assigned when there is a specific deviation from Canadian Coding Standards, DAD Abstracting Manual and coding conventions in ICD-10-CA and CCI. Examples of when this reason code would be used are:

- Not following coding standards, DAD Abstracting Manual
- Not following coding conventions in ICD-10-CA/CCI (index look up, inclusion/exclusion notes)

Appendix D: Definitions of Diagnosis Typing[18]

(M) Most responsible diagnosis - The one diagnosis or condition that is accountable for the greatest portion of the length of stay or greatest use of resources.

(1) Pre-admit co-morbidity - Conditions that exist pre-admission and satisfy the requirements for determining co-morbidity. Selection of a condition as a type (1) depends on whether it satisfies the requirements of significant, according to diagnosis typing definitions. Refer to Appendix E.

(2) Post-admit co-morbidity - Conditions that arise post-admission (while in hospital) and satisfy the requirements for determining co-morbidity. Selection of a condition as a type (2) depends on whether it satisfies the requirements of significant, according to diagnosis typing definitions. Refer to Appendix E.

(3) Secondary diagnosis - Conditions for which a patient may or may not receive treatment and does not satisfy the requirements for determining co-morbidity. Refer to section 6.2.1 for mandatory type (3).

(0) Optional - Secondary or supplemental conditions mostly associated with newborn cases. Diagnosis type (0) is used to distinguish babies born via caesarean section from those born vaginally. It can also be used for purposes other than newborn cases. However, type (0) was only mandatory in the study when applicable to newborns.

(9) External cause/place/activity - Diagnosis type (9) is applicable to External Causes of Morbidity and Mortality that is mandatory with codes classifiable to Chapter XIX – Injury, poisoning and certain other consequences of external causes. Type (9) also applies to Place of Occurrence (U98), which is mandatory with all accident and poisoning codes in the range W00-Y34, with the exception of Y06 and Y07. Type of Activity (U99) is optional and not a requirement in the study.

(W)(X)(Y) Service transfer diagnosis - An ICD-10-CA diagnosis code associated with the first/second/third patient service transfer. It must fit the criteria for a significant co-morbidity in order to apply.

(6) Proxy MRDx - (in effect FY 2005/2006, not applicable in study) A diagnosis type (6) is assigned to an asterisk code, the manifestation in a dagger/ asterisk convention when it fulfills the requirements stated in the definition of diagnosis type (M). Assign a diagnosis type (6) to an asterisk code on the second line of the diagnosis field of the abstract whenever the manifestation (which is identified with an asterisk symbol in ICD-10-CA) rather than the underlying cause is responsible for the greatest length of stay and/or resources used during hospitalization.

[18] Taken from the Canadian Coding Standards for ICD-10-CA and CCI, 2003

Appendix E: Criteria In Determining Significance[19]

Co-morbidities

Co-morbidities are all conditions that coexist at the time of admission or develop subsequently and demonstrate at least one of the following:

- significantly affects the treatment received
- requires treatment beyond maintenance of the preexisting condition
- increases the length of stay (LOS) by at least 24 hours.

Consider the following in determining whether a condition qualifies as a co-morbidity.

1. To support a determination of significance, there must be documented evidence in the physician's notes or discharge summary that the condition required at least one of the following:

   - clinical evaluation/consultation, excluding pre-operative anesthetic consults, where a new or amended course of treatment is recommended and instituted
   - therapeutic treatment/intervention with a code assignment of 50 or greater from Section 1 of CCI
   - diagnostic intervention, inspection or biopsy with a code assignment from Section 2 of CCI
   - extended the length of stay (LOS) by at least 24 hours

2. A post-procedural condition becomes a co-morbidity when any one of the following situations exist:

   - the condition appears in the physician's documentation as a complication of the procedure
   - the condition is present at discharge
   - the condition persists post-procedurally for at least 96 hours.
   - Some signs, symptoms and conditions may occur in the post procedural period but are NOT on their own regarded as post procedural co-morbidities. Further information may be found in the Canadian Coding Standards for ICD-10-CA and CCI. Examples of such conditions are:

     - anaemia
     - confusion
     - headache
     - difficulty walking
     - paraesthesia
     - vomiting
     - cough
     - cardiac arrhythmia
     - electrolyte imbalances
     - abnormal blood pressure reading
     - nausea
     - urinary retention
     - flatulence
     - dysuria

[19] Source: The Diagnosis Typing Definitions, Canadian Coding Standards for ICD-10-CA and CCI, 2003

3. Diagnoses that are only listed on the Front Sheet, Discharge Summary, Death Certificate, History & Physical or pre-operative anesthetic consults qualify as a diagnosis type (3) — secondary diagnosis. If there is documentation elsewhere in the chart that the condition affected the treatment received or required treatment beyond maintenance of the preexisting condition or increased the length of stay (LOS) by at least 24 hours it then must be determined if it is a type (1) or type (2) co-morbidity.

4. Nurses notes, pathology reports, laboratory reports, autopsy reports, medication profiles, radiological investigations, nuclear imaging, and other similar investigations are valuable tools for identifying the appropriate diagnosis code. To be classified as co-morbidities, these diagnoses must be supported by documentation as identified in number 1.

Note: The documentation of ongoing medication for treatment of a preexisting condition does not in itself denote significance. Conditions not qualifying as co-morbidities, if coded, should be classified to diagnosis type (3).

Appendix F: List of Hospitals in Reabstraction Report

Facility Key	Facility
A	Lakeridge Health Corporation
B	Quinte Healthcare Corporation – North Hastings
C	Credit Valley Hospital
D	Quinte Healthcare Corporation – Belleville
E	William Osler Health Centre – Georgetown
F	Quinte Healthcare Corporation – Trenton
G	University Health Network
H	St. Michael's Hospital
I	Quinte Healthcare Corporation – PE County
J	Ottawa Hospital
K	William Osler Health Centre – Brampton
L	Trillium Health Centre
M	Mount Sinai Hospital
N	William Osler Health Centre – Etobicoke
O	London Health Sciences Centre
P	Arnprior and District Memorial