RELIABILITY OF THE HMRI (CIHI) DATABASE:
A RE-ABSTRACTING STUDY

by

LINDA ANNE BROWN

B.A., The University of British Columbia, 1985

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF
THE REQUIREMENT FOR THE DEGREE OF
MASTER OF SCIENCE

in

THE FACULTY OF GRADUATE STUDIES

Health Services Planning and Administration Programme
Department of Health Care and Epidemiology

We accept this thesis as conforming
to the required standard

THE UNIVERSITY OF BRITISH COLUMBIA

15 September 1995

Copyright Linda Anne Brown, 1995
In presenting this thesis in partial fulfilment of the requirements for an advanced degree at the University of British Columbia, I agree that the Library shall make it freely available for reference and study. I further agree that permission for extensive copying of this thesis for scholarly purposes may be granted by the head of my department or by his or her representatives. It is understood that copying or publication of this thesis for financial gain shall not be allowed without my written permission.

Department of Health Care & Epidemiology

The University of British Columbia
Vancouver, Canada

Date Dec 14/95
In presenting this thesis in partial fulfilment of the requirements for an advanced degree at the University of British Columbia, I agree that the Library shall make it freely available for reference and study. I further agree that permission to extensive copying of this thesis for scholarly purposes may be granted by the head of my department or by his or her representatives. It is understood that copying or publication of this thesis for financial gain shall not be allowed without my written permission.

(Signature)  

Department of Health Care and Epidemiology  
The University of British Columbia  
Vancouver, Canada  

Date  

Dec 6, 1995
ABSTRACT

Accuracy of the information in health care database systems is essential throughout the health care system for functions such as planning and research. This study examines the reliability of the HMRI database, recently amalgamated under the CIHI. It compares selected data items from the in-patient hospital discharge database against a re-abstracted set of data items from the original health record in a sample of six Vancouver acute care hospitals for the fiscal year 31 March 1986 - 1 April 1987. Cases were restricted to acute care medical and surgical cases. A total of 606 cases using the ICD-9 classification system were re-abstracted using the original health record. Results demonstrated non-diagnostic variables demonstrated an overall agreement of 92.4%. The agreement for the Most Responsible Diagnosis (MRD) to four-digits is 61.4%, while individual hospital scores ranged from 52.0 to 69.6%. For the MRD to three-digits agreement increased to 72.1%, with individual hospitals ranging between 62.4 and 79.4%. The Principal Procedure (PP) agreement to three-digits was 64.8% with individual hospital scores ranging from 56.3 to 75.6%. For the PP at two-digits, agreement was 72.9%, with individual hospitals ranging from 61.4 to 85.4%. Denominators for secondary diagnoses and secondary procedures reflect the total number of diagnoses and/or procedures recorded. Secondary diagnoses to four-digits had agreement scores of 67.4% by number of diagnoses recorded and secondary procedures to three-digits of 80.7% by number of secondary procedures recorded. Total diagnoses and procedures combined demonstrated an overall agreement score of 68.3% with individual hospitals ranging from 61.8 to 73.1%. Agreement by case, where all relevant diagnostic and procedural codes in the entire record matched, dropped significantly to 34.5% for secondary diagnoses and to 59.7% for secondary procedures. The greatest frequency overall for the type of discrepancy was for clerical errors, especially for code books not used properly to determine specificity of the diagnosis. Specificity of the code is required, the information is available in the record, but specificity is not determined by the coder. The greatest frequency of discrepancy for the MRD was 73.1% for clerical errors. For the PP, 46.9% of discrepancies were in the selection of principal procedure and 42.9% for clerical errors. This study did not demonstrate a significant difference between individual coders by years of experience, by credentials or by years of experience and credentials. It was determined that the data are unsuitable for a quality of care study where the data are utilized beyond the individual hospital site. Care must be taken when utilizing those data for research purposes.
# TABLE OF CONTENTS

- TITLE PAGE ................................................. i
- ABSTRACT .................................................. ii
- TABLE OF CONTENTS ........................................ iii
- LIST OF ABBREVIATIONS .................................. xii
- LIST OF TABLES ............................................. xiv
- LIST OF FIGURES ........................................... xvi
- LIST OF APPENDICES ....................................... xvii
- ACKNOWLEDGEMENTS ....................................... xix

__CHAPTER 1 - LITERATURE REVIEW__

1.0 INTRODUCTION ............................................. 1

1.1 GROWTH OF HEALTH RECORD ASSOCIATIONS & EDUCATIONAL PROGRAMS ............................................. 3
  1.1.1 The American Experience ................................ 3
    1.1.1.1 Hospital Schools .................................. 4
    1.1.1.2 Technical Programs ................................ 4
    1.1.1.3 Independent Study .................................. 5
    1.1.1.4 Baccalaureate Degree Programs ...................... 5
    1.1.1.5 Representation on Educational Advisory Boards .... 6
    1.1.1.6 Graduate Education .................................. 6
    1.1.1.7 Continuing Education & Progression ................ 7
1.7 COMPONENTS OF THE HMRI (CIHI) SYSTEM

1.7.1 Most Responsible Diagnosis
1.7.2 Comparison of the MRD and the Principal Diagnosis
1.7.3 Positive Features
1.7.4 Problem Areas
1.7.5 Case Mix Groupings (CMG's)
1.7.6 Resource Intensity Weights (RIWs)

1.8 OTHER CRITICAL ISSUES

1.8.1 What is Reliability?
1.8.2 Sources of Reliability
1.8.2.1 Input Reliability
1.8.2.2 Instrument Reliability
1.8.2.3 Observer (Rater) Reliability
1.8.2.3.1 Inter-Rater Reliability
1.8.2.3.2 Intra-Rater Reliability
1.8.3 The Coding Process
1.8.3.1 Three-dimensions of Coding Accuracy
1.8.3.2 Classification Rules
1.8.3.3 To Classify a Term Not Listed
1.8.4 Hospital Procedures

1.9 LITERATURE REVIEW ON QUALITY OF HEALTH CARE DATABASES

1.9.1 Hendrickson and Myers - Pennsylvania (1969)
1.9.2 Roos et al. - Manitoba Health Services Commission (1971)
1.9.3 George & Maddocks - England (1972-1973)
1.9.4 Institute of Medicine (IOM) - U.S. National Studies (1974)
1.9.5 New Jersey Department of Health Study (NJDH) (1977)
1.9.6 IOM's Third National Study (1977)
1.9.7 Corn - New Jersey Hospitals (1977)
1.9.8 Barnard & Esmond - Chicago, Illinois (1979)
1.9.9 Doremus & Mickenzi - Cleveland, Ohio (1978)
1.9.10 Connell, Blide & Hanken - Washington State (1978-1979)
1.9.12 Lloyd & Rissing - Georgia (1982)
1.9.13 Massanari et al. - Iowa (1984)
1.9.14 Hsia et al. - Maryland (1984-1985)
1.9.15 Iezzoni et al. - Boston (1984-1985)
1.9.16 Lovison & Bellini - Italy (1985)
1.9.18 Wood et al. - Bristol, England (1989)
1.9.19 Recent Canadian Studies
1.9.19.1 OHA, OMOH & HMRI - Ontario (Fiscal 1988/89)
1.9.19.2 Armstrong et al. (UBC) (1985)
CHAPTER 2 - METHODOLOGY

2.0 RATIONALE .............................................. 119
2.1 OBJECTIVE AND PURPOSE .......................... 121
2.2 THE QUESTIONS PROPOSED FOR THIS STUDY .... 122
2.3 HYPOTHESES ............................................. 123
2.4 NULL HYPOTHESES ..................................... 124
2.5 SAMPLE SIZE CALCULATIONS ...................... 125
   2.5.1 Three Distinct Diagnostic Categories Selection ..... 125
   2.5.2 Stratified Random Categories Selection ........... 126
   2.5.3 Total Sample Size ................................ 128
2.6 SAMPLING PLAN ......................................... 129
   2.6.1 Selection of Hospitals ............................ 129
   2.6.2 Sampling Frame .................................... 129
   2.6.3 Selection of Cases for the Distinct Diagnostic Categories .... 130
   2.6.4 Selection of Cases for Stratified Random Sample .... 130
   2.6.5 Alternate Records .................................. 131
   2.6.6 Subsample for IRR Reliability .................... 131
2.7 RESEARCH DESIGN .................................... 132
2.8 DATA SOURCE .......................................... 133
2.9 RESEARCH INSTRUMENTS .................................................. 134
2.9.1 Re-abstracting form ...................................................... 134
2.9.2 Hospital Questionnaire .................................................. 134
   2.9.2.1 Purpose of Questionnaire ...................................... 135
2.9.3 ICD-9 Classification System .......................................... 136

2.10 STUDY PLAN .................................................................. 137
2.10.1 Individual Hospital Research Committees ....................... 137
2.10.2 Updating Diagnostic & Procedural Code Books ................. 137
2.10.3 Principal Investigator Training ...................................... 138
2.10.4 Selection of Cases and Preparation of Original Records .... 138
2.10.5 Hospital Questionnaire Interview .................................. 140
2.10.6 Re-Abstracting Process ............................................... 141
2.10.7 Compiling Information and Re-Filing Health Records ......... 142
2.10.8 Reconciliation Process ................................................ 143
2.10.9 Subsample for Intra-Rater Reliability ............................. 144
2.10.10 Computer Input ....................................................... 145
2.10.11 Statistical Data Analysis ............................................. 145

2.11 PROJECT TEAM ................................................................. 148
2.11.1 Principal Investigator ................................................... 148
2.11.2 Assistant .................................................................. 148
2.11.3 Computer Analyst ....................................................... 149
2.11.4 Hospital Personnel ...................................................... 150

2.12 TYPE OF DISCREPANCY CATEGORIES .................................. 151
2.12.1 System Discrepancies .................................................... 151
   2.12.1.1 Judgement ........................................................... 151
   2.12.1.2 Classification ....................................................... 152
   2.12.1.3 Procedural .......................................................... 152
2.12.2 Clerical Discrepancies ................................................... 152
   2.12.2.1 Data Conversion .................................................. 153
   2.12.2.2 Charting Errors .................................................... 153
   2.12.2.3 Coding/Abstracting Discrepancies ......................... 153
2.12.3 Documentation Discrepancies ....................................... 154
   2.12.3.1 Incomplete Record ............................................... 154
   2.12.3.2 Discrepant Records .............................................. 154
2.12.4 Other Discrepancies .................................................... 155
2.12.5 Extra .0 in MOH Data .................................................. 155

2.13 CONFIDENTIALITY ............................................................. 156

2.14 SIGNIFICANCE OF FINDINGS ........................................... 157
CHAPTER 3 - RESULTS OF DATA ANALYSIS

3.0 INTRODUCTION ......................................................... 165

3.1 AGREEMENT BETWEEN PRINCIPAL INVESTIGATOR AND CODING SUPERVISOR ......................................................... 167

3.2 MOST RESPONSIBLE DIAGNOSIS ........................................ 169
   3.2.1 Frequency of Match - All Hospitals .................................. 169
   3.2.2 Range of Match - Individual Hospitals ................................. 172

3.3 PRINCIPAL PROCEDURE .................................................. 174
   3.3.1 Frequency of Match - All Hospitals .................................. 175
   3.3.2 Range of Match - Individual Hospitals ................................. 176

3.4 SECONDARY DIAGNOSES ............................................... 178
   3.4.1 By Number of Secondary Diagnoses Recorded .......................... 178
      3.4.1.1 Frequency of Match - Four-digit, Raw Score ..................... 179
      3.4.1.2 Frequency of Match - Four-digit, Adjusted Level ............... 179
   3.4.2 By Number of Cases .................................................... 180
CHAPTER 4 - DISCUSSION AND CONCLUSION

4.0 INTRODUCTION .............................................. 222

4.1 AGREEMENT SCORES FOR DIAGNOSES ................. 222
  4.1.1 Most Responsible Diagnosis ...................... 222
  4.1.2 Secondary Diagnoses .............................. 224
  4.1.3 Total Diagnoses .................................... 225

4.2 AGREEMENT SCORES FOR PROCEDURES ............... 226
  4.2.1 Principal Procedure ................................ 226
  4.2.2 Secondary Procedures ............................. 227
  4.2.3 Total Procedures .................................... 227

4.3 TOTAL SCORES BY NUMBER OF CASES .................. 228

4.4 AGREEMENT SCORES FOR NON-DIAGNOSTIC DATA .... 229

4.5 AGREEMENT SCORES FOR DIAGNOSTIC ................. 230

4.6 TYPE OF DISCREPANCY ..................................... 231
  4.6.1 MRD .................................................. 231
  4.6.2 Secondary Diagnoses .............................. 232
  4.6.3 Principal Procedures ............................. 232
  4.6.4 Secondary Procedures ............................. 233
  4.6.5 Overall Discrepancies ................................ 234
  4.6.6 Non-Diagnostics ................................... 234

4.7 CODERS .................................................... 235
  4.7.1 Agreement Scores for Coders ...................... 235
  4.7.2 Coder Training ..................................... 236
  4.7.3 Testing of Coders .................................. 241
  4.7.4 Credentialling and Continuing Education ......... 242
LIST OF ABBREVIATIONS

ADT - Admission Discharge Transfer System
AHIMA - American Health Information Management Association
AMA - American Medical Association
AMRA - American Medical Record Association
AMRLA - American Medical Record Librarians Association
ART - Accredited Records Technician

BCIT - British Columbia Institute of Technology

CAMRL - Canadian Association of Medical Record Librarians
CCC - Council on Clinical Classifications
CCHRA - Canadian College of Health Record Administrators
CCP - Canadian Compendium of Procedures
CHA - Canadian Hospital ASsociation
CHDC - California Health Data Corporation
CHRA - Canadian Health Record Association
CIHI - The Canadian Institute for Health Information
CMG - Case Mix Groups
CPHA - Commission on Professional and Hospital Activities

FTE - Full Time Equivalent

HIS - Health Information Services Department (Douglas College)
HMRI - Hospital Medical Records Institute
HDT - Health Data Technologist
HRA - Health Record Administrator
HRABC - Health Record Association of B.C.
HRT - Health Record Technician
HUP - Hospital Utilization Project
H-ICDA - Hospital adapted version of ICDA-8
H-ICDA-2 - 2nd revision of H-ICDA

ICD - International Classification of Diseases
ICD-9 - International Classification of Diseases, 9th revision
ICD-9CM - International Classification of Diseases, 9th revision, American version
ICD-10 - International Classification of Diseases, 10th revision
ICD-A - Adapted for Indexing Hospital Records by diseases and operations (1959, revision 1962)
ICDA-8 - Adapted for use in the United States (ICD-8 version)
IOM - Institute of Medicine
IRR - Intra-rater reliability
MOH - Ministry of Health
MRA - Medical Record Administrator
MRD - Most Responsible Diagnosis
MRL - Medical Record Librarian
MRT - Medical Record Technician

NJDH - New Jersey Department of Health

OHA - Ontario Hospital Association
OMA - Ontario Medical Association
OMOH - Ontario Ministry of Health

PAS - Professional Activity Study
PP - Principle Procedure

QA - Quality Assurance
QUEST - Quality, Utilization, Effectiveness, Statistically Tabulated (Database)

RIW - Resource Intensity Weights
RRA - Registered Record Administrator
RRL - Registered Record Librarian

SNOMed - Systematized Nomenclature of Medicine
SNOP - Systematized Nomenclature of Pathology, 1965

UBC - University of British Columbia

WHO - World Health Organization
LIST OF TABLES

1. AGREEMENT ON CORRECT CODES BETWEEN PRINCIPAL INVESTIGATOR AND CODING SUPERVISOR - BY NUMBER OF CASES FOR INDIVIDUAL HOSPITALS ............................................. 167

2. FREQUENCY OF MATCH FOR MOST RESPONSIBLE DIAGNOSIS - BY NUMBER OF CASES - OVER ALL HOSPITALS ................................................................. 170

3. FREQUENCY AND PERCENT OF MATCH FOR MOST RESPONSIBLE DIAGNOSIS - BY INDIVIDUAL HOSPITAL ................................................................. 172

4. FREQUENCY OF MATCH FOR PRINCIPAL PROCEDURE - BY NUMBER OF CASES - OVER ALL HOSPITALS ................................................................. 175

5. FREQUENCY AND PERCENT OF MATCH FOR PRINCIPAL PROCEDURE - BY INDIVIDUAL HOSPITAL ................................................................. 176

6. FREQUENCY OF MATCH OF DIAGNOSES RECORDED BY ORIGINAL CODER - FOR SECONDARY DIAGNOSES, 4-DIGIT LEVEL, RAW SCORE ........................................... 179

7. FREQUENCY OF MATCH OF DIAGNOSES RECORDED BY ORIGINAL CODER - FOR SECONDARY DIAGNOSES, 4-DIGIT LEVEL, ADJUSTED ........................................... 180

8. FREQUENCY OF MATCH FOR SECONDARY DIAGNOSES, 4-DIGIT LEVEL, ADJUSTED - BY NUMBER OF CASES ................................................................. 181

9. FREQUENCY OF MATCH FOR SECONDARY PROCEDURES, 3-DIGIT LEVEL - BY NUMBER OF PROCEDURES RECORDED ................................................................. 183

10. FREQUENCY OF MATCH FOR SECONDARY PROCEDURES, 3-DIGIT LEVEL - BY NUMBER OF CASES ................................................................. 184

11. FREQUENCY OF MATCH FOR NON-DIAGNOSTICS - BY NUMBER OF CASES - INDIVIDUAL HOSPITAL RATES ................................................................. 186

12. RANKING OF BIRTHDATE - BY NUMBER OF RECORDS WITH DISCREPANT BIRTHDATES - OVER ALL HOSPITALS ................................................................. 187
13. FREQUENCY OF MATCH FOR COMBINED VARIABLES - BY NUMBER OF DIAGNOSES AND PROCEDURES RECORDED ............... 189

14. FREQUENCY OF MATCH FOR DIAGNOSTIC AND PROCEDURAL CODES - BY NUMBER OF CASES .................................................. 192

15. FREQUENCY OF MATCH FOR MOST RESPONSIBLE DIAGNOSIS, 4-DIGIT LEVEL, ADJUSTED - BY DIAGNOSTIC CATEGORIES - OVER ALL HOSPITALS .................................................. 194

16. FREQUENCY OF MATCH FOR PRINCIPAL PROCEDURE, 3-DIGIT LEVEL - BY DIAGNOSTIC CATEGORIES - OVER ALL HOSPITALS ... 196

17. FREQUENCY OF "TYPE OF DISCREPANCY" - AS A PERCENTAGE OF DIAGNOSES AND PROCEDURES RECORDED - ALL HOSPITALS .. 198

18. FREQUENCY OF DISCREPANIES BY TYPE OF DISCREPANCY, MOST RESPONSIBLE DIAGNOSIS - FOR INDIVIDUAL HOSPITALS .......... 204

19. FREQUENCY OF DISCREPANIES BY TYPE OF DISCREPANCY, PRINCIPAL PROCEDURE - BY NUMBER OF PROCEDURES RECORDED BY ORIGINAL CODER - FOR INDIVIDUAL HOSPITALS .................................................. 206

20. FREQUENCY OF DISCREPANIES BY TYPE OF DISCREPANCY, SECONDARY DIAGNOSES - BY NUMBER OF SECONDARY DIAGNOSES RECORDED BY ORIGINAL CODER - FOR INDIVIDUAL HOSPITALS ... 208

21. FREQUENCY OF DISCREPANIES BY TYPE OF DISCREPANCY, SECONDARY PROCEDURES - BY NUMBER OF SECONDARY PROCEDURES RECORDED BY ORIGINAL CODER - FOR INDIVIDUAL HOSPITALS .................................................. 210

22. FREQUENCY OF DISCREPANIES BY TYPE OF DISCREPANCY, NON-DIAGNOSTICS - FOR INDIVIDUAL HOSPITALS ................. 212

23. FREQUENCY OF DISCREPANIES BY TYPE OF DISCREPANCY, COMBINED DIAGNOSES AND PROCEDURES - BY NUMBER OF DIAGNOSES AND PROCEDURES RECORDED BY ORIGINAL CODER (EXCLUDING NON-DIAGNOSTICS) - FOR INDIVIDUAL HOSPITALS ... 214

24. INTRA-RATER RELIABILITY FOR PRINCIPAL INVESTIGATOR FOR DIAGNOSTIC CODES TO 4-DIGITS AND FOR PROCEDURAL CODES TO 3-DIGITS BY TOTAL NUMBER OF DIAGNOSES AND PROCEDURES RECORDED .................................................. 222
LIST OF FIGURES

1. Integration into a classification of disease for SNOMed ................. 46
LIST OF APPENDICES

A. DEFINITIONS OF TERMS ........................................... 299
B. LIST OF VENDORS .................................................. 302
C. HMRI (CIHI) PAPER ABSTRACT .................................. 303
D. HMRI GUIDELINES FOR SELECTION OF MOST RESPONSIBLE DIAGNOSIS, ICD-9 ............................... 304
E. LISTING OF HOSPITALS IN GREATER VANCOUVER REGIONAL HOSPITAL DISTRICT SELECTED FOR STUDY ..................... 313
F. SAMPLE SIZE CALCULATIONS, NUMBER OF SELECTED CASES PER CATEGORY AND SAMPLING FRAME ......................... 314
G. STANDARD ERROR CALCULATIONS AND CONFIDENCE INTERVALS ................................................................. 316
H. MOST FREQUENTLY OCCURRING DIAGNOSES IN B.C. FISCAL 86/87 ............................................................... 317
I. ICD-9 CATEGORIES FOR STRATIFIED RANDOM SAMPLE ............................ 318
J. HOSPITAL PROGRAMS DATA FILE (B.C. MINISTRY OF HEALTH) .......... 319
K. PERMISSION TO ACCESS DATA - MINISTRY OF HEALTH .................... 320
L. PERMISSION TO ACCESS DATA - LIONS GATE HOSPITAL .................. 321
M. PERMISSION TO ACCESS DATA - ST. PAULS HOSPITAL ...................... 322
N. PERMISSION TO ACCESS DATA - VANCOUVER GENERAL HOSPITAL ................................................................. 323
O. PERMISSION TO ACCESS DATA - MOUNT SAINT JOSEPH HOSPITAL .............. 324
| P. | PERMISSION TO ACCESS DATA - UNIVERSITY HOSPITAL*, UBC SITE AND SHAUGHNESSY SITE | 325 |
| Q. | RE-ABSTRACTING FORM | 328 |
| R. | HOSPITAL QUESTIONNAIRE | 329 |
| S. | SUMMARY OF STUDY PLAN | 337 |
| T. | APPROVAL FOR ETHICAL REVIEW | 338 |
| U. | APPROVAL FROM THESIS SCREENING PANEL | 340 |
| V. | TRAINING SESSION DETAILS - BACK TO BASICS WITH ICD-9 | 341 |
| W. | TESTING RESULTS | 343 |
| X. | HOSPITAL SCHEDULE DATES & CONTACTS | 344 |

* During the fiscal year 1986/87, University Hospital, UBC Site and University Hospital, Shaughnessy Site were two separate hospitals. Both sites are treated as independent hospitals for the purpose of this study.
ACKNOWLEDGEMENTS

I would like to take this opportunity to express my gratitude to members of my thesis committee for their assistance, advice, encouragement and support during the course of this study.

- Dr. Morris Barer, Supervisor, Professor, Department of Health Care and Epidemiology, Faculty of Medicine, University of British Columbia and Director, Centre for Health Services and Policy Research;
- Dr. Steve Marion, Assistant Professor, Department of Health Care and Epidemiology, Faculty of Medicine, University of British Columbia; and
- Dr. Robert Armstrong, Associate Professor, Faculty of Medicine, University of British Columbia and Vice-President, Medicine, Sunnyhill Health Center for Children in Vancouver.
- Dr. Geoff Anderson, Assistant Professor, Department of Health Care and Epidemiology, Faculty of Medicine, University of British Columbia until his move to the Institute for Clinical Evaluative Studies, Toronto, until his move.

This study would not have been possible without the cooperation of the participating hospitals. Numerous individuals at each of the hospitals were extremely helpful during the data collection stage. Their support is gratefully acknowledged.

Assistance is acknowledged from Richard Chamberlayne and Bo Green from the Health Database Linkage Project, Centre for Health Services and Policy Research for drawing the samples, and from Teresa Peters for assisting at each of the hospital sites.

Support from Deanna Hickie from HMRI is also gratefully acknowledged.

I am ever so grateful to my loving children: to Devon and Michelle, for their patience and understanding throughout the course of this study and to Michael for his encouragement and support. I also want to express my appreciation to Faith and Phil, who were there for me when the pressure was on.

A very special thank you to you all.
CHAPTER 1 - LITERATURE REVIEW

1.0 INTRODUCTION

Information stored in a health care coding and classification database is essential for effective functioning of the entire health care system. Health care databases are a means towards the accomplishment of tasks, such as epidemiologic research, which can lead to a better understanding of the utilization of the health care system. Dr. Jérôme-Forget (1984), then Assistant Deputy Minister, Policy Planning & Information Branch, Health & Welfare Canada, declared in her opening address to the International Conference of the Medical Informatics Association in 1984 that information systems are invaluable tools for determining priorities in the health care system by providing statistics on various diseases and causes of mortality and morbidity.

A uniform, comprehensive, and reliable information system is essential to plan for the health needs of the nation, to evaluate the system’s effectiveness, and to assure the availability of adequate, but necessary, information. The accuracy of data coded for health care statistical purposes is extremely important because the analysis of these data can lead to major decisions in planning, budgeting and allocation of resources of various parts of the total health care system (Wright, 1983).

This thesis examines the reliability of a patient-centred health care database. The study encompasses a re-abstracting study of the Hospital Medical Records Institute (HMRI) data at six Vancouver hospitals utilizing original health records for the fiscal year 31 March
1986 to 1 April 1987. HMRI has since amalgamated with three other health databases to form the Canadian Institute of Health Information (CIHI).

This chapter describes the growth of health record associations and educational programs, the evolution of health records and clinical database systems in Canada up to and including the CIHI, nomenclatures and classification systems, the history of the International Classification of Diseases (ICD), the ICD-9 Classification System, the ICD-10 Classification System, and other critical issues. It reviews the literature on the quality of health care databases and describes the aims of the research. Chapter 2 describes the methodology, Chapter 3 describes the data results, Chapter 4 provides a discussion of the analysis and the conclusion.
1.1 GROWTH OF HEALTH RECORD ASSOCIATIONS & EDUCATIONAL PROGRAMS

This section reviews the growth of health record associations in the United States and in Canada and documents the evolution of education programs in the two countries. It also reviews coder certification and the future of health record personnel in Canada.

1.1.1 The American Experience

The first North American association to recognize a specific "type of person" in charge of hospital records was the Association of Record Librarians of North America in response to the Congress of American College of Surgeons' annual meeting in 1928 (Szabo, 1980). The association was organized in 1928 and incorporated in 1929 with 58 charter members (Huffman, 1972). Canadians were members of this organization. Its main objective was "to elevate the standards of clinical records in hospitals, dispensaries, and other distinctly medical institutions" (Huffman, 1972, p. 23).

The name of the association was changed to the American Medical Record Association (AMRA), and in 1991 was changed to the American Health Information Management Association (AHIMA). Today, AHIMA has more than 35,000 credentialed professionals (Skurka, 1994). An AHIMA member belongs to both the national and the state association in the same member category (Huffman, 1972), ensuring representation and membership in both levels of organization.

As early as 1941, the AMRA requested the American Medical Association (AMA) to assume responsibility for the approval of schools for Medical Record Administrators (MRAs).
Today, the responsibility for accrediting agencies in the United States is shared between the AMA and the AHIMA, an unheard of circumstance in Canadian history.

1.1.1.1 *Hospital Schools*

In 1932, the basis of an educational foundation in North America was conceived. Examinations were formulated for registrants who could, if successful, take on the professional status of Registered Record Administrator (RRA). The establishment of an American formal educational program began in four hospital-based schools in 1935. St. Mary’s hospital in Duluth, Minnesota was affiliated with the College of Scholastica and was the first program to grant a baccalaureate degree in medical records. Other programs gradually made this transition. By the 1960’s, most hospital schools were phased out.

1.1.1.2 *Technical Programs*

In 1951, a shortage of RRA’s necessitated curriculum planning for ancillary health record workers, those less trained and experienced than the RRA’s; thus an Accredited Record Technician (ART) program was established in 1953, again largely hospital-based (Huffman, 1972; 1975).

During the 1960’s, two-year training programs for technicians were founded in colleges and universities that grant associate degrees. Currently there are about 134 accredited schools for technicians. More than 60% of the active membership in AHIMA is comprised of ARTs (Grostick & Slovensky, 1989).
1.1.1.3 *Independent Study*

Enrollment in independent study correspondence program of 18-24 months duration began in 1962 to meet the demand for ancillary medical record personnel. A high school diploma was required for entrance. The curriculum consisted of 25 lessons in medical record science and medical terminology. Students could write the national examinations for technicians. In 1979 the curriculum was changed dramatically to a more comprehensive program. Currently the program contains over 97 lessons and is extended over a 36-month duration. The high school graduation requirement was maintained. Students may now write the national exam for technicians after completing 30 semester hours of college credit (Huffman, 1985).

1.1.1.4 *Baccalaureate Degree Programs*

The Americans have maintained four-year baccalaureate programs for Medical Record Administrators (MRAs). In addition, candidates with any four-year degree, with the addition of a nine-month post baccalaureate certificate program in Health Information Administration are also eligible to write the national certification examination offered by the AHIMA (Szabo, 1980; Barer et al., 1982; University of Washington¹). There are 52 accredited degree programs across the United States. All either grant degrees or require them for entrance.

At the University of Washington, applicants need to have pre-requisite course work in human anatomy and physiology along with laboratory, medical terminology, statistics, principles of management and computer applications. The curriculum includes courses in

---

¹University of Washington program information.
health data systems, organizational theory, computer systems in health care, health
information systems analysis, professionalism and leadership, quality assurance, finance and
legal issues. All students complete a health information management internship and project
before being eligible to write the national exam.

1.1.1.5 Representation on Educational Advisory Boards

The Educational Advisory Board for the Health Information Administration program
at the University of Washington includes members from a wide range of health services,
including professors from medical schools, practising RRA's and several top management
people from a variety of positions within health care organizations. Other American schools
follow a similar format. The schools' philosophy implies that representation from other
health care professions ensures a proactive and liberal role in the education of health
information administrators. This process results in the expertise of the profession being held
as a legitimate function within the health information industry. The competency of those in
the profession ensures such a legitimate role.

1.1.1.6 Graduate Education

There are three master's programs in Health Information Management affiliated with
Medical Record Administration programs at the University of Alabama at Birmingham
(Birmingham, Alabama), Ohio State University (Columbus, Ohio), and University of
Pittsburgh (Pittsburgh, Pennsylvania). Of 20,093 active medical record professionals in 1989
in the United States, 5.5% held graduate degrees, while an additional 5.5% indicated they
were working on masters' degrees (Grostick & Slovensky, 1989). In Canada, there is no equivalent educational opportunity at the baccalaureate level, let alone the graduate level.

1.1.1.7 Continuing Education & Progression

The AHIMA also recognizes that a true professional never finishes learning. Since 1975, AHIMA has required 15 hours per year over a 5-year cycle for RRAs and 10 hours per year for technicians to retain credentials. Credentials can be removed if upgrading is not fulfilled. Recently these 5-year cycles were changed to 2-year cycles with the same number of education hours per year required. Every individual is responsible for self-assessment of her personal competence and maintenance of compliance through appropriate educational programs (AMRA, 1987c). Progression from an ART level to an RRA level can only be obtained through registration in a formal education program.

1.1.2 The Canadian Experience

Canadian health record personnel saw the need for independence from their American beginnings. The Ontario Association of Medical Record Librarians was founded as early as 1935 as a breakaway from its American counterpart. The Canadian Association of Medical Record Librarians (CAMRL) was organized in 1942 and obtained Dominion Charter in 1949. However, total membership was low and Szabo (1980) notes that only three members from the province of British Columbia were registered at that time.

The provincial Health Record Association of British Columbia (HRABC) was established in 1950. The mission of the HRABC is "to contribute to the quality of health
care in this province by promoting excellence in health information management and professional practice" (HRABC, 1990). Unlike its American counterpart, membership in the national association does not automatically confer provincial membership. As a result, membership in the provincial associations is relatively higher than membership in the national association. Many members belong to the provincial associations, few belong to both. Policies and procedures are implemented provincially and are not necessarily consistent throughout the country.

CAMRL's name was changed to the Canadian Health Record Association (CHRA) in 1973. The purpose of the CHRA is "to ensure that health record professionals are recognized leaders in health information management" (CHRA, 1990). The board of directors for the CHRA consists entirely of health record representatives from each of the provinces, through a nomination and ballot process and the Executive Director as an ex-officio member.

The CHRA assigned accreditation and standard setting powers to the then newly founded subsidiary, the Canadian College of Health Record Administrators (CCHRA). The CCHRA is the professional registering body that sets the national educational standards for the health record field. The objective of the CCHRA is "to advance health record science for the benefit of the citizens of Canada and to set or establish standards in the field of health record science which may be attained through educational institutions" (CCHRA, 1990). The composition of the board of directors for the college consists of board members from the CHRA. In addition, a Canadian representative to the International Federation of Health Record Organizations and the Executive Director of the CHRA are ex-officio members without voting rights.
Membership categories for the association are fellow, certificate, associate and other. Certificate level is for the administrators, such as the Health Record Administrators (HRAs) and the Medical Record Librarians (MRLs), while associate category is for technicians, such as the Accredited Record Technicians (ARTs), the Health Record Technicians (HRTs), and the Health Data Technologists (HDTs). There is a requirement that members be active in the health record field to hold these member categories. The "other" category is reserved for other professions and inactive health record personnel. The designations allocated by the registering body are CCHRA(C) for certificate members and CCHRA(A) for associate members.

At the inception of the CCHRA, the titles associated with membership changed from that of MRL to HRA and all previous MRL registrants were eligible to write the certificant exam, regardless of previous educational requirements.

1.1.2.1 Hospital Schools

The Canadian educational requirements for health record personnel are not as stringent as those in the United States. Curricula are not as encompassing and duration of the programs is not as long. In 1936, the first hospital schools for training MRLs were started. Incumbents were limited to those students with high school education and the course was one year in length.

The New Westminster school in B.C. was started in 1955 but in some years it graduated only one student or sometimes no students at all (Szabo, 1980). The maximum number of students enrolled was nine. The school was hospital-based and closed in 1974.
shortly after the advent of the two-year program at the British Columbia Institute of Technology (BCIT).

The Americans linked hospital schools with university programs and provided an ongoing cooperation with the American Medical Association for accreditation, but the Canadians did neither. Hospital schools were closed and with only one university program in place in Canada, the norm became a two-year diploma program.

1.1.2.2 Independent Study

By the 1950's, the CAMRL and the Canadian Hospital Association (CHA) jointly set up a two-year correspondence course for persons already employed in Medical Record Departments of hospitals. High school graduation as a requirement was not necessary. Graduates became known as Medical Record Librarians (MRL's) and were eligible to write a registration examination after five years' practical experience. The education clause was eventually changed and graduates needed high school graduation to write the registration exams. Szabo (1980) notes that these students were not trained in computer science, statistics or management. For almost ten years, two-year correspondence students with no previous educational requirements were granted the MRL status, the same status given a four-year baccalaureate graduate in the United States. The two-year correspondence course ended in 1960.

With the demise of the two-year correspondence program, the association still recognized the need for a technician program to supply trained health record personnel who lacked formal preparation in health records. Therefore, during the 1960's, the CHA, jointly
with the CAMRL, started a one-year technician program through correspondence. High school graduation was not a requirement for admission. Originally graduates were called Accredited Record Technicians (ARTs), with the name change in 1973 to Health Record Technicians (HRTs). Graduates could write the professional examinations at the associate level.

For over 20 years, the CHA, currently in conjunction with the CCHRA, has offered the one-year certificate correspondence program for HRTs. Although the curriculum has changed over the years, the curriculum encompassed medical terminology, anatomy and physiology, the study of disease, and health record management. In 1993-94 the program was extended to two years in length and the curriculum dramatically changed from 8 to 16 units, each unit representing a different topic area and appropriate lessons. Students take four weeks to complete each unit. Statistical methods, coding, and a small component on management were added to the curriculum. Coding training was added to six units in the second year of the program. Applicants must be employed in a health record department and high school graduation is still preferred, but this requirement may be waived for mature students. The three week structured internship at another hospital was retained and extended at the end of both years. The program is recognized by the CHRA/CCHRA and graduates are eligible to write the professional examination at the associate level.

1.1.2.3 Technical Programs

In the 1960 and 70's, one-year programs for technicians were started at community colleges. In 1971, an 18-month program for Health Data Technologists (HDTs) was started
at the British Columbia Institute of Technology (BCIT). These HDT graduates were only given the equivalent of associate status in the CCHRA because the students were considered technicians, although the program content was deemed equivalent to that of an administrator program (Szabo, 1980). Szabo, previously a Notre Dame instructor, states that the HDT program was clearly patterned after the two-year Health Record Administrator (HRA) programs elsewhere in Canada, yet graduates did not maintain the right to write registration exams at the certificant level. The HDT program was discontinued in 1978.

By 1979, the BCIT program began two-year training programs for HRA’s eligible to sit the CCHRA certificant exam. By 1981, BCIT offered evening continuing education courses for HDT’s to upgrade to the HRA level. The BCIT program was short-lived and by 19892 was transferred to the Health Information Services (HIS) Department at Douglas College in New Westminster.

Until recently, technical education programs at community colleges in British Columbia seldom enabled the transfer of credit into a university program. Associate degrees are a new phenomenon at the college level in British Columbia, and the health record diploma programs are currently not eligible for any transfer of credit to university programs.

In British Columbia, training options are limited to those programs offered by Douglas College in New Westminster and the CHA. Currently Douglas College is the only educational institution in British Columbia to offer training for HRTs and HRAs. Both programs are recognized and accredited by the CHRA/CCHRA.

---

2The above information was obtained through the calendar produced by Douglas College and brochures developed by the Health Information Services Department.
The one-year certification program at Douglas College is designed to educate HRT students in the technical aspects of health record and information systems. The curriculum includes anatomy and physiology, applied medical terminology, applied pathophysiology, health record systems, introduction to quality assurance, introduction to coding, coding and data retrieval, introduction to computers, communications and health information reporting. Students become competent in evaluating health records according to specific documentation standards, coding and abstracting, collecting and presenting statistical reports, controlling the confidentiality and release of health information, using computers in relation to health record functions, and providing information to support quality assurance programs. An 8-week practicum completes the first year. Graduates are eligible to write the associate level exam through the CCHRA.

The two-year diploma program includes the curriculum offered to the first year students. In addition to this curriculum, HRA’s also study the management of health record departments and their resources in relation to the health care system, including the design and implementation of computerized health information systems, health record administration, employee relations, development of health record policies and procedures, budgeting, work measurement, financial aspects of clinical information, quality review programs and research. A 10-week practicum completes the second year. Graduates write the certificate level exam of the CCHRA. It should be noted that the difference in curricula between the one-year and two-year programs is administrative. Both programs offer the technical skills required by the health record personnel.
The difference between a health record technician and a health record administrator appears to be mainly administrative in nature. The anatomy, physiology, medical terminology, disease process and coding skills associated with the one-year technician students, is the same offered in the two-year training program. Administrators are further trained in administrative skills and analysis of the health record databases.

Unlike its predecessor, the program at Douglas College did not offer evening courses for ARTs to upgrade to the HRA level early in its inception. Recently, however, the one-year program at Douglas college has been phased out and evening education classes have started to enable progression from a technician to an administrator. These directives are in response to the changes recommended by the CCHRA with its INFOCUS program, discussed in section 1.1.2.7. All health record personnel in British Columbia currently are being trained at the administrator level. There are few technician-level training programs remaining at Canadian institutions.³

1.1.2.4 Degree-Granting Programs

During the 1960's, the CAMRL attempted to match the education of the AMRA, which was granting four-year baccalaureate degrees at that time. A four-year baccalaureate degree granting institution for MRLs was developed in 1963 at Notre Dame University in Nelson, B.C.. Graduates were eligible to write registration exams for MRLs. But the

³It should be noted that the one-year correspondence course offered through the Canadian Hospital Association has also been changed to a two-year program.
program was short-lived. The program lost its accreditation status with the CAMRL in 1973, and was eventually closed in 1977.

The Canadian association was unsuccessful in its attempt to require further degree programs. The HRABC in 1975 presented a proposal for a new four-year program to the University of British Columbia Faculty of Medicine, but was rejected (Barer et al., 1982). The national association finally acquiesced and the two-year program became the standard for the Canadian health record field. Szabo (1980) states that, unlike the AMRA, the CAMRL chose not to seek the sponsorship of the Canadian Medical Association and tried to rely upon its own resources. Perhaps sponsorship by the Canadian Medical Association might have produced a different result.

In September 1982, the four-year baccalaureate program for Health Information Science (HIS) was started at the University of Victoria in British Columbia. Health Information Science is the study of the nature of information and its processing, application and impact within a health care system. Although the desire of the health record association was to see the evolution of this program as the replacement for the Notre Dame program, this was not to be the case. The program is a four-year Cooperative Education program. Curriculum is quite different from the health record curriculum. Currently HRA graduates are given credit for some, but not all, courses in the HIS curriculum in Victoria.

1.1.2.5 Representation on Education Advisory Boards

Health care professional groups do not have representation on the Education Advisory Boards of health information programs. In British Columbia, none of the medical
professions, hospital administrators, medical researchers, statisticians, financial experts, or other health care professionals are involved in the training of health record personnel at Douglas College in New Westminster, which currently operates the only formal training centre for health record personnel in the province. Currently that program's advisory board consists of practising health record personnel, a representative from the B.C. provincial association, an informal representative from the national association (i.e., a member who is a member of the CHRA, but not necessarily appointed by the CHRA body), representation from HMRI and one representative from the provincial Ministry of Health.

1.1.2.6 Continuing Education & Progression

In Canada, once certified as an ART or HRA, there is no required upgrading for the duration of an incumbent's career. Credentials cannot be removed and progression can be attained without registration in a formal education program. Although progression from the lowest rank of associate to the highest rank of fellow is possible, to date, no one has obtained the rank of fellow. Progression is based on a combination of further education that is undefined, on an accumulation of credits for attendance at association events, and for work experience. Szabo (1980) claims this progression is based on a lack of meaningful standards and process is not clearly established. For instance, a total of 60 credits were required to upgrade from an ART/HRT/HDT to a HRA. However, one technician may receive half the total required credits for the one-year correspondence course for Hospital Department Management sponsored by the CHA, while another may not receive any credits at all for

4 CHA requirements
completion of a university degree in a relevant field, such as administration. Since the main difference between the one-year and two-year programs appears to be administrative, the rationale for which member receives what credits is purely the subjective interpretation of the education committee. Szabo also claims that the level of education of an HRA or an ART/HRT/HDT corresponds to the status levels within the association only, and is not related to status within the clinical setting. This appears to be as relevant today as it was in 1980.

1.1.2.7 **INFOCUS**

The CHRA is committed to the continuing development and delivery of its strategic plan. In May 1991, the Board introduced INFOCUS, a project to address the profession’s future role and educational requirements. This project involved a series of strategies to define practitioner competencies, enhance curriculum, determine the academic facilities best suited to deliver the education programs and implement all changes and enhancements at all required levels. INFOCUS recognizes the need to train health record personnel better to meet the changing requirements within the health care industry (Reece, 1991).

INFOCUS identified a need for a degree program for health record managers but have not currently identified a viable university site. It is working with the University of Victoria towards better transfer of credit for health record personnel into the Health Information Sciences Program in Victoria.
However, in February 1995, the CHRA Board and INFOCUS Management Committee recommended a move to a single future professional⁵, the Health Information Professional. This recommendation stemmed from feedback from association members intending to clarify professional identity, increase their presence by taking a strong leadership role in health information, and play a part in streamlining the health care infrastructure. This move is in line with the move to eliminate the one-year technician programs and change the correspondence course to a two-year course. Bridging mechanisms will enable associate and certificant members to advance into this new category, leaving both the associate and certificant categories dormant over time.

There is a great deal of variation in the level of skills required within a health record department. These skills range from clerical to coding to administrative. Given the diverse range of skills required, there could be several tasks that do not require skills intended for a "health information professional". This move appears to be similar to an earlier move by the Canadian Nurses Association to over-credentialize nurses in their move towards "Baccalaureate by the Year 2000".

⁵Letter dated 28 February 1995 from the CHRA to its membership.
1.1.2.8 Other Canadian Training Programs & Workshops

Staff of the Hospital Medical Records Institute (HMRI), now part of the Canadian Institute for Health Information (CIHI), conduct training seminars across the country for new hospitals joining their system (Winger & Wu, 1988). Basic Training includes International Classification of Diseases, Ninth Edition, (ICD-9) and ICD-9CM coding and the HMRI (CIHI) abstracting. Both ICD-9 and ICD-9CM are reviewed in section 1.5. Errors, Corrections and Report Interpretation is offered as advanced Basic Training to all participating hospitals. However, a knowledge of health record curricula and experience with classification systems is assumed to exist already.

A refresher correspondence course in Coding and Diagnosis Typing through the Canadian Hospital Association is also available. The course is designed for people who wish to review basic coding and diagnosis typing theories and to keep current with changes in this field. A certificate is granted for students who complete all six units and achieve an overall average of 60%.

1.1.3 Coder Certification

So exactly what is a coder? Coding skills across the prospective health record designations should be interpreted as being equivalent, given the basic understanding that the difference between the designations appears to be only administrative. Although most health record clerks have no medical training (Wood et al., 1989), they do perform an extremely important task. Coding is normally a function of the health record department. The task, by tradition, is performed by both health record administrators and health record technicians.
Albeit an important function, coding is a mundane and clerical role. It takes a special kind of person to perform the task successfully.

There is no specific coder certification in Canada, nor has it been contemplated at any level of health record association for personnel recognized as coding specialists. On the other hand, coder certification in the United States has been incorporated by the AHIMA since 1992 (Skurka, 1994). Coder certification is a designation given to the health record technician or health record administrator for recognition of the special skills required for coding within the health record field. In response to the study completed by Hsia et al. (1988), a task force was established by AHIMA to review the improvement in the quality of coding. It recommended a certification program for coding specialists. This was based on a concern that the national examinations for MRT's and MRAs tested only entrance-level competency in coding and classification systems and that the average percent on the coding section was 68% (Amatayakul, 1991). There was no other mechanism in place to measure the competence of a coder.

Only recently has there been a standard training program for coding specialists. Coder certification in the United States requires high school graduation. The examination for certification is one full-day and is based on knowledge and skills beyond the entry level as examined in the national exam for ARTs and RRAs. Professionals other than those with an ART or RRA designation, such as registered nurses, are also eligible to write the exam. The exam consists of questions on anatomy and physiology, medical terminology, disease processes and coding principles. It is a rigorous format consisting of multiple choice and open-ended questions. A second section involves the coding of actual medical records. In
1993, the exam consisted of fourteen in-patient cases representing a variety of diseases and increasing difficulty (Amatayakul, 1994). The passing grade was set at 80%. The process for preparing testing material was extensive. The entire country was canvassed for record samples. An initial consultant identified three times the number needed for the initial screening. An expert panel recommended 14 inpatient cases for the 1993 examination (one less than the 1992 examination). Another panel of coding specialists developed multiple choice questions. Sixty questions were drawn for the examination. The examination itself was pilot tested.

Results for the 1992 examination showed 46% of candidates (3677) passed, while 52% of candidates (1251) passed in 1993 (Amatayakul, 1994). The mean score in 1992 was 78.9%, while in 1993, the mean score was 77.8%. The highest score in 1992 was 93%; in 1993 it was 95%. Other personnel, such as registered nurses, did not fare as well as the health record profession, and even worse than candidates with no credentials at all. Only 39% of the other credentialed personnel passed the exam. When broken down by number of years experience, only 49% of those candidates with one to two years experience passed and only 22% with less than a year passed. Credentials were not a variable in the years of experience analysis.

Should coders have the ability and responsibility to infer the codes based on the documentation in the record? If so, health record coders need to be sufficiently educated to challenge the diagnoses or documentation within a chart. It must be recognized that the coder does not have clinical evidence, other than that which is documented in the health record. When guidelines clearly state that the Most Responsible Diagnosis (MRD) is that
diagnosis which determines the longest length of stay, the authority should lie within the documentation within the entire chart, not that listing of diagnoses documented solely on the face sheet, the first page of information of the health record. Coders need to be educated at a level where they can demonstrate competence at coding. Competence of those in the profession is essential to ensure the quality of the data within the database. The standard of the profession needs to be high to ensure a high quality of data. Coders should also be tested on a periodic basis to ensure intra-rater reliability on the accuracy of the codes. Therefore, the specialty field of coding needs to be recognized as an addition to the basic health record education preparation to ensure quality data systems that are consistent throughout the nation.

1.1.4 Is there a Future for Health Record Personnel in Canada?

In her master’s thesis, The future of health record personnel in Canada, Szabo (1980) clearly documented the problems faced in the industry with the lack of educational consistency in training. Since that time, the nature of health record education has not changed much. MRAs may have become HRAs by virtue of a name change, but they have not overcome their deficiencies. Courses have changed from one-year to two-year, but curriculum has stayed relatively the same. However, the educational needs of the occupation are in a continual state of flux. Name changes and program lengths are currently being discussed, but program content remains virtually unchanged.

---

*Details on the content of the face sheet are found in section 1.8.3 on the coding process.*
Szabo (1980) claimed that neither local, regional nor provincial level participation is sought from the HRA occupation when decisions are made regarding some aspects of health information. Even today, the B.C. Ministry of Health has not had a health record consultant on staff since the mid-eighties. HRAs, she claimed, gave all their attention to the clinical needs for information and did not develop the expertise necessary to use this information for management and for planning health services. She recommended more extreme curriculum revision.

In Canada, education programs for health record personnel are quite different from those in the United States. The national associations have not maintained the same educational requirements for their respective members. The Americans have maintained four-year baccalaureate programs for RRAs all along and are currently seeking programs at the master’s level. RRA professionals in the U.S. sit on several regional and national government committees. The association supports the health record profession in two tiers, the managers and the technicians.

Coding patients’ medical condition is an area of considerable difficulty, yet the importance of coding health records has not been sufficiently recognized in Canada. The controversy over the utilization of health record staff is that they are either underutilized according to the health record profession (Dietrich, 1983; Walton, 1983; Pursel & Terry, 1986; Hogan, 1987; Harper & Holmes, 1988) or under-trained, according to others (Hendrickson & Myers, 1973; George & Maddocks, 1979; Szabo, 1980; Bulatow, 1988).

The role of the health record departments within the health care field is constantly changing. The profile of health record departments and that of the health record professional
is on the rise along with the value of the information that flows from those departments. Health record associations in both the United States and Canada are trying to maintain their role of "keeper of the records" at the same time trying to enhance the role of the health record profession as the guardian of health information.
1.2 HEALTH RECORDS AND DATABASE SYSTEMS IN NORTH AMERICA

One source of information for patient-centred health databases is the hospital health record, a fundamental recording of the events surrounding each hospital in-patient stay. Health records are, and have always been, an essential information element in the practice of medicine. This section reviews the differences between health records and clinical databases. It also describes a few independent healthcare database vendors that provide information to various governmental agencies.

1.2.1 Evolution of Health Records

A medical record, by tradition, implies the medical record of a patient treated by a physician (Waters, 1983). A health record is used here to refer to a clinical document which records the events, treatments and/or observations, and describes facts, opinions and knowledge relating to the patient's treatment during a hospital stay provided by all professions involved in the treatment of the patient. The difference between a medical record and a health record, or clinical document, is the all encompassing use of the latter record by all health care professionals in all levels of health care delivery, rather than by physicians alone (Skurka, 1994). Among many health records' uses, they promote communication among all health care professionals treating the patient so that continuity of care can be provided. They facilitate medical education and research, and provide documentary evidence of the course of the patient's illness and treatment for third parties concerned with the patient (Huffman, 1972; Blois, 1984).
The health record itself is the foundation of coding accuracy. The major responsibility for accurate coding resides in the health record department in conjunction with physicians and other health care professionals. However, extracting the data from the health record for secondary databases traditionally has been the responsibility of health record coders. Coded data are, in turn, the foundation that supports health care planning, clinical program evaluation, and in some provinces, funding of the health care system. The American Medical Record Association (AMRA) stresses that to achieve the highest accuracy of coded data, qualified and competent coders must be employed (AMRA, 1987b). Key problems in going from a health record to a database are the accuracy of the information in a health record and the process of the conversion.

The earliest purpose for the coding of health records was to provide access to them by means of diagnoses and procedures, so that the original records could be retrieved for medical research (Thompson & Slee, 1978). Today, coding is not done simply as a retrieval guide for the researcher. Diagnostic and operation codes are used to make decisions regarding the management of health delivery systems. Automation of health information systems has created an explosion of information useful throughout the entire health care delivery system.

1.2.2 Clinical Databases

A clinical database contains retrospective data. It is designed to summarize the record for easy access to cumulative patient data. Although it can be utilized to access one individual’s clinical data, its purpose is to enable a review of the cumulative data, such as the
number of patients admitted with a specific diagnosis or the number of patients treated by a specific physician. There was a general movement from an interest in individual patients to an interest in a group of patients or populations with particular characteristics. Generally, decision-makers have little interest in the entire content of a single record. Rather, they require a subset of records, or a set of specific items from all records for analysis. For example, descriptive epidemiology, the study of the distribution of disease, is of central importance in health planning (Reinke, 1988) and requires some information from all records in a database.

The first step in a planning process is to describe, forecast and analyze a system's activities. This is where the database can be a useful tool. Historical data provide a basis for initial forecasts. However, the availability of data alone is necessary, but not sufficient for successful health care planning (Reeves et al., 1984). To provide statistics for planning within the health care system, databases must not be utilized in isolation from other factors. Planning also requires knowledge of future circumstances to choose between appropriate alternatives. Societal values and political considerations also enter the formulation of the forecaster and decision-maker. Statistics will be quoted and used by those responsible for planning and resource allocation. Therefore, accuracy of a database is vital for the successful utilization of it.

Professor Blois (1984) claims that the most difficult process in recording and processing medical data is the construction of a suitable classification system and the representation of the clinical data so that they can be processed by a computer program. The International Classification of Diseases (ICD) utilized within the computer programming by
the Hospital Medical Records Institute (HMRI) is an example of such a process. A review of the HMRI database and the ICD Classification system is necessary to understand the processing of information. Both are described below in section 1.5 and 1.7A.

1.2.3 Independent Systems

There are several private, non-profit hospital discharge database systems in North America. Throughout their history they have evolved, operated and merged with other systems. In the 1920's, the Committee on the Costs of Medical Care in the United States collected the first national data on the socioeconomics of medical care (Funk, 1973). By 1935-36, the Public Health Service National Health Survey was the first source of information on the use of health statistics. By the 1950 and 60's, several health data abstracting systems emerged. All were independent, but had similar system characteristics and data items. The minimum dataset collected by the majority of the systems was fairly common and consisted of such variables as the hospital number, patient number, date of admission, date of discharge, principal diagnosis, secondary diagnoses, complications, attending physician, principal procedure, secondary procedures, date of procedure, and surgeon. The minimum dataset has been fairly standard between the systems and over the years and is fairly representative of what exists today.

A few of these hospital discharge systems utilized by Canadians are listed below: the Professional Activity Study (PAS), the Hospital Medical Records Institute (HMRI), and the Canadian Institute for Health Information (CIHI).
1.2.3.1 *Professional Activity Study (PAS)*

A prominent health data abstracting system utilized by most Canadian hospitals was the Professional Activity Study (PAS). Hendrickson & Myers (1973) claim that it was the largest single repository of patient information in the United States. PAS utilized the International Classification of Diseases (ICD) classification system. For a service fee, both Canadian and American hospitals shared the computer facility and the staff.

The PAS was established in 1953 under grants from the W.K. Kellogg Foundation (CPHA, 1959; 1972). It was founded as a national body to provide information to medical staff and hospital administrations on patterns of clinical practice. It was operated by the Commission on Professional and Hospital Activities (CPHA), a non-profit non-governmental research and education centre dedicated to the improvement of hospital and medical care. PAS was sponsored by the American College of Surgeons, the American College of Physicians, the American Hospital Association and Southwestern Michigan Hospital Council (CPHA, 1959; 1972; Funk, 1973). Its members represent the fields of medicine, hospital administration, nursing, health records, education, systems analysis, data processing and statistics. In 1972, its membership consisted of 1,450 hospitals. In 1975 membership peaked at 6,466 hospitals. By 1983 membership had dropped to 919 hospitals. Today, PAS has been incorporated into a larger system.

1.2.3.2 *The Canadian Hospital Medical Records Institute (HMRI) Database*

Prior to the introduction of the first Canadian clinical database, a major contractor for Canadian hospitals was CPHA in Ann Arbor, Michigan which utilized the Professional
Activity Study (PAS). Until April 1987, most Canadian hospitals were serviced by the CPHA.

HMRI was established to better meet the needs of the Canadian hospital population. It was a not-for-profit, federally chartered medical data processing company that processed hospital discharges across Canada since 1963 (Boyd, 1985; Yungblut, 1988). HMRI was reorganized in 1986 as a joint venture between the CPHA of United States and the Hospital Medical Records Institute (HMRI) of Canada (McLean & Winger, 1986a). The HMRI company was founded by the Ontario Hospital Association (renamed the Ontario Health Association) and the Ontario Medical Association (Boyd, 1984; Croke, 1985; Yungblut, 1988). The HMRI system is a retrospective record classification and indexing system for coding and abstracting for both hospital in-patients and ambulatory patients. The main focus of this database system, however, is in-patient utilization.

HMRI was initially started in Ontario, but gradually expanded to other provinces. In 1972, the Ministry of Health (MOH) in Ontario required all acute care institutions to utilize the HMRI database (Boyd, 1985).

During the 1970's, the majority of hospitals in B.C. began participating in abstracting/data processing programs (Nusbaum, 1983). The MOH in British Columbia has recognized that reporting and analysis of discharge information was, and still is, considered essential to the administrative and medical management of medium-sized to large hospitals (Nusbaum, 1983). However, smaller hospitals could still provide information through admission/separation records. The data were then further coded and analyzed by the Ministry. Since April 1983, the Ministry required hospitals to provide data through their
abstracting/data processing vendors, such as HMRI and CPHA, rather than providing the admission / separation records directly. HMRI assumed medical record data collection services on behalf of CPHA in the fiscal year 1987. Other systems were used initially, but the HMRI system has been maintained and other vendors either use the HMRI software or have adapted software to fit the HMRI system (see HMRI’s listing of accepted vendors in Canada in Appendix B). The advantage was that the same data used by the hospitals for research, planning, and budgeting would now be utilized by the Ministry to perform its functions. This produced a significant reduction in the data collection workload. Instead of completing both a government abstract and an abstract to serve their own requirements for each inpatient separation, hospitals would only perform the function once.

HMRI has been widely used since 1987 and currently maintains over 80% of Canadian hospital in-patient discharges (Hickie, personal communication) which consists of over 600 hospitals from coast to coast (Crewson, 1994). In 1994 it was recognized as the national body for patient-centred health care information in Canada. The database has been expanded to include data on day surgery, rehabilitation, and chronic diseases.

HMRI’s Board of Directors represented the health care community that they served. Membership consisted of the past Chairman and current Chief Executive Officer at HMRI, and representatives from each province as determined through each provincial User Group (Yungblut, 1988). The number of members were determined by the volume of in-patient cases received from each province. Each province had at least one member, with a maximum of four from the largest provinces.
The HMRI User Group in each province had representatives from its respective provincial Ministry/Department of Health, the provincial Medical Association, the provincial Health Association, and the provincial Health Records Association (Yungblut, 1988). The main purpose of each group was to keep the region advised on HMRI activities and to keep HMRI appraised of the effectiveness of its services in the region by monitoring existing programs, evaluating suggestions for change or modification both regionally and nationally, recommending appropriate regional modification as well as contributing to and evaluating educational services (McLean & Winger, 1986).

The database provides the health care community with access to local, provincial and national data and provides a broad range of comparative data. The data can be extracted by single hospitals or groups of hospitals, making inter-hospital comparisons possible. However, as the literature review demonstrates, because of the lack of standards and quality evaluation to ensure comparable data, very little inter-hospital comparison actually occurs.

This database allows Canadian hospitals to access national and provincial information and provides the ability to compare data and base decisions on information that is unique to the Canadian population. The establishment of this database by the two organizations fitted with both organizations’ objectives - the improvement of patient care and hospital efficiency through high quality of patient care data.

HMRI information can be utilized for epidemiologic and utilization studies to determine information such as disease prevalence and patterns of hospital care. It can be potentially useful in a variety of activities including quality assurance, utilization reviews, risk management, research projects, accreditation, and government reviews. It may even
play an important role in determining policy directions. Research can provide guidance for making policy or management decisions about procedures, programs, planning, operations and finances.

1.2.3.3 The Canadian Institute for Health Information (CIHI)

In 1989, the National Health Information Council was created through the efforts of the provincial Deputy Ministers of Health to improve health information in Canada. In May 1990, this council set up a National Health Information Task Force which recommended creating a national institute to coordinate development of national health information systems. The Canadian Institute for Health Information (CIHI) incorporated as a national not-for-profit corporation in December 1993 and by January 1994 both the Management Information Systems (MIS) Group and the HMRI merged their functions under the CIHI umbrella.

The corporation has since combined some resources of both Health Canada and Statistics Canada - Health Division. The mission of the institute is "to provide and coordinate the provision of accurate and timely information required for the establishment of sound health policy, the effective management of the Canadian health system, for generating public awareness about factors affecting health, while at all times respecting the personal privacy, and safeguarding individual record confidentiality and system security" (Crewson, 1994; CIHI, 1994). Their role is to develop a comprehensive health information system for Canada that will, among other functions, develop and promote standards and guidelines for national health information. The 28-member Board of Directors represents a balance between government and non-government members (CIHI, 1994). Membership includes national and
provincial representatives and a broad range of health care professionals. Representatives include the medical profession, the provincial health associations, chief executive and chief operating officers, statisticians, and researchers. All stakeholders are broadly represented. CIHI has maintained the provincial User Group implemented by HMRI.
1.3 NOMENCLATURES AND CLASSIFICATION SYSTEMS

Information stored in health databases is essential to health planning and research. The identification and classification of disease processes is essential to utilizing this information for comparison and statistical purposes.

A medical nomenclature is simply a listing or a catalogue of approved names for morbid conditions (WHO, 1977; Côté & Rothwell, 1989). A nomenclature, by necessity, must list every recognized disease irrespective of its significance. It standardizes terminology. The arrangement of entries is without concern for their frequency of occurrence (CPHA, 1973). The standard medical nomenclature utilized in medicine today has evolved through a recognized system of preferred terminology for naming diseases. Not all terms in a nomenclature are listed. The nomenclature attempts to have each disease labelled under one term. As medicine advances, so the terminology adapts to address the changes.

A classification system is a method of generalization (WHO, 1977), of describing and recording clinical and pathological observations (Côté & Rothwell, 1989) to produce necessary statistical information (Huffman, 1985). It emphasizes the grouping of related entities. Unlike nomenclatures, classification systems try to contain all terms, not just terms considered appropriate by the nomenclature. A medical classification system standardizes medical conditions or procedures which are grouped together. Classification systems are a way of organizing the data for easy accessibility and retrieval.

A medical classification of diseases may be defined as a system of groups or categories to which morbid entities are assigned according to some established criteria (WHO, 1977) or related phenomena (Côté & Rothwell, 1989). A medical classification system
standardizes the medical conditions grouped together while a medical nomenclature standardizes terminology. The classification must be confined to a limited number of categories that will encompass the entire range of conditions (WHO, 1977). A specific disease entity should have a separate title in the classification only when its separation is warranted because of the frequency of its occurrence, or its importance as a morbid condition. Every disease or morbid condition must have a definite and appropriate place as an inclusion in one category of the classification.

For scientific purposes, according to Feinstein (1988), a system of classification must fulfil three major requirements:

1. the classification must have a suitable organizing principle;
2. the categories must be labelled with standard titles;
   and
3. the members of each category must be identified with a suitable set of operational criteria.

Huffman (1985) adds that the organizing principle needs to be a single one, such as anatomical site, that the categories must be exhaustive to include every possible entity available, and that membership within each category must be mutually exclusive so that a given term is placed in one and only one category.

The scientific requirements of a classification must allow the nosology to be useful in its focus, and to be standardized, uniform, and reproducible. After categorical titles are developed, operational criteria are needed to ensure that the same things are catalogued consistently in the same way. There may be several possible choices for these criteria. The classification chosen will depend upon the use of the statistics compiled.
No single classification will fit all the specialized needs of all the groups utilizing such a system, but it should provide a common basis of classification for general statistical use, such as storage, retrieval, tabulation and analysis of data (Feinstein, 1988; WHO, 1977). Coding a medical term represents an alphanumeric equivalency to the language of words which contain the disease entity. Alphanumeric coding of natural language allows the data within the health record to be organized. Symbolic codes, such as these, are essential to allow the data to be manipulated by computer technology. Classification systems enable a unique code number for each and every concept. The main classification system used throughout the world today is the International Classification of Diseases (ICD).
1.4 HISTORY OF THE INTERNATIONAL CLASSIFICATION OF DISEASES

This section reviews the history of the International Classification of Diseases (ICD) up to the current version, the ICD-9. It also describes the Council on Clinical Classifications (CCC) and the Systematized Nomenclature of Medicine (SNOMed).

The first known attempt to classify disease in humans was in the 17th Century with the London Bills of Mortality, when John Graunt estimated the proportion of liveborn children who died before the age of six years (Huffman, 1985; Côté & Rothwell, 1989). Thirteen classes of disease in young children were listed.

Prior to the 19th Century, François Bossier de Lacroix, better known as Sauvages, and William Cullen published works entitled Nosologia Methodica and Synopsis Nosologiae Methodicae, respectively (World Health Organization (WHO), 1977; Côté & Rothwell, 1989). The classification most generally used was that of Cullen.

By 1853, the First International Statistical Congress was held in Brussels and headed by Dr. William Farr, the Congress’ first medical statistician (WHO, 1977; Feinstein, 1988; Côté & Rothwell, 1989). Farr discussed principles that should govern a statistical classification of disease and urged the adoption of a uniform classification based on anatomical site. By 1855, a simple classification of 138 separate groups was designed under five major headings, classified by anatomical site, epidemic diseases, constitutional disease, developmental diseases and trauma. These were subsequently revised in 1864, 1874, 1880

---

7The authors have cited Greenwood, M. "Medical statistics from Graunt to Farr" in Biometrika, 1942, 32:204
and 1886. Farr's anatomically oriented classification became accepted as the major basis for a classification (Feinstein, 1988).

In 1893, Dr. Jacques Bertillon developed the *Bertillon Classification of Causes of Death*, based on Farr's principle of anatomical site. The first grouping had 44 titles, which expanded to 99 and then 161 separate headings (Côté & Rothwell, 1989). General diseases were distinguished from those localized to anatomic sites (Feinstein, 1988). In 1898, the American Public Health Association recommended its usage in the United States and Canada and further suggested a revision every ten years. ICD-1 was born. It consisted of 3-digit codes. The ICD is a monoaxial system in that it offers only one name for each disease. There were revisions in 1910 and 1920. In 1920, the ICD-3 had only 205 codes (Slee, 1978). The principle of classifying diseases by anatomical site has survived to this day as the basis of the *International List of Causes of Death*. After Bertillon's death in 1922, the Health Organization of the League of Nations published the fourth and fifth revisions (1928 and 1938 respectively).

In 1946, the World Health Organization (WHO) accepted responsibility for the sixth revision. In 1948, the sixth revision was published under *International Statistical Classification of Diseases, Injuries, and Causes of Death* (ICD-6). This revision marked a new era in international health statistics. For the first time, lists for the tabulation of morbidity and mortality were included (WHO, 1977; Huffman, 1985; Feinstein, 1988). International rules for selecting underlying cause of death were agreed upon (WHO, 1977; Feinstein, 1988). The conference also recommended a program of international cooperation including the establishment of national committees to coordinate statistical activities and to
serve as a link between national statistical institutions and the WHO (1977). Member nations agreed by treaty to abide by the classification and rules in publishing health statistics.

In the early 1950's, the United States took the lead in using ICD for clinical purposes. Clinical adaptations in the United States began with ICD-6 when the groups were subdivided and the rules peculiar to cause-of-death coding were deleted. A consolidation of these changes, Standard Nomenclature of Diseases and Operations (SNDO) was published by the American Medical Association in July 1954. This paper listed the ICD diagnostic codes\(^8\) and for each of them, the SNDO codes that would fall into the ICD group.

In 1955, the WHO sponsored the seventh edition, ICD-7, with only minor essential changes and amendments of errors or inconsistencies.

In 1958 the relative merits of SNDO and ICD for hospital indexing were studied by the American Hospital Association and the American Association of Medical Record Librarians. At that time, the newest version of the ICD, the ICD-7, was too broad in its groupings. This revision did not meet the needs of the physicians or the Medical Record Association (Côté & Rothwell, 1989). By 1959, a working party made up mainly of Medical Record Librarians prepared the first ICD-A, entitled, the International Classification of Diseases Adapted for Indexing Hospital Records by Diseases and Operations. The changes consisted mainly of additions of fourth-digit diagnostic codes\(^8\). The first three-digits denoted the major category or title, while the fourth digit represented more specificity and detail. However, not all the three-digit categories in ICD-A and ICD-7 were compatible. All ICD-A

---

\(^8\)See Appendix A for definition of diagnostic codes.

\(^9\)See Appendix A for definitions of four-digit codes.
codes that were different from ICD-7 were indicated by an asterisk. In 1962, the ICD added an alphabetic index and a procedure classification, but the basic structure of the classification and general philosophy of the classification of diseases remained.

The eighth revision (ICD-8), published in 1969, listed many new causes of morbidity and mortality but left unchanged the basic structure of the classification and philosophy of classification of three-digit codes according to etiology rather than manifestation. That is, diseases are grouped according to the problems they represent. ICD-8 served as the basis for coding diagnostic data for official morbidity and mortality statistics in the United States and proved useful to hospitals in indexing hospital records (Huffman, 1985). The ICD was highly suitable for hospital indexing: coding was somewhat faster and charts could be accessed faster. The name was changed to the *Manual of the International Statistical Classification of Diseases, Injuries, and Causes of Death*. There were 3004 codes (Slee, 1978).

The ICD-8 had to be adapted for American use as it no longer met their needs. It needs to be stressed here that the ICD was designed as an international classification system. Not all countries have agreed to all its applications. Robert A. Israel (1984), Deputy Director of the National Center for Health Statistics in the United States, stated that the ICD is a worldwide statistical standard and, at best, is a compromise between the needs of both developing and developed countries. However, it is the only accepted system that provides a common basis of classification of diagnoses for both national and international statistical use (Wood et al., 1989). Physicians and the American Association of Medical Record Librarians recognized that the classification served the purpose of a statistical classification, but often
failed to provide the specific information needed for meaningful analysis of the records. Thus, the ICDA-8. The "A" no longer meant *Adapted for Indexing Hospital Records by Disease and Operations*, but *Adapted for Use in the United States*. The ICDA-8 was the same as the ICD-8 to the three-digit level. The fourth digit was added which made it different from the ICD-8.

At this time, some hospitals and physicians informed the Commission on Professional and Hospital Activities (CPHA) that some aspects even of ICDA-8 lacked adequate detail (Slee, 1978). The 1960's witnessed a nation striving for specificity. The hospitals, therefore, created their own codes. In 1968, the CPHA published the Hospital Adaptation of ICDA, H-ICDA, for use with the Professional Activity Study (PAS). Five years later (1973), a second edition and revision, H-ICDA-2, was published. H-ICDA-2 was the same structure as the H-ICDA, but approximately 30% of the code numbers were different.

Coding with the ICD system is relatively easy, but retrieval of the data is not. Except for some very general statistics, it is not always possible to extract specific diagnostic data in any detail. For example, a record retrieval study is likely to include many records in which the researcher is not interested, but which were properly coded and retrieved under the same diagnostic code.

1.4.1 **Council on Clinical Classifications**

The Council on Clinical Classifications, a division of the CPHA, was formed in 1976 to form a permanent organization so that clinical medicine could speak with one voice on classification matters (Côté & Rothwell, 1989). The sponsors consisted of the American
Academy of Pediatrics, the American College of Obstetricians and Gynecologists, the American College of Physicians, the American College of Surgeons, the American Psychiatric Association, and the CPHA. Efforts for clinical modification were mobilized under Robert Israel, Deputy Director of the National Center for Health Statistics and head of the recently established WHO Center for Classification of Diseases for North America.

1.4.2 Systematized Nomenclature of Medicine (SNOMed)

At the same time that the ICD-9 was published in 1977, the College of American Pathologists published the two-volume edition of the *Systematized Nomenclature of Medicine* (SNOMed). A second edition was published in 1979. SNOMed is a comprehensive multiaxial nomenclature of medical terms designed for use by the entire health care system (Côté & Rothwell, 1989). A multiaxial nomenclature is one where there may be several components to one code. For instance, a multiaxial system may identify the name of the disease, the severity of the disease, as well as other clinical categories of patient status, such as prognosis (Feinstein, 1988).

SNOMed was a major expansion of the Systematized Nomenclature of Pathology (SNOP) published in 1965 by the American College of Pathologists¹⁰. SNOMed is a two-volume edition consisting of a numeric listing and an alphabetical index. Each entry consists of a five-digit core code representing anatomical site, patients signs, symptoms, problems and disease components as well as final diagnosis.

¹⁰Information on the axes was derived from the College of American Pathologists, Committee on Nomenclature and Classification of Disease.
SNOMed has logical open-ended modules that allow the incorporation of additional axes in future editions. The current edition has seven axes. The first, Topography, is based on anatomical site or structure. This field represents all parts and regions of the body down to the cellular and subcellular structures. Beyond this, the intracellular units have been considered functional units and are found in the Function field, described below. Within the major one-digit sections, there are two-digit headings within which the terms are arranged in a hierarchical fashion up to the five-digit level, depending on the specificity of the description. For instance, the lung is at T-28--- level. The right lung is represented by T-281--, T-282--, T-283--, and T-284--. The right upper lobe, apical segment is represented by T-28210. Most diagnoses should be anchored to some Topography number. The Topography field has almost 10,000 entries.

The second field is Morphology, which contains all the terms that represent a change in the form of normal anatomy of the body. These M codes need to be coupled to a T site for more specificity. For instance, if a health record states the diagnosis of sulfuric acid burn of skin of hand, the topography site would be skin of hand, T-02650, the morphologic change would be chemical burn, M-11150 and the etiologic agent causing this chemical burn is sulfuric acid, E-5026.

The third field or axis is Function, which incorporates all functions, functional states and functional units. Functional units include elements, molecules and compounds and substances normally present in the body and which are often measured to assist diagnosing and monitoring of disorders and diseases, e.g., signs and symptoms, such as headache and fever. The Function field terms stand in some contrast with those in Etiology in that many
will appear to be logically suited to stand alone, without needing terms from other fields to make them clear or useful. However, all of the terms are essentially related to the Topography or other fields. For example, T-70400 (Female Genital System) and F-30200 Menstrual cramps / dysmenorrhea.

The fourth axis, Etiology, expresses some causes or causal agents of pathologic anatomy or pathologic function identified in other fields. These are demographic and epidemiologic in character and are not likely to appear in admission or discharge diagnoses very often. This kind of collection creates a need for detailed supplementary coding or recording.

The fifth axis, The Disease Field represents the classification part of the system. Practically all entries could be further specified by terms from the four previously described axes which make up the nomenclature. Because certain diseases or syndromes consist of multiples of the Topography, Morphology, Function, and Etiology categories, the Disease Axis, was created to classify the disease. This axis is equivalent to the International Classification of Diseases (ICD). The most common example is tuberculosis. See this example shown in Figure 1, Integration into a classification of disease for SNOMed.

The sixth axis, called Procedure, is the administrative, preventive, diagnostic or therapeutic action of the health care team. These are to help the coder see possible relationships with other fields in order to code a diagnosis accurately. For example, a frequently used medical term cauterization effect (P-1503) can be found under burn injury (M-11100). The reference to P means that the burn injury called cauterization effect is related to a procedure called cauterization and this procedure should be coded also. The
procedural activity (P) is usually linked to the appropriate site (T) to express the entire medical concept. The Procedure field consists conceptually of three parts: Administrative and General Procedures, Diagnostic and Monitoring Procedures, and Preventive and Therapeutic Procedures.

The seventh axis is an open-ended procedural module.

Both Huffman (1985) and Skurka (1994) designate SNOMed as the most comprehensive nomenclature available in the health care field. The purpose of the creation and design was for computer handling and storage to provide a medical coding system that would be more specific and have a statistical classification to enhance retrieval of information. It would enable transfer of the information contained in the medical record to a codable representation.

![Diagram of Nomenclature and Classification]

**Figure 1: Integration into a classification of disease for SNOMed**

*Adapted from Côté & Rothwell, 1989*
In 1982, SNOMed was reprinted and updated for the third time. There were close to 50,000 discrete entities encompassing the entire field of medicine including each of its specialties. The CHRA was the first national association to sponsor workshops of SNOMed concepts and coding (Côté & Rothwell, 1989).

An advantage of a multiaxial system is that it allows maximal detailed expression of medical concepts. It ensures that every medical concept can be coded, stored and retrieved. It allows the full expression of important relationships, such as cause and effect. It allows for an infinite number of unique combinations of medical terms and concepts. A limitation is that more digits are required for storage and that several codes are needed to express one concept, thereby the accuracy of the data could be more at risk. Some uses of medical concepts may not require that much detail.

Côté & Rothwell (1989) have termed SNOMed an integrated nomenclature-classification system. The authors feel that future integration with the ICD would resolve the nomenclature-classification issue. However, to date, the WHO is not changing the basic structure of the ICD. The authors also claim that multi-axial coding of medical information presents obstacles to coding, but feel these obstacles could be overcome with the advent of an encoding system. Automated encoding would be a desired technological advantage with such a system to help prevent clerical errors from occurring. Currently a computer program is being developed that will create an equivalency table from SNOMed classifications to ICD codes (Côté & Rothwell, 1989).
1.5 ICD-9 CLASSIFICATION SYSTEM

This section reviews the ICD-9 System, the American adaptation, ICD-9CM, Canadian usage of the ICD-9CM and Problems with the ICD-9.

The *International Classification of Diseases, Ninth Revision, 1977* (ICD-9) is a two-volume index, monoaxial statistical classification system, which became effective on January 1, 1979 (Israel, 1984). The one axis provides only one name for the disease entity: it does not provide for other factors, such as severity of disease. The entities included morbid and mortal diseases and classification of impairments and handicaps. The *Classification of Procedures in Medicine* for therapeutic and diagnostic procedures was also published with the 9th revision. The ICD is used to translate diagnoses and procedures into alphanumeric codes that allow easy storage, retrieval and analysis of data.

The general arrangement of the ninth revision was much the same as the 8th, but with much more detail. The purpose was to secure uniformity in the recording of causes of death. In ICD-9, diseases are grouped according to the problems they present. For example, all neoplasms, regardless of anatomical site, are grouped under the neoplasm system and tuberculosis is grouped under infectious and parasitic diseases.

Care was taken to ensure that the categories were meaningful at the three-digit level. Although the lack of balance within the disease categories of the 8th revision remained, there were five major innovations in the 9th revision (WHO, 1977). These include:

---

\[\text{See Appendix A for definitions of diagnostic and procedural codes.}\]
ICD-9 consists of 17 major categories, broken down into approximately 800 three-digit codes with further subclassification to four- and five-digits, separated by a decimal point. The purpose of the fourth and fifth digit subclassification is to give greater specificity of diagnosis. For example, diabetes mellitus is given a three-digit code of 250. Diabetes mellitus with ketoacidosis is given a four-digit code of 250.1. Diabetes mellitus with ketoacidosis, juvenile type onset is assigned a five-digit code, 250.11. The fifth digit is optional\textsuperscript{13}. At the time of printing, ICD-9 had three times as many codes as ICD-8, 9607 versus 3004, respectively (Slee, 1978).

Volume 1 of the ICD is the Tabular List containing the alphanumeric diagnostic categories. The terms used in the categories of the Tabular List are not exhaustive; they serve as examples of the content of the category. For example, the three-digit code 411 in Volume 1 has several examples under the title, "Other acute and subacute forms of ischaemic heart disease". The terms unstable angina, intractable angina, crescendo or accelerated

\textsuperscript{12}See Appendix A for definition of fifth digits.

\textsuperscript{13}HMRI drops the optional fifth digit when submitting data to the appropriate Ministries/Departments of Health.
angina are not specified, but fall within this code. To date, the ICD-9 contains over 10,000 specific diagnostic classes. The introduction of Volume 2, states that it is not possible to express all variations in the assignment of a diagnostic term in the Alphabetical Index; therefore, Volume 1 should be regarded as the primary coding tool (WHO, 1978).

Volume 2 of the ICD is an Alphabetical Index to the Tabular List of Volume 1. The Index is an essential adjunct to the Tabular List since it contains many diagnostic terms that do not appear in Volume 1. There are over 120,000 medical terms in common usage (Thompson & Slee, 1978; American Medical Record Association (AMRA), 1987a; 1987b). The classification system attempts to accommodate all medical terms, whether standard, colloquial, old or new. It is intended to include all diagnostic terms currently in use. The Index is organized as lead terms with various levels of indentations. The code numbers that follow the terms in the Alphabetic Index are those of the three-digit categories to which the terms are assigned in Volume 1. For instance, Angina (attack) (cardiac) (chest) (effort) (heart) (pectoris) (syndrome) (vasomotor) is followed by the three-digit code 413. The first level of indentation of the subclassifications includes, among others, decubitus 413, pre-infarctional 411, and streptococcal 034.0. In general, if the three-digit category is subdivided into four-digit categories, the appropriate fourth digit is also given in the code number in the alphabetical index. Since its inception, the ICD has added several categories.

The ninth revision reflects the broadening use of the classification in the indexing of medical records, the indexing of ambulatory care records, and the grouping of data for evaluation of medical care. A fourth use, the furnishing of clinical descriptions of patients, is a new and largely unrecognized development (Slee, 1978).
One major innovation in the ICD-9 is that it provides for dual classification in some categories which can combine two descriptions of disease entities into a single code number (WHO, 1977; Huffman, 1985) or provide an alternate code for placing certain conditions in classes determined by cause of manifestation, depending on the interests of the user (WHO, 1977; Slee, 1978). However, the manifestation part of the code is optional and it is only available for a selected number of categories. For example, specified pathological lesion in the kidney with diabetes mellitus is coded as 250.3, diabetes with renal manifestations. This is the compulsory primary code. An optional code to reveal the exact manifestation is 583.8, the code that specifies the pathological lesion in the kidney. The three-digit code, 250, falls under the endocrine, nutritional, metabolic and immunity disorders, while the three-digit code, 583, falls under the genitourinary system.

Although the main advantage of a single axis coding system is simplicity, it also implies limitations in the expression of larger concepts. Each successive digit in the decimal system expands the possibilities tenfold. Single axis does, however, fit some special purposes where the total number of variables is small.

1.5.1 The American Adaptation, ICD-9CM

The Americans developed an adapted version, the *International Classification of Diseases, Clinical Modification*, 1978 (ICD-9CM), based on WHO’s ICD-9. ICD-9CM was developed in May 1978 in response to a need by hospitals for a more efficient basis for storage and retrieval of diagnostic data (Slee, 1978; Israel, 1984; AMRA, 1987a; 1987c).
ICD-9CM is administered as a division of CPHA and is sponsored by the original Council on Clinical Classification group. It is supported by the United States National Center for Health Statistics. Despite the improvement in ICD-9 over ICD-8, the CCC persuaded the federal government in September 1976 that a modification was needed for American clinical use. The Americans felt that ICD-9 was adequate for coding and analyzing causes-of-death as reported on death certificates, but it did not contain enough detail for the growing needs of hospitals and other users in clinical settings.

To paint a clinical picture of the patient, the codes needed to be more precise than those needed only for statistical grouping and trend analysis. Therefore, for morbidity applications, a modification was introduced which maintained full compatibility with the international version but permitted fifth digit expansion in some categories.

The disease classifications have been expanded to include health-related conditions and to provide greater specificity at the fourth and fifth digit level of detail. For instance, in ICD-9, acute myocardial infarction (AMI) is coded as 410, a three-digit code. There are no fourth or fifth digits available. In ICD-9CM, AMI can be coded to incorporate specificity and location. It provides detail about whether the patient is receiving care for initial treatment of AMI or whether the patient is being observed, treated or evaluated within eight weeks of onset of AMI, (CPHA, 1978; Finnegan, 1989; Hickie, personal communication). For example, AMI is subcategorized into 410.0 anterolateral wall, 410.2 inferolateral wall, 410.6, true posterior wall infarction, etc. The fifth digit is subdivided into '0' for episode of care, unspecified, '1' for initial episode of care, and '2' for subsequent care which includes observation for treatment or observation following an acute or healing state.
ICD-9CM contains a supplementary procedural classification code which consists of up to four-digits (01.00 to 99.96). When used, it makes the system a dual classification system.

ICD-9CM replaces both the ICDA-8 and H-ICDA-2 versions. The ICD-9CM exceeds the ICD-9 in the number of codes provided (Harrington, 1987). At the time of printing, the added detail, 9607 codes in ICD-9 to 10,241 codes in ICD-9CM, was achieved by adding fifth digits (Slee, 1978). Today, ICD-9CM has over 11,000 codes. The ICD-9CM provides an opportunity for individual hospitals to maintain their data in greater detail than is possible with ICD-9.

ICD-9 is a bound book, 6" x 9-1/2", which cannot accommodate change easily. Making changes in ICD-9 is cumbersome because changes need to be added or deleted manually. Since the current edition has not been revised since 1977, it contains many changes and additions. Because the pages of the volumes are small, it is not often easy to add entire sections with ease.

The typography and headings of both ICD-9 and the ICD-9CM are sometimes very difficult to work within. For example, essential information, such as headings, inclusion and exclusion notes, is noted on a previous page. There are no comments or references to guide the reader to turn to the previous page for these important details or notations. Hypertensive disease is such an example with the four-digit subdivisions and exclusions on one page, while the specific codes are on following pages (WHO, 1977, page 261). Another example in the CCP manual, shows the heading for the two-digit category for Operations on Retina, Choroid
ICD-9CM, on the other hand, has several choices of text. One is not bound, but is provided in a binder, approximately 8" x 10". Changes can be made easily by extracting or adding a page. There is also a choice of three individual volumes or the combined volumes in one bound text. Typography and headings are easier to understand because the inclusion and exclusion notes are made more visible. Mandatory fifth digits are enclosed by a flag to show the coder exactly when the codes need the extra digits. Colour coding is also used to denote the need for mandatory fifth digits. CM also handles errata by utilizing preprinted pages that can be pasted directly into the existing manual.

Concordance of ICD-9CM with ICD-9 is being maintained through an agreement that ICD-9CM is exactly collapsible back into ICD-9, to permit international exchange of data (Slee, 1978). Although the original intent was to keep the two classifications completely compatible in this manner, some ICD-9CM code categories have changed dramatically from those of ICD-9.

For example, ICD-9CM provides a new code for chronic obstructive asthma. The code 493.2 has been provided to specifically identify chronic obstructive pulmonary disease (COPD) patients in whom the major manifestation is asthma, without mention of bronchitis. ICD-9 provides codes for chronic obstructive lung disease as 496, coexisting with codes for bronchitis as 491.2 or asthma 493.- with subcategories, .0 for extrinsic, .1 for intrinsic or .9 for unspecified.
It is also noted that for ICD-9 or ICD-9CM, there are no codes for differentiating stages of disease, e.g. acute exacerbation of COPD. There are also incorrect assignments of codes. For example, Volume 2 claims deformity metacarpus is assigned to code 736.0; however, Volume 1 states 736.0 excludes the finger.

The Institute of Medicine's (IOM) third national study had concerns with the coding of ICD-9CM to five digits, given that the four-digit specificity coding had a greater number of clerical errors than the three-digit coding with H-ICDA and ICDA-8 (IOM, 1980).

1.5.2 Canadian Usage of ICD-9CM

As part of the agreement between CPHA and HMRI since 1987, Canadian hospitals have had the option of coding with ICD-9CM. Currently, both the ICD-9 and ICD-9CM classification systems for diagnoses are utilized by Canadian hospitals. HMRI has the capacity to incorporate both. The two classification systems, ICD-9 and ICD-9CM, are compatible, but hospitals coding with the ICD-9CM have access to more finely detailed clinical data for comparative and summary reporting (Harrington, 1987).

Although Canada uses the WHO's classification of diseases (ICD-9) it has developed its own nomenclature for surgical procedures, the Canadian Classification of Diagnostic, Therapeutic, and Surgical Procedures, 1986 (CCP). Formal development of the CCP was undertaken by Statistics Canada when it became apparent that the WHO's classification would not meet Canadian needs (Statistics Canada, 1986). Only the Canadian CCP for procedures is utilized in Canada. The achievement of accurate and consistent medical classification using the ICD and the CCP are constantly reviewed (Statistics Canada, 1985).
When summarizing the pooled data for the provincial Ministries/Departments of Health, HMRI collapses the ICD-9CM data into ICD-9. It uses equivalency tables to regroup the data into the ICD-9 nomenclature for homogeneity before sending the information to the appropriate agencies. ICD-9 is the common base of the HMRI system (Hickie, personal communication) according to the international agreement. Where a Canadian hospital utilizes ICD-9CM, it is only that hospital, and other hospitals utilizing the inter-hospital comparisons, which can utilize these codes.

1.5.3 Limitations of the ICD

There are several known limitations of the ICD system of classification, some of which are:

1.5.3.1 ICD Design

ICD was designed for vital statistics on morbidity and mortality. It was based on cause of diseases rather than manifestations (Slee, 1978). Feinstein (1988) states that the coding conventions depend on the final diagnosis, not on the manifestations that may have precipitated admission. Although there are some categories with dual classification, the decision to use these codes is optional. In Canada, the definition of Most Responsible Diagnosis (MRD) is that diagnosis that determines the greatest proportion of the length of stay.

The task of achieving international agreement on the names of diseases appears impossible. There is a great deal of discussion about the suitability of a single coding system
to satisfy the needs of all users of health-related information, such as reimbursers, researchers, vital statisticians, health care providers, and manufacturers. Users' needs range from classifying causes of death for vital statistics to indexing of hospital records for retrieval purposes; from the federal and provincial level to academia; from large health care institutions to the individual practitioner. The uses vary from scientific to administrative to financial. It should also be reasonable to assume that a single coding system cannot provide all of the necessary detail and that other methods of collecting information are required.

There is a rapidly increasing demand for more flexibility within the system. There are attempts to change the lack of balance in categories with the tenth revision. Working with a classification system updated every ten years or so is inadequate, considering the rapid changes and new advances in medical diseases and procedures. Codes are not specific enough. However, more frequent updates, if consistency and ability to map precisely from the most recent to previous systems is not maintained, would play havoc with any longitudinal or time-series research.

Some clinicians feel that the non-medical factors, such as age, discharge disposition and insurance status, although important are better collected independently of the coding system (Mullin, 1988). But if we lose the ability to link across data sets because of independent collection, then the costs may outweigh any possible benefits of independent collection.
1.5.3.2 Mono-axial Orientation

Another problem with the ICD taxonomy, according to Feinstein (1988), is that the mono-axial orientation of the taxonomy offers a single name for each disease, but prognosis and therapy in modern medicine depend on multi-axial consideration of such features as decompensation of organ systems, severity of illness, rapidity of disease progression, and existence of major ailments. The diagnosis may identify the patient's ailment, but does not identify how sick the patient is, what is going to happen prognostically, and what to do therapeutically. All these other features of patient care require additional categorizations of the patient's clinical status. A classification system should be comprehensive, easily expanded and capable of constant updating. There should also be some method to capture severity of illness or disease staging. HMRI (CIHI) captures severity through the RIW system, outside of the main disease classification system. RIWs are explained below in section 1.7.6.

The limitation of a single axis coding system is that only a limited number of concepts are possible. The single axis code already incorporates a number of expressions, and yet does not express the needs of medicine as a whole. Minor variations in wording are gathered into a code which puts unlike conditions together. Special purpose codes are quite useful in specific applications, but do not communicate with medicine as a whole. Coding of an extensive medical condition into a single code number loses data and detail. When the procedural classification is utilized, as in most American and Canadian health institutions, the classification system becomes a dual axis system.
Feinstein (1988) notes that two patients diagnosed as having coronary artery disease may be substantially different in prognosis and therapy because the first patient, having been detected during a screening procedure, is asymptomatic, whereas the second patient has severe angina pectoris, unstable atrial fibrillation, and moderate congestive heart failure. These distinctions among patients with the 'same' disease are not reflected in monoaxial diagnoses that cite etiologic causes while omitting pathophysiologic effects. At the ICD-9 conference, the World Health Organization (WHO, 1977) felt that the multi-axial classification often destroys the ability to retrieve disease terms, and therefore maintained its mono-axial orientation.

1.5.3.3 Rules and Guidelines

There is apparently no written documentation of rules, guidelines and interpretations and there are multiple sources supplying conflicting information (Mullin, 1988). Currently, hospitals in B.C. need to consult individual hospital policies, the ICD guidelines, Statistics Canada, HMRI guidelines and the provincial Ministry of Health (MOH) requirements. There are also exceptions to many rules (Hickie, personal communication). When guidelines are not clear, or when there is no code available in the code books, hospitals may contact any one of the above sources for information. Where there is documentation, it is usually scant, and not in a form of one policy or procedure manual, but several. When this information is received, information to other hospitals is not disseminated until the Statistics Canada errata addendums are published. This time frame could sometimes be years later. For example, information on new disease entities, such as Human Immunodeficiency Virus Infection did
not reach the coders until several years after the disease was discovered. However, there
were individual hospitals using some of these codes long before others because they
specifically asked how to code it. The errata from Statistics Canada do not necessarily reflect
the information hospitals receive from the other two sources, HMRI and the provincial
Ministries/Departments of Health. At times, the MOH in B.C. has requested that hospitals
code a certain way. There is potential for conflict of information.

Mullin (1988) claims the WHO council appeared more concerned with the placement
of diagnostic information, such as whether strokes should be in the nervous or cardiovascular
section, rather than with the actual defining of disease entities and the establishment of rules,
guidelines and interpretations. He also feels that the coding rules for acute myocardial
infarction (AMI) in ICD-9 do not accurately reflect the incidence of AMI. For instance, all
AMIs "specified as acute or with a stated duration of eight weeks or less" are to be coded as
410, acute myocardial infarction (WHO, 1977). Thus, any admission to hospital within eight
weeks of the original AMI is given a code denoting an acute infarct. It is not uncommon for
a patient to have several admissions within an eight-week period of the original acute infarct.
Each of these admissions will be statistically counted as a separate AMI, which could
potentially bias the national database.

In addition, many suspected but eventually disproved entities, such as "rule out
myocardial infarction" may be coded as though the entity had actually occurred. Although
the Canadian convention utilizes the definition of MRD, if a patient is admitted to "rule out"
any suspected disease entity, that entity is coded if no other diagnosis is presented. In the
case where the diagnosis is ruled out, the convention is to code the symptoms presented during admission.

The policy to code "rule out", "suspected", or "question" infarcts as actual AMIs has made it impossible to determine the real incidence of AMI. This would be the case with any disease entity. ICD-9CM has attempted to correct this by utilizing fourth and fifth digits, as previously described in section 1.5.1. ICD-9CM provides a way for the coder to distinguish between admissions for an actual AMI and admissions for care and treatment following the initial episode. This is not the case with ICD-9.

1.5.3.4 Criteria for Classification Systems

Feinstein (1988) also claims that the ICD has never attained the third requirement of a scientific system of classification. It offers no operational criteria to provide reproducible identification for each disease. Without operational criteria, clinicians can use the standard names of diseases authorized in the ICD, but the names are chosen with enormous variability by different users. Feinstein, (1988) cites an example by which the same clinical ailment was found diagnosed in the United States as *pulmonary emphysema* and in the United Kingdom as *chronic bronchitis*. More recent examples occur with the combination of diseases, whereby one physician may prioritize the diseases differently than another, thereby presenting a different MRD. Another occurs where two specialists may each give a separate disease entity

---

for the same conditions of one patient. This problem occurs more in the diagnosing of symptoms rather than the classification system itself.

1.5.3.5 Incomplete Categories

Feinstein (1988) claims that apart from the inconsistencies in the use of the nomenclature, technological progress has also led to new disease entities and new diagnostic methods of identifying them. He refers not only to the addition of new entities, but to the transformation of disease entities into altered nosology terms. For example, because of new technology, the label of "tuberculosis" would no longer be given today unless accompanied by evidence of pulmonary lesions and microbial demonstration of the tubercle bacillus. Therefore, the incidence of the disease is not consistently traced through the ICD system.

1.5.3.6 Imprecision with Procedural Categories

Feinstein (1988) states that many classifications of surgical procedures are imprecise. Surgical procedures may be coded according to the outcome, rather than the approach used. Mullin (1988) goes further to say that physicians play a very minor role in the development of the procedural categories. He also claims that physician input into the entire process is mostly through various specialty societies.

Some procedure codes do not have clear guidelines for their usage. For example, does one use the diagnostic code, V64.3 (procedure not carried out) or procedure code 99.07 for an operation abandoned before onset. HMRI guidelines specify that V64.3 be used when a procedure is cancelled. Some explanations the investigator heard from practising health
record personnel during this study ranged from "it depends on whether they were set up in
the operating room, waiting for this patient" to "it doesn't make any difference". This type
of coding is left up to the individual hospital to devise a policy for scheduled surgery
cancelled for any number of reasons.

1.5.3.7 Accuracy and Consistency

Accuracy and consistency in the selection of diagnostic and procedural codes are a
concern at every level in the health care system where coded information is utilized.
Statistics Canada produces national statistics based on coded information received from a
variety of sources, one being health care facilities. The problems of inaccuracy and
inconsistency have a potential magnifying effect due to the volume of data being processed.
Errors in these national statistics could have broad effects on both short and long range health
services planning.

1.5.3.8 Subjectivity

Every disease or abnormality should have a definite and appropriate placement in one
category within the classification. Most do. That is, direct coding is possible to most of
these classes. However, as noted earlier, there are over 120,000 diagnostic terms in common
use (AMRA, 1987a; 1987b). Therefore, decisions must be made about where each diagnosis
belongs in the classification. For example, there is no term in Volume 2 for unstable angina.
Some coders use 411, others use 413. HMRI guidelines state that 411 be used.
The practice of classifying is the placing of one diagnosis in a class or group of diagnoses related to each other in some manner. Subjectivity of the coder involves making a judgement regarding some classifications. Coding subjectivity is poorly recognized and inadequately addressed (Lloyd & Rissing, 1985). Several studies have demonstrated judgement errors as the main cause of discrepancy (IOM, 1977a; 1977b; Demlo, Campbell & Brown, 1978; NJDH, 1978; Barnard & Esmond, 1981; Corn, 1981; Hsia et al., 1988).

Whatever classification scheme is used, it should not lose sight of the fact that more than one classification is needed for some diseases (Reinke, 1988). No single classification will satisfy everyone’s needs and no degree of sophistication will produce an ideal, all-purpose system of classification. Even modest sophistication may be unwarranted when the reported diagnoses are of questionable accuracy in the first place.
1.6 ICD-10

The ICD-10, the *International Statistical Classification of Diseases & Related Health Problems, Tenth Revision* is the latest in the series of ICD classifications formulated in 1893 as the Bertillon classification. The title was amended to make the content and purpose clearer, but the more familiar ICD acronym was retained (WHO, 1993). It consists of three volumes: Volume 1, the Tabular List, Volume 2, the Instruction Manual, and Volume 3, the Alphabetical Index.

The Instruction Manual is new and provides a basic description of the ICD. It provides guidelines for recording and coding, although it is not intended to provide detailed training. It was also developed for ease of handling when reference to both Volumes 1 and 3 need to be made simultaneously. Currently, the instruction manual is part of the introduction to Volume 1. When instructions need to be reviewed, the coder may find him or herself flipping back and forth in the volume. With ICD-10, the instruction book can be accessed without interruption to the coding manuals. Another advantage is that the instruction book can be changed without changing the entire first volume.

This revision of the ICD does attempt to address some concerns with the previous editions. Its newer alphanumeric coding scheme replaces the previous numeric one and consists of one letter of the alphabet followed by three numbers at the four character level. For example, diabetes mellitus (DM) is coded under the nutritional and metabolic section designated by the letter ‘E’. DM is given a three-digit code between E10 and E14. Within this version, the major subdivisions are: E10 - insulin dependent DM, E11 - Non-insulin dependent DM, E12 - malnutrition-related DM, E13 - other specified DM, and E14
unspecified DM. Ketoacidosis is coded as the fourth character (.1). The fourth-digit character of .7 is used for multiple complications. Diabetes mellitus with ketoacidosis, juvenile type onset is assigned a five-digit code, E10.1.

ICD-10 provides a larger coding frame than ICD-9 and leaves room for future expansion providing the flexibility the previous version did not maintain. Three-digit categories increased from 1200 in ICD-9CM to 2000 in ICD-10. The dagger and asterisk system of dual classification to classify according to manifestations, although highly criticised, was retained and extended but again, was not mandated. The use of these codes is still optional. The new version consists of 21 chapters, as opposed to the ninth revision's 17 chapters and two supplemental classifications. The supplemental classifications are now included in the core classification. Two new chapters were created by splitting two of the previous chapters.

Proposals containing outlines for biaxial classification for the ICD-10 were turned down by the WHO (Côté & Rothwell, 1989). These proposals recommended letters to be used for the division of anatomy and the disease or abnormality. For instance, the first axis was for anatomical or topographical site. The second was for disease or abnormality. The code for anatomical site made up the first half of the code, the disease entity made up the second. An example of a code may look like F11 + B10, where F would be the cardiovascular system, 10 the myocardium NOS. The letter B represents a congenital anomaly and 1- a congenital absence. The ICD-10 has maintained the monoaxial orientation.

The new revision makes it easier to see the inclusion, exclusion and other notations. Chapter headings are larger print and in bold type. The alphanumeric category blocks are
shaded and easily visible to the reader. Inclusions, exclusions and notes follow the chapters, three-digit blocks and category titles. When fourth-digit subdivisions are provided under the chapter heading, the three-digit block has a note to refer to the chapter heading for a listing of the respective notes.

The core classification is a three-character code. The fourth character follows the decimal point. The mandatory level of coding is to the four-character level. Where the fourth character is not used, the letter X is recommended to fill its position so that codes are the standard length for data processing. ICD-9 left this position blank. In the HMRI conversion from ICD-9CM to ICD-9, codes that should have been left blank were given a "0". For instance, ICD-9CM provides a fourth digit for 410, such as 410.10, AMI of anterolateral wall. ICD-9 only provides a code for 410 with no fourth digit. With the conversion ICD-9 codes show 410 as 410.0, although site is not specified.

Work on the ICD-10 started as early as September of 1983 and was originally scheduled for release on 1 January 1993 (Slee, 1978; Mullin, 1988; Côté & Rothwell, 1989). In the recent past, WHO-sponsored revisions were scheduled on a ten-year cycle. However, the ninth revision was on a modified schedule and was expected to be in force for approximately 15 years. Although the ICD-10 was published by 1993, member nations are at liberty to implement later, based on their own needs and circumstances (Taylor, 1992). Both the American and the Canadian implementation will occur later than the modified 15-year schedule. Both countries are adapting the system.

The National Health Information Council (NHIC) is a federal-provincial council that provides direction and sets priorities for Canadian health information issues. NHIC has been
established as the umbrella organization for implementation of the ICD-10. As noted previously, the NHIC was also the organization that created the task force that resulted in the formation of the CIHI. NHIC has several concerns with early implementation of the ICD-10 (Taylor, 1991). The concerns are:

1. the use of ICD for the collection of both causes of death and morbidity information and its use in various disease registries,
2. modifications of the ICD currently in use and difficulties surrounding the use of more than one version within the country,
3. the need for an updated procedure classification,
4. modifications needed to computer systems,
5. the amount of lead time necessary before any change is introduced, and
6. the need for a coordinated approach to training and implementation.

With these time lags, the Canadian implementation will be delayed until approximately the end of 1995 or 1996 (Taylor, 1992). The American procedural classification has not yet been determined, which is causing delay with the implementation of ICD-10. Implementation is not expected in the U.S. until 1996 (Taylor, 1992; Skurka, 1994) and there are some predictions that it will be as late as the year 2000 (Skurka, 1994).
1.7 COMPONENTS OF THE HMRI (CIHI) SYSTEM

This section reviews the Most Responsible Diagnosis (MRD), comparison of the MRD and the principal diagnosis, and the advantages and the disadvantages of the system. A short review of Case Mix Groupings (CMGs) and Resource Intensity Weights (RIWs) completes this section. HMRI records are computerized; however, hospitals previously submitted data by paper abstracts. For a prototype of an HMRI paper abstract, see Appendix C.

This database is organized by variables, such as the Most Responsible Diagnosis (MRD) and the Principal Procedure (PP). Complications, i.e. those conditions that develop subsequently to admission, and comorbid conditions, i.e. those that coexist at the time of admission, are coded as diagnoses and accessed through the diagnostic reports. The HMRI reports represent an overview of patterns of patient care and illustrate the utilization of resources associated with each in-patient admission.

1.7.1 Most Responsible Diagnosis

In Canada, the MRD is not the first diagnosis or the diagnosis that occasioned the patient's admission to hospital, but that responsible for the greatest length of stay. It is the diagnosis that describes the most significant condition of a patient which 'causes' the longest portion of the length of stay in hospital (Winger & Wu, 1989; HMRI, 1990).

In a case where multiple diagnoses may be classified as the MRD, HMRI policy states that the diagnosis 'responsible for the greatest length of stay' is coded. The MRD may not necessarily be associated with the diagnosis that made hospitalization necessary in the first
place. However, it is not always easy to determine the MRD, given the inter-relationship of various conditions.

The MRD chosen by an individual coder has implications all the way up to the Canadian national database. All statistics published by Statistics Canada, by the provincial Ministries/Departments of health, by the Teaching Hospital Associations, by Community Hospitals, and monthly health record reports within each participating hospital all begin with the MRD.

1.7.2 Comparison of the MRD and the Principal Diagnosis

It is worth noting here that the Canadian definition for the MRD differs from the American definition of principal diagnosis. The American definition (by CPHA) is that which is "the condition established, after study, to be chiefly responsible for occasioning the admission of the patient to the hospital for care" (Winger & Wu, 1989).

Although there are similarities, comparison of the two variables must be done with caution. The most significant difference between the MRD and the principal diagnosis is the fact that the latter captures the "reason for admission". During treatment a patient's condition can change a great deal. The MRD is designed to capture the "reason" for the patient being in hospital the longest. For many hospital admissions, the reason for the patient being in hospital may be quite different from the diagnosis on admission.
1.7.3 Positive Features

There are unequivocal advantages to a Canadian database. The database represents the federally-provincially funded health care system. It has the capability of relating directly to Canadian medical and health care practices. It can facilitate inter-hospital comparisons so that hospitals can more effectively monitor use of hospital resources. It can standardize the reference database for hospitals and government and support the creation of profiles of unique categories of Canadian institutions, such as specialty hospitals. It can enhance the capability to monitor trends and incidence of disease throughout the country. But more important, particularly in the context of this study, such a database has the capacity to standardize definitions and documentation that can provide uniformity and consistency from coast to coast.

The HMRI system is relatively inexpensive and easily understood by health record personnel. It has the capacity to summarize health record information for the clinician, the medical researcher, and the administrator. It also offers the capacity to draw comparisons across hospitals or specialities and it makes possible graphic formatting of information.

In an attempt to ensure coding is as accurate as possible, HMRI has many edits in place. Edits are computer-generated checks to ensure the accuracy of documented information or to determine missing information. In both classification systems, ICD-9 and ICD-9CM, age and sex edits are in place. For example, an error would be generated if a hospital submitted an abstract with the sex identified as male and a procedure code of total hysterectomy. There are also certain codes that should not be used as the MRD, such as history of neoplasm codes (V codes) and causes of injury and poisoning (E codes) with both
the ICD-9 and ICD-9CM systems. The edits ensure these codes do not appear in the MRD position. A neoplasm code, not the history of neoplasm code, and the actual injury itself, rather than the cause of injury warrants the MRD position.

1.7.4 Problem Areas

As with many systems, just as there are favourable features, there are also several problem areas. A major problem is that the data are assumed to be complete and accurate. However, there is a lag in the submission of data from some hospitals, rendering the pooled data incomplete. Only error-free abstracts determined by the computer edits are submitted to the Ministries/Departments of Health (Buchanan, 1983; Winger & Wu, 1988). No extensions are given. Even if the hospitals do not submit quarterly data within the 90-day period, the data are pooled by HMRI for submission to the provincial Ministries/Departments of Health. The period for completion allows 60 days at the end of each quarter for submission of original abstracts and an additional 30 days to submit and process corrections. Revisions or tardy submissions after this date are corrected for the individual hospital data, but not for the Ministries/Departments of Health. This process can easily lead to an incomplete database. It is noted that pathology reports sometimes appear two to three months after the patient’s death. The original record is coded long before this event and sometimes the MRD on the pathology report is not the same MRD as stated on the record at the time of death. Even when coders change the data for their own hospital’s database, the national database is left unchanged.
All hospitals are required to update the data requested in the default report. The default report stems from the edit check. HMRI requests hospitals to make changes to or additions to the submitted information. The purpose is to ensure that the Master File has been updated for any reporting to the appropriate Ministry/Department of Health (Winger & Wu, 1988) as well as for individual hospital records. However, not all the corrections are completed within the 30-day period and not all hospitals have all the data to HMRI at the end of the 60-day period. If a hospital does not submit its data to HMRI on time, these same corrections will not form part of the national database.

There can also be deficiencies in submitted hospital information which render it incomplete and not a reflection of the hospital’s actual experience. Within the HMRI system, only in-patient recording of information is mandated: daycare and ambulatory coding is optional. Although daycare recording is mandatory in BC, ambulatory (out-patient) and intermediate care coding is not. Intermediate and extended hospitals reporting is only mandated if they are associated with an acute care facility. Free standing facilities reporting is not mandated (Hickie, personal communication). At the time of writing, very few hospitals in B.C. currently utilize the ambulatory option (Hickie, personal communication). Although ambulatory patients are not recorded, they do represent a substantial workload within most health care facilities.

Another problem is the lack of information about inter- and intra-rater reliability of coders across and within the provinces, which raises doubt concerning the consistency of the way in which the information is coded. HMRI requires no testing of coders and has no requirements for the training of coders, although it assumes that trained health record
personnel do actually code the charts. However, under the current system, individuals need only be employed by a hospital. The training of coders was reviewed in section 1.1.2.

Although HMRI has edits in place in an attempt to ensure accurate codes, there are no audits performed to review the accuracy of the coding and abstracting process. No check is done to ensure that what is coded is actually representative of the information in the health record. To take it one step further, no check is done to ensure that the documentation in the health record actually represents the events and illnesses of the patient.

It has been mentioned previously that Canadian health care institutions have the option to utilize either ICD-9CM or ICD-9. HMRI converts the ICD-9CM data to ICD-9 to provide a common denominator for the pooled database. In preparing the data for the appropriate Ministries/Departments of Health, essential elements of codes are deleted. For example, Q prefixes are added to the code when the diagnosis is questionable. In the transformation, the Q prefix is deleted, thus making the diagnosis appear more definitive than it is. Fifth-digit diagnostic codes that are not mandatory are removed. All fifth-digit procedural codes are removed. Two examples of fifth digit procedural codes are an instance where a patient had a procedure done, but in another hospital (i.e. 0 added to the end of the code) or where the operation was not completed (i.e. an 8 added). Removal of these digits provides great potential for the mis-interpretation of the remaining procedure codes.

Although encoding systems are currently being developed, there is no encoding system currently in the province of B.C. An encoding system is the computer-assisted conversion of data from one type (i.e., written diagnoses and procedures) to another (i.e., alphanumeric codes). Vendors are currently marketing computer-assisted encoders that enable a coder to
determine the numeric or alphanumeric code based on its written diagnosis or procedure. Branching logic is used to lead the coder to the desired code (Skurka, 1994). Software options include coding, grouping, editing and optimizing. Consistency and accuracy are major advantages with such a system. Coding errors, such as the transposition of numbers and judgement errors to determine the MRD may be avoided with an encoding system.

1.7.5 Case Mix Groupings (CMG's)

ICD-9 codes are too detailed to allow comparison and grouping of homogeneous patients. It is assumed that patients could be classified into clinically cohesive groups, similar in resource consumption, which would then be employed to describe a case mix for each hospital. This would form a basis for logical decision-making on resource allocation across institutions (Linton, 1984). Therefore, HMRI first introduced the Case Mix Groupings (CMGs) in 1983 for better homogeneity of patient groups and for organizing discharge information on patients (Fiskel, 1992). CMGs are a way to aggregate the data so that the hospital inpatient activity is defined in a meaningful way. The CMG's were adapted from the American coding system of Diagnostic Related Groups (DRGs), originally developed by a group at Yale University and modified to the Canadian ICD-9 coding system. There are 553 CMG groups in the 1990 version (HMRI, 1990).

Diagnosis "typing" is important in CMG assignment. Besides the MRD, diagnoses may be recorded in several ways, but only three are described here: primary, or comorbid (type 1), complications (type 2), and secondary (type 3). Primary diagnoses are diagnoses

\[15\]CMG is a registered trademark of HMRI
other than the MRD that are also important conditions of the patient and have a significant influence on the patient’s length of stay. A complication describes a condition arising after the beginning of treatment. A secondary diagnosis describes a condition for which a patient may (or may not) have received treatment, but which did not significantly contribute to the patient’s length of stay in hospital. This database splits the major clinical CMG categories between those requiring surgery and medical cases. It is further delineated by age of patient. Additional primary diagnoses and complications are used only when clinically relevant. Length of stay plays no role in determining the assignment of a CMG.

1.7.6 Resource Intensity Weights (RIWs)\textsuperscript{16}

The advancement of the CMG has necessitated the development of Resource Intensity Weights (RIWs). RIWs were developed by HMRI to measure the relative expected difference in the cost of providing care to patients in different CMGs. The RIW for an individual CMG represents the ratio of the average total cost (from admission to discharge) for a patient within that CMG, to the average total cost of a hypothetical "standardized" patient. The lower the RIW for a CMG, the lower the expected total cost of care (on average) for the patients within the CMG. RIWs have been used to support utilization management, physician impact analysis, planning and budgeting. Some Canadian provinces are also moving towards incorporating the RIW in their funding of in-patient care.

The original RIWs were created by taking New York State "Service Intensity Weights" and New York cost information for DRGs, mapping the DRGs to CMGs, and

\textsuperscript{16}RIW is a registered trademark of HMRI
adjusting the weights to reflect length of stay differences for patients in New York versus Canada (HMRI, 1990). New York data were used because of the unavailability of Canadian CMG-specific cost information. Because there is not a one-to-one comparison between CMG's and DRGs, the HMRI team has considered issues, such as the differences in coding between Canada and the United States. It has compared the differences in populations, medical practices, health care system policies, and diffusion of technology, among other factors, between Canada and New York. New York costs were adjusted to reflect the apparent differences between the two populations and hence, the Canadian RIW was born. A review of the definition of the "typical" case within each CMG and an approach to determination of relative resource use for non-typical cases was also considered. HMRI is working to ensure that CMGs reflect the state of the art in case mix technology as applied to the Canadian health care system (HMRI, 1990).
1.8 OTHER CRITICAL ISSUES

Just how accurate are the data in these databases? Do we really know that the information in these databases actually reflects the hospitalization experience of the nation? Examination of databases is essential to ensure reliability of the information they claim to provide.

The question addressed by this study is whether the Canadian HMRI database is reliable. It examines its reliability specifically for six Vancouver hospitals within the British Columbia provincial database. This section reviews the critical issues addressed by the study and not previously examined.

1.8.1 What is Reliability?

The quality of the coded data within the databases is a constant topic of discussion among health care professions. The elements of data quality are validity, comprehensiveness, and timeliness of the available data (McLean, 1986).

The three different goals of scientific measurement are consistency (otherwise known as reliability), accuracy, and suitability (Feinstein, 1987). These three goals encompass the validation process.

The validity of a measure refers to the adequacy with which the method of measurement does its job: that is, how well does it measure that which the investigator actually wants to measure? Reliability refers to the degree to which an instrument can be consistently used over time to achieve the same results (Feinstein, 1987). It is a measure of
the comparison of two or more independent measurements. Therefore, if a measure is not reliable, the validity is reduced.

Abramson (1990) states that high reliability does not necessarily mean that a measurement is a satisfactory one. However, he does state that the less reliable an instrument, the less useful it will be. Ideally, the agreement rate for inter-rater reliability should be 100%. However, he states that complete reliability is not essential. Roos et al. (1982) argue that it cannot even be achieved. What is essential is to know how much unreliability there is and what the intended use of the information is. Feinstein (1987) concurs that the information is reliable if the measurement obtained with a particular procedure ‘closely agrees’ with the results obtained with a reference or standard procedure. Reliability for clinical use for specific patients will require a higher standard of accuracy than the reliability for purposes of population-based utilization analysis.

1.8.2 Sources of Reliability

Feinstein (1987) suggests three sources of reliability: input, instrument and observer (or rater):

1.8.2.1 Input Reliability

Input variability arises when the source of information itself is altered between observations. For example, events documented within the health record should not change after discharge or death. In reality, however, record documentation is often augmented because of delayed reports. But the record itself should not be altered. It is common
knowledge that very few records are completed by the time of discharge or death. Since hospital personnel code from health records, coding guidelines require that the chart be completed before the coding process takes place (Winger & Wu, 1988b). In practice, however, this is not always the case. For several reasons, such as administrative policy, budget restraints, or even meeting HMRI deadlines for completion, health record personnel do code from incomplete charts.

Other areas of input variability include situations where documentation has arrived after the chart has been coded. For instance, a pathology report could change the discharge diagnosis of a patient. If it is just filed onto an already coded record, without any changes to the database, a discrepancy between the record and the database will be created.

1.8.2.2 Instrument Reliability

Instrument reliability varies when the methods or procedures used in the collection of the data have their own source of variation. The classification system itself often leads to discrepancies due to the subjective nature of the coding process. For example, the standards and procedures used within the classification system may not be clear in every instance, and can be left to subjective interpretation by the individual coder.

Common medical terms may be coded differently depending on the policies and procedures of each hospital. There may not be a one-to-one correspondence between the medical terms in the chart and the medical terms in the classification system. For example, there is no ICD-9 classification for recurrent tonsillitis: the coder must choose between acute and chronic. This is a common medical term for children admitted for a tonsillectomy. Each
hospital may address the coding dilemma differently. Another example: peripheral edema is not found in the Alphabetic Index, but cardiovascular edema is found under congestive heart failure, code 428.0 (ICD-9). However, the term, peripheral edema, is not found under the category 428.0 in the Tabular List. Category 428.1 is left heart failure with acute pulmonary edema. There is no one-to-one correspondence between the term, peripheral edema, and any specific ICD-9 category.

Discrepancies can also be introduced because of differences in hospital policy. Variations in hospital procedures can create different sets of instructions among the various hospitals. For example, the use of dual classification coding in ICD-9 is optional. Each hospital needs to address whether it will choose to code the manifestations of the disease entity. When hospitals follow different guidelines, variation in the database results. This can lead to an incongruence within the aggregate database.

1.8.2.3 Observer (Rater) Reliability

An index has observer, or rater, reliability if it yields similar results when repeatedly applied to the same entity by different observers. People who apply the procedures may have their own variability when they use them.

1.8.2.3.1 Inter-Rater Reliability

Inter-rater reliability is the variation among individuals making the same observation, not from changes in the characteristics or from the measuring instrument itself. For example, coders who extract data from records may have their own interpretation of the documentation.
within the chart. This variation in interpretation may be due to a diversity in coding skills, differences in perceptions, or a propensity to make mistakes.

Inter-rater variability is of greater concern than intra-rater. Inter-rater reliability is a source of variability that could seriously affect the outcome of any study and should be measured in studies where more than one observer is used to collect data (Beard et al., 1988). It is conceivable that misclassification could invalidate a study. Beard et al. (1988) suggest that low inter-observer reliability might not necessarily indicate poor agreement between the observers, but may indicate observations or variables that are difficult to classify. However, while low inter-observer reliability may be because of difficulty in classification, it must by definition indicate poor agreement between the observers.

1.8.2.3.2 Intra-Rater Reliability

Intra-rater reliability is the variation among observations of the same events by the same person on different occasions.

Both intra- and inter-rater reliability can and do introduce misclassification error into data collected for research purposes and both are of concern during data collection activities. Both are a potential source of variability.

1.8.3 The Coding Process

The health record itself is the foundation of coding accuracy. The major responsibility for accurate hospital coding resides in the health record department. Extracting
the data from the health record for this database traditionally has been the responsibility of health record coders. Health record coders perform an extremely important task. This section reviews the process, the three dimensions of coding accuracy, classification rules and how to classify a term not listed in the ICD.

Diagnostic and procedural coding systems are the basic building blocks for the development of almost any type of patient classification system. Diagnostic and procedural data are compiled by coding and classifying the diagnostic and procedural events recorded in each patient's health record for each episode of care. Problems in the coding system are imported directly into the database itself (Mullin, 1988). Extracting data from unstructured or incomplete records is time consuming and prone to error.

The face sheet and the discharge summary are the two documents most widely used within the health record. A summary, or face sheet, is the first sheet of the health record and summarizes such identification variables as the admission date, the separation date, the calculated length of stay, the MRD, the PP, secondary diagnoses, secondary procedures and other pertinent non-diagnostic information. This information is documented by the physician, the nursing staff and the health record personnel. Sometimes, the face sheet contains the physician's summary, in paragraph form, stating the outcome of the patient's stay.

A discharge summary, on the other hand, is a concise document prepared by the physician that summarizes the patient's illness, investigation and treatment in hospital (Huffman, 1977). It should also include a note regarding the condition on discharge and follow-up plan for the patient. According to the Canadian Council on Health Facilities Accreditation, documentation should be completed within 48 hours of discharge.
Converting a record into codes in computer readable form adds another step in which error may be introduced (Hendrickson & Myers, 1973). After the data are coded and abstracted, they are submitted to CIHI for processing. Currently, most hospitals use electronic processing as opposed to the previously-used paper abstracts. This speeds up the process for the hospitals and for CIHI.

1.8.3.1 *Three-dimensions of Coding Accuracy*

Thompson & Slee (1978, p. 44) define three dimensions of coding accuracy:

1. the accuracy of individual codes
2. the accuracy of the totality of codes
3. the accuracy of the sequence in which the codes are recorded.

For example, for one specific patient, each individual code may be correct, and even the totality of codes may be correct, but the sequence in which the codes are recorded may not be. This could lead to a wrong conclusion being drawn about that patient and his episode of care.

Thompson & Slee (1978, p. 48) maintain that the definition of coding accuracy be that the code "must describe the patient's condition as completely as possible *within the constraints of the classification used*". Thus, the criterion for determining whether a diagnosis is complete "*cannot be less than* the degree of specificity that is possible within the classification".

It is the total content of the health record that represents all that is known about the patient. The entire content of the record must be reviewed to determine whether there is information present which allows for a more specific code than that which would be obtained
if only the face-sheet diagnosis were used. A coder must thoroughly read and interpret the
document to select the appropriate MRD and PP. Assessing the adequacy of the
documentation is necessary to ensure that the documentation supports the assigned codes.
This process is time-consuming and prone to error (Porter & Tibbott, 1986).

Although physicians document diagnoses and procedures on the face sheet, they may
not necessarily reflect the word-for-word terminology in the alphabetic index of the ICD or
the correct selection of the MRD. Many coders express reticence to abstracting information
that does not appear on the face sheet or in the discharge summary (New Jersey Department
of Health (NJDH), 1978). Coders are also reluctant to omit any diagnosis or part of a
diagnosis reported by a physician.

Often, physicians are not aware of the coding system or the rules, guidelines and
interpretations applied to the system and will usually not consider these requirements when
completing the face sheet (Thompson & Slee, 1978; Doremus & Mickenzi, 1983; Lloyd &
Rissing, 1985; Feinstein, 1988; Mullen, 1988). Wood et al., (1989) suggest that
responsibility for coding has not been accepted generally by clinicians. The authors have
given several reasons for this: some physicians may not be happy with the system adopted in
the ICD, others may argue that their busy clinical schedule does not afford them time, or that
it is impractical to carry two volumes of the ICD around the health care facility and the task
of locating the directory would be both laborious and time-consuming. The database may be
a threat to professional autonomy. Physicians need to be educated about and involved in
activities that support the hospital database.
The requirement for accurate sequencing of codes is not directly related to coding itself, but rather to the way in which coded data are abstracted from the health record and transferred into the data system. When the first listed diagnosis on the face sheet was routinely abstracted as the principal, discrepancies in sequencing were likely to occur, since sequencing is based on total record review (NJDH, 1978). What to code rather than the selection of codes is a difficult process and accuracy usually depends on the skill of the coder. The circulatory system, especially the cardiovascular and cerebrovascular, are the most difficult areas to code (Lloyd & Rissing, 1985).

All diagnoses affecting the current stay and all procedures performed must be assigned a code. Each individual diagnosis or procedure must be assigned a correct and complete code or codes. This requirement includes all conditions that coexist at the time of admission (comorbid) or develop subsequently (complications) that effect the treatment received or the length of stay. Codes must be sequenced correctly to understand the chronology of events.

Thompson & Slee (1978) argue that the primary sources of errors are clerical (carelessness) and procedural (incorrectly following coding procedures). For example, experienced coders may assign a code from memory that may result in consistently assigning an incorrect code number (Huffman, 1985).

1.8.3.2 Classification Rules

Classification rules need to be followed. The Nosology Reference Centre within Statistics Canada (April 1985) stresses that coders need to utilize the classification manuals correctly. Coders need to understand and apply the conventions used in the classification
systems and should always refer to the Alphabetic Index of the classification first and then verify the code number in the Tabular Listing. The instructional notes in both the Tabular Lists and the Index are necessary before making a final decision on a code number.

In abstracting and collecting coded diagnostic and procedural information from hospital records, one must consider many rules and guidelines when deciding how much or how little information to code. One must act within the structure and rules of classification being used (e.g. ICD-9 or ICD-9CM) and within the definitions and guidelines of the data processing system involved (e.g. HMRI/CIHI) in such a way as to ensure retrieval of consistent, accurate and useful data (Taylor, 1984). In addition where coded information is being collected for an agency (i.e. the provincial Ministry of Health) any rules or guidelines established by that agency must also be followed.

Although HMRI/CIHI has strict coding guidelines, there are no standards on coding procedures. This implies that each individual hospital may extract the same information differently to suit the needs of the individual participants. For instance, historical diagnoses are entirely optional to code. Some hospitals abstract these as current disease, others do not.

The term, "overcoding", is used among health record personnel. Coding all the events in a record would be considered overcoding, but there are no objective criteria for defining this term. It is defined differently by each individual. The results are that the health record coder interprets the details in the record to determine whether the code is of relative importance or not. There is a fine line between overcoding and not coding all the relevant aspects of the patient’s diagnosis.
For a variety of reasons, medical staff or research committees of individual hospitals may request diagnostic information that is slightly different from the system's standards. To accomplish this, they may have to go against HMRI/CIHI guidelines or the rules set out by Statistics Canada. For example, one of the hospitals in this study requested that congestive heart failure (428.0) be coded as left heart failure (428.1). When a hospital changes these guidelines to suit the requirements of individual hospital research, the entire pooled database can be affected. See hospital procedures, section 1.8.4.

Updates of the classification manuals, changes to HMRI/CIHI policies, and individual hospital policies and procedures are in a constant state of flux. Changes in research and new terminology necessitate changes to the classification system. Therefore, the classification manuals need to be updated periodically. Statistics Canada sends the addendum of ICD manuals out to provincial health record organizations and Ministries/Departments of Health when compiled to be forwarded on to the appropriate bodies for usage. Health record personnel then have to update the ICD manual to keep it current. This is a time-consuming process and if not done, can lead to incorrect coding\textsuperscript{17}.

1.8.3.3 To Classify a Term Not Listed

Most medical terminology used to describe disease processes appears as inclusions within a classification. However, there are several exceptions. To classify a

\textsuperscript{17}Making changes in ICD-9CM is not as cumbersome as in the ICD-9. In preparing for this study, the principal investigator worked full-time for two weeks to upgrade the 1977 revision of the Tabular List, the Index and the CCP manual to update the reference books for the fiscal year 1986/1987. There have been several updates to the classification system since that time.
medical term not listed in the Alphabetical Index of the ICD, the coder must understand the meaning of the term, the structure and organization of the classification scheme, and the principles of taxonomy (Thompson and Slee, 1978). Each step requires subjective judgement and can lead to unreliability in the classification and coding of terms that are not specifically listed in the indexes.

It is noted in the introduction to the ICD-9 classification that categories are not meant to be all encompassing, but only guidelines. It states that the terms in the Tabular List are not exhaustive, but serve as examples in the content of the category (WHO, 1978). It must also be noted, however, that health record coders do not have medical training.

For some multiple conditions, it is impossible to state with absolute certainty which diagnosis should be regarded as the most responsible. For some patients, several diagnoses may have contributed equally to the necessity for hospital admission. Several studies have been conducted to evaluate the reliability of medical data. These suggest that the presence of multiple diagnoses reduces the reliability of the data because an incorrect diagnosis may be sequenced as the principal diagnosis (Corn, 1981; Barnard & Esmond, 1981; AMRA, 1987a; 1987b; Massanari et al., 1987; Hsia et al., 1988; Wood et al., 1989). However, coders must decide which diagnosis is the most responsible. They are often faced with conflicting or ambiguous information and need to refer to sources other than the health record for clarification.
1.8.4 Hospital Procedures

For this study, the principal investigator developed a questionnaire designed to determine whether policies and procedures differ across hospitals (see Appendix D for the questionnaire). It has been suggested throughout the literature that hospital policies and procedures do indeed differ (IOM, 1977a; 1977b; Demlo, Campbell & Brown, 1978; Demlo & Campbell, 1981; Barnard & Esmond, 1981; Hsia et al., 1988).

The New Jersey Department of Health (1978) found that coding procedures unique to a hospital often produced discrepancies with the re-abstracted coding procedure. The problem is that no one knows by reviewing the database that specific hospitals have changed their guidelines. There is no requirement for submission of these changes. The researcher does not know when the data processing has been changed. When individual hospital procedures change, the aggregate database may also change and produce a bias for future studies.

When provincial Ministries/Departments of Health mandate policies that provincial hospitals have to abide by, the aggregate database may also change. When individual hospital policy differs from that of the HMRI or ICD guidelines, the aggregate database is affected. Where hospital policy restricts coders from going beyond the face sheet, discrepancies are likely to occur since coding should be based on an entire review of the medical record (NJDH, 1978). The lack of specific guidelines on selection of the MRD or in choosing procedural codes all affect the aggregate database. The aggregate database can only be assured of consistency if all hospitals and all coders use the same guidelines and procedures to code the charts.
1.9 LITERATURE REVIEW ON QUALITY OF HEALTH CARE DATABASES

This section reviews some related literature on studies that have examined the reliability of similar health care databases. Several studies regarding the quality of health care databases examining coding and classification systems have been selected for discussion. A review of these studies is necessary to determine the cause of inaccuracy of data systems. The studies are listed chronologically according to the date of data selection.

1.9.1 Hendrickson and Myers - Pennsylvania (1969)

One of the earliest studies in 1969, by Hendrickson and Myers (1973), examined the coding of routine laboratory tests using the Professional Activity Study (PAS) database in a single teaching hospital setting. The authors found a "variable but significant error rate" ranging from 31-100%. Because such a high error rate was found, the investigators limited the study to examining only routine lab tests. "The information obtained was so obviously inaccurate, that it [was] unlikely that anyone would consider using it".

They noted that the training of coders was the leading source of error and recommended improvement in training programs for hospital coders. However, the investigators noted that adequate knowledge would not necessarily guarantee a greater degree of accuracy, given the nature of the coding process. Training programs for coding personnel needed to be more effective. Individual hospital practices and procedures also contributed to the errors.

The authors’ recommended that PAS standardize the training of coders and make specific changes to the PAS abstracts. They also called for computer edit checks to verify the
data for the obvious inaccuracies. The authors’ conclusion was that such errors caused the data to be "unsuitable for a quality of care study". Obvious inconsistencies discredited all the data as a basis for such studies. They stated that "the use of computerized data to increase standards of health care must be accompanied by scrupulous care in the collection, programming and analysis of the data".

1.9.2 Roos et al. - Manitoba Health Services Commission (1971)

Roos et al. (1982) investigated the reliability of the Canadian database in one province using the Manitoba Health Services Commission data based on the ICD-8 nomenclature. Procedures and diagnoses from the hospitals’ claims and the physicians’ medical claims were compared. Research focused on those areas in which problems existed and in which the data could be relied upon. A small-scale re-abstracted record was compared with the database to check on the reliability of other information collected. The results demonstrated that the reliability for demographic and non-medical data was the highest at 92%. Diagnostic information presented the most problems. For gallbladder, diagnostic agreement was 89 to 92%. However, some conditions were more reliable than others. Agreement with procedural data was lower than with diagnostic data. Levels of agreement using the three-digit ICD-8 codes were higher than the four-digit codes. The three-digit agreement was higher than comparable three-digit data in the American studies.

The sources of error in the re-abstracted record check were reported to be the under-reporting of in-hospital consultations and some procedures and the inadequacy with ICD codes. Reporting of major operations and diagnoses were more consistent than others. The
authors' concluded that face sheet information and data on performance of major surgical procedures were found reliable in this databank. With appropriate caution, serious operations and life-threatening conditions could, they concluded, be meaningfully investigated using this databank.

1.9.3 George & Maddocks - England (1972-1973)

In 1972 & 1973 in England, George & Maddocks (1979), both physicians, compared the difference in coding between doctors and hospital clerks by examining hospital discharges using the ICD-8 nomenclature. Physicians' notes on each patient were examined and allocated a diagnostic code. These codes were compared with those listed on the form that serves as the input document for the computerized Hospital Activity Analysis, a system that enables analysis of administration, personal and clinical details of patients. This form is completed by hospital clerks.

The authors found the reliability of diagnostic codes to be 80.8% to four-digits. The differences between the doctors and clerks were confined to a few diagnoses. However, there was no account of the under-recording of diseases. The accuracy of coding for children was greater than for adults, presumably reflecting the documentation by the paediatricians, many of whom were involved in producing the Cardiff Diagnostic Classification, and presumably knowledgeable in the determinants of diagnostic classifications.

Major sources of error were determined to be completion of the record by medical staff and coding of the record by health record personnel. Errors were due to the lack of understanding of terminology in clinical medicine by the non-medical personnel. The authors
concluded that the possibility of inaccuracy must be considered when aggregated data are used for planning purposes. They suggested that medical staff, not clerks complete the diagnostic sections of the input document for the computer.

1.9.4 Institute of Medicine (IOM) - U.S. National Studies (1974)

For the calendar year 1974, the Institute of Medicine (IOM) initially reviewed the reliability of several private abstracting systems (PAS, HUP, QUEST & CHDC) in two national studies (IOM, 1977a; 1977b). Both H-ICDA and ICDA-8 nomenclatures were used. The first study grew from a concern to identify an existing database to measure the impact of the Professional Standards Review Organization (PSRO). It concentrated on hospital separation data derived from private abstracting services and compared these data to an independent abstracting of hospital records.

In the second study, the Health Care Financing Administration (HCFA) requested a similar study of Medicare claims for reimbursement. Data collection and objectives were similar but the scope was broadened to permit greater attention to methodology. The Medicare study was undertaken because of the outcome of the private abstracting services. Independent abstracting of hospital records was compared with the Medicare records.

The findings from both studies showed demographic data and admission and discharge dates to be highly reliable. However, diagnostic and procedural information were shown to be less reliable. The first study of the private abstracting services, showed the overall diagnostic data reliability rate at the four-digit level to be 65.2%; the second study established
the reliability rate of Medicare claims at the four-digit level to be 57.2%. Principal procedural data were slightly better at 73.2% and 78.9% respectively.

In both studies, reliability varied with diagnosis. For example, the results demonstrated reliability with a principal diagnosis of cataracts was 97.3% and 94.3% respectively for private abstracting services and Medicare patients, while reliability for patients with diabetes mellitus was 39.5% and 39.2% respectively.

Results from both studies also showed the three-digit coding for diagnoses to be more reliable than the four-digit. The reliability of most coded diagnoses ranged between 0 to 9.4 percentage points higher with the three-digit comparison. Some diagnoses, however, showed greater increases. For example, reliability for the Medicare patients with acute myocardial infarction was 25.7% higher with three-digit coding.

The main reasons for the diagnostic discrepancies were erroneous selection of principal diagnosis, face sheet coding, and lack of review of the record. For procedural coding, the reasons were incomplete review of the health records and differences in administrative procedures within the hospitals (Demlo, Campbell & Brown, 1978; Demlo & Campbell, 1981). Diagnoses were often stated in the health record, but not recorded on the face sheet, and therefore were not coded. From this study, recommendations were made to improve the coding of health record information, and especially to improve the inadequacies in the face sheets of health records. This study also recommended that hospitals set up programs to increase the quality of information documented in the health record, routinize hospital procedures, and set up education programs for physicians. The authors
recommended changing the face sheet to prioritize diagnoses so that the principal diagnosis is listed first.

These IOM national studies prompted many other abstracting studies that have revealed principal diagnosis reliability to range from 18.0% to 90.9% (NJDH 1978; IOM, 1980; Corn, 1981; Barnard & Esmond, 1981; Doremus & Mickenzi, 1983; Connell et al., 1984; Johnson & Appel, 1984; Lloyd & Rissing, 1985; Massanari et al., 1987; Hsia et al., 1988; Iezzoni et al., 1988; Beard et al., 1988; Lovison & Bellini, 1989; and the Ontario Health Association (OHA) et al., 1991).

1.9.5 New Jersey Department of Health Study (NJDH) (1977)

The NJDH (1978) examined two major abstracting services utilizing ICDA and H-ICDA nomenclatures. Fifty hospitals participated in the study and 2,475 records were examined. The study compared re-abstracted diagnostic and procedural data (both operative and non-operative) to the abstracting services data.

The results demonstrated non-medical data to be of 96.9% or greater accuracy. Admission dates and discharge dates were 99.7% accurate. Principal diagnosis was 72.6%, secondary diagnosis 66.5%, and other diagnoses 52.9% to the four-digit level. Principal procedure was 66.4%, secondary procedures were 59.3% and other procedures were 44.0%. Occurrences of discrepancies at the fourth digit of the code indicated that diagnostic coding was much more reliable at the three-digit level. Singular diagnoses appeared to be more consistently coded than multiple. Multiple diagnosis were likely to involve discrepancies
because of the increased number of options for the sequencing and increased requirement for coding judgement. Complex records demanded more sifting of information.

One source of error was demonstrated to be incomplete records at the time of coding. Charts were coded before completion by physicians to adhere to deadlines for abstract submission. Other problems included individual hospital policies and practices, diagnostic information in the medical record that was often not sequenced by the physician, documentation in the medical record that was often insufficient to determine with certainty which of two or more diagnoses listed on the face sheet should be abstracted as principal, and the fact that the coding classification was often inadequate to represent specific diagnostic and procedural information.

The authors’ recommended that physicians be urged to complete the medical record in a timely fashion and to clearly designate the diagnosis responsible for the patient’s admission; that abstractors be encouraged to review the entire medical record; and that hospitals establish ongoing quality control procedures for the coding process. They also explained the need for coding consistency at the four-digit level for case mix analysis, and for a thorough investigation of data consistency in anticipation of national coding of future ICD-9CM.

1.9.6 IOM’s Third National Study (1977)

The IOM’s third national study in 1977 (1980) affirmed the findings of their previous two studies. Both ICDA-8 and H-ICDA nomenclatures were used. Medical diagnoses and surgical procedures were independently abstracted and compared to the data of the abstracting services. The principal diagnosis to four-digits agreed to 63.4%. At three-digits, agreement
was 72.3%. Greater levels of specificity are captured with the four-digit code and not surprisingly, agreement decreased at this level. The principal procedure agreement was 71.4%.

The largest range in errors found in the circulatory system was caused by the over-coding of chronic ischemic heart disease. These findings demonstrated that data accuracy had not significantly increased from the data examined in 1974. Reliability was associated with the general organization, orderliness and logic of the health record. The study warned of concerns for the reliability of the introduction of five-digit coding with future national implementation of ICD-9CM.

1.9.7 Corn - New Jersey Hospitals (1977)

Corn (1981) examined ICD-9CM abstracting and coding errors for New Jersey hospitals during 1977. Methods were similar to those used in the IOM studies. Reliability was 81.5% for principal diagnosis, 74% for secondary diagnosis and 83.4% for principal procedure. When converted to DRG assignment, reliability was 82.9%. He found that DRG prospective reimbursement was fairly insensitive to errors in patient data. Although there were both over- and under-payments, errors were symmetrically distributed.

The most frequent cause of error was in judging which condition caused the patient to be admitted to hospital. Corn found sequencing errors to be a major cause of error in reimbursement. He recommended that both the major diagnosis and the condition responsible for admission be coded.
1.9.8 Barnard & Esmond - Chicago, Illinois (1979)

Barnard & Esmond (1981) compared billing data using a concurrent diagnosis with the health record using retrospective diagnosis for one hospital for the year beginning 1 May 1979. In January 1979, Medicare stopped using ICDA-8 and started using the ICD-9CM nomenclature. The investigators' purpose was to review the appropriateness of the data for use in case mix reimbursement for the Health Care Financing Administration.

Reliability was 23.1%! ICD-9-CM codes were compared to only three-digits, because the two decimal places were irrelevant to the assignment of a DRG. The major reasons for discrepancy were differences between the definition and selection of diagnosis, and errors in hospital coding.

The conclusion was that diagnostic and surgical data were not reliable for case mix reimbursement, management reporting, quality assurance, utilization review, diagnostic costing, program planning or research. The authors recommended that the data be improved and implementations of case mix reimbursement deferred until the reliability of the data is determined.

1.9.9 Doremus & Mickenzi - Cleveland, Ohio (1978)

Doremus & Mickenzi (1983) examined 262 re-abstracted records of Medicare discharges by comparing them to the original medical records at one teaching hospital using the ICD-9CM classification system during 1978. There were 52.3% agreement in diagnoses before the "coded discharge order" and 65.2% agreement with the first listed additional diagnosis. Procedures agreed to 58.4% before the coded discharge order. The results
demonstrated divergent diagnostic and procedural data that resulted in significant variation in DRG classification and reimbursement. In 32.3% of cases, discrepancies occurred in the wording of the principal diagnosis by the physician. The physician's wording of principal diagnosis on the discharge summary was not the same wording as on the original discharge order form. This resulted in errors in specificity coding at the fourth digit. It was often the case that the re-abstractor listed more additional diagnoses than the physician listed on the discharge summary.

Twenty-nine percent of procedure discrepancies were due to the order in which the procedures were listed. That is, if a biopsy was listed first, it was coded first chronologically, rather than the major procedure being a more important procedure. In 27.6% of these cases there was no listed procedure on the discharge summary. In 20.7% of these cases there was a third-digit recording error in specificity.

The study demonstrated incomplete recording of diagnostic and procedural information in the health record and significant variation between data sources. It also demonstrated that human error factors in coding were less important than the problem of recorded information in the health record. The authors suggested educating physicians on the use of discharge information. The investigators concurred with Barnard & Esmond (1981) that the data needed improvement before being used for case mix reimbursement purposes.

1.9.10 Connell, Blide & Hanken - Washington State (1978-1979)

Connell, Blide & Hanken (1984) examined 574 patients with diabetes-related conditions to assess principal diagnosis coding validity using the ICDA nomenclature. This
study was undertaken to determine whether results from a previous study that demonstrated differences in admission rates for diabetes would be due to differences in coding practices or errors. Re-abstracting of sample discharges were compared against the original medical record charts.

The study demonstrated that when coded as the principal diagnosis, diabetes could be correctly coded 60-83% of the time. Seventeen percent (17%) of cases coded as principal should have been coded as secondary. Sixty-eight percent (68%) of these were due to incorrect sequencing of codes. Only 60% of the sample was found unequivocally correct. In another 23%, the correct principal diagnosis could not be determined with certainty.

There is considerable ambiguity in how the diagnosis of diabetes can or should be applied. In over 50% of the discharges where diabetes could reasonably have been considered a principal diagnosis, the choice of correct diagnosis could not be definitely specified. These ambiguities are not a result of coding errors, but of classification and sequencing of the diabetic codes.

The implications were that with further studies of reliability, diagnostic coding should explicitly state the criterion or standard against which data are to be evaluated. The use of principal diagnosis for case finding analysis was considered highly imperfect. Ambiguities in diagnostic coding have dramatic impact on health care research that entailed the analysis of diagnosis-specific hospital admission rates. Analysis of local variations in admission rates could be impeded if hospitals have different coding practices. The problem rested with the concept of the principal diagnosis, which allowed only one condition to be coded as the cause of admission. The authors stated that enhancements in the coding of multiple diagnoses
needed to be developed. They concluded, as did Corn (1981), that despite substantial errors and ambiguities in coding, there was little variability in DRG grouping.


Johnson & Appel (1984) compared the Medicare case mix of 26 Minneapolis-St. Paul hospitals using ICD-9 with the actual case mix of hospitals as described in the health record for 1980 and 1981. Reliability for principal diagnosis in 1980 was 49.4%, while in 1981 it was 53.0%. The range of agreement by hospital for 1980 was from 31% to 74%. For 1981, agreement by hospital ranged from 36% to 71%.

The study demonstrated that hospital case mix was inaccurate about 50% of the time. The authors claimed the prospective payment DRG system had serious flaws because it offered little incentive for treating intensely ill patients and too much for less intense cases. The study demonstrated a statistically significant understatement of hospitals’ case mix which had implications for setting of DRG prices and equitable hospital reimbursement.

1.9.12 Lloyd & Rissing - Georgia (1982)

Lloyd & Rissing (1985) examined 1829 records of five Veterans' Hospitals using ICD-9CM nomenclature during fiscal 1982. Eighty-two (82) percent of the cases differed in at least one item. There was only an 18% rate of agreement. The 'source of discrepancy' analysis showed that 62.1% of the errors were physicians, 89% of which were for failure to report a procedure or diagnosis that led to inaccurate coding. Most of the procedural data not recorded were not performed in the operating room, and not usually performed by the
primary physician. The specificity required for surgical procedure coding was not provided. The second source was in coding errors that resulted in 34.5% of discrepancies. The most common coding error was an incorrect decision about what to code rather than the use of an incorrect code. The study demonstrated a substantial difference between coded data and the medical record.

1.9.13 Massanari et al. - Iowa (1984)

Massanari et al. (1987) evaluated the effectiveness of abstracting techniques in identifying nosocomial infections at discharge. Concurrent chart reviews included all acute medical and surgical patients. Only those charts where the nosocomial infection was identified prior to discharge were included in the analysis. The impact on reimbursement under the prospective payment system was assessed from September to November 1984 for one institution using the ICD-9CM nomenclature.

Reliability was 57%. Most of the infections were just not recorded. Sources of coding errors were failure of the physician to document the infection, miscoding and transcribing errors, and lack of space on the discharge abstract when patients had multiple diagnoses exceeding the allowable limit.

The authors concluded that even if the institution implements programs for careful documentation in the discharge abstract, the cost of the added codes would far exceed any projected reimbursement anticipated through reassignment of DRGs. The problem was not serious for the intended purpose, but may be serious for other uses of the data.
1.9.14 Hsia et al. - Maryland (1984-1985)

Hsia et al. (1988) measured the accuracy of coding in 7,050 health records for DRGs in 239 hospitals using ICD-9CM and receiving Medicare reimbursement during the period October 1984 through March 1985. They examined the amount of incorrect coding that occurs under the prospective-payment system. Medical records of discharged patients were re-abstracted to determine the correct DRG assignment and then compared with those originally specified by the physician and hospital administration.

This study demonstrated reliability to be 79.2% for DRG assignment. The results demonstrated that incorrect coding translated to a 20.8% error rate for DRG assignment. Of the incorrect coding errors, 61.7% resulted in more reimbursement for the hospitals.

The reasons for coding errors included mis-specification in the selection of the diagnosis, miscoding, and re-sequencing the order of the diagnoses when a secondary diagnosis produced a higher reimbursement DRG category. Physicians' erroneous selection of the principal diagnosis accounted for most of the errors. The authors concluded that DRG "creep" does occur in the coding for DRG's.

1.9.15 Iezzoni et al. - Boston (1984-1985)

Iezzoni et al. (1988) evaluated the appropriateness of diagnostic coding of acute myocardial infarction (AMI) across 15 teaching and nonteaching hospitals in metropolitan Boston between October 1984 and September 1985. Re-abstracted medical records were compared to the original hospital records. Criteria for AMI were designated. Cases
qualifying for inclusion in the study had to have at least two out of three criteria demonstrated in the medical record.

Reliability at tertiary hospitals for the diagnosis of AMI was 58.3%, while at nonteaching hospitals it was 90.9%. Two factors that contributed to the error rate were the imprecise use of terminology by physicians and the ICD-9CM coding rules. Abstractors were instructed to code diagnoses accompanied by phrases such as "rule out", "suspect", or "question" as if the disease had actually occurred. Those patients for whom AMI was ruled out on discharge were assigned the same code as those with the acute condition. This study revealed the implicit ambiguities in the nomenclature that can lead to inaccurate diagnostic coding.

1.9.16 Lovison & Bellini - Italy (1985)

Lovison & Bellini (1989) analyzed the quality of 993 hospital separation records using the Diseases of Circulatory System, ICD-8 nomenclature. A sample of records stored in the regional registry was re-abstracted and compared to the original record. No dates of the study were given, but the authors stated the study was similar to Demlo et al. (1978; 1981), George & Maddocks (1979), Hsia et al. (1988) and differed from Lloyd & Rissing (1985); therefore, data selection probably occurred after 1985.

The results demonstrated disturbing levels of inaccuracy in assignment of primary diagnosis. The study demonstrated a lack of reliability in abstracting health records. The authors suggested that the data may be seriously biased and recommended caution in using the separation data.
Sources of error included ambiguity between symptoms and the disease entity and an inadequacy with the ICD nomenclature. The use of unique codes for classification of complex clinical profiles was inadequate. Although the re-abstractors were physicians, reliability was similar to studies using health record abstractors. The authors concluded that the performances in identifying and coding primary diagnoses were unsatisfactory. The authors claimed that personnel, both medical and non-medical, were negligent in their coding techniques.


Beard, Bergstralh & Klee (1988) re-abstracted 50 records at a Mayo Clinic. Three trained registered nurses reviewed the medical records. Each record was reviewed independently by two out of the three nurses and compared with the original data. Inter-observer variability was reviewed as part of a Quality Assurance Program for monitoring medical efficacy of lab tests. No record of a classification nomenclature was noted.

The results demonstrated 82% agreement for classification of diagnostic codes, 93.4% agreement for interpretable and nominal data and 99.5% for numerical entries such as lab results and dates. The investigators concluded that inter-observer variability was a source of bias that could affect the outcome of any study and render it invalid.

1.9.18 Wood et al. - Bristol, England (1989)

Wood et al. (1989) developed a booklet using ICD codes for classifying neurology patients at one hospital in England. The purpose was to enable physicians to determine the
codes while they were assessing the patient. Medical files were coded by five independent observers and compared with the original records. Using this booklet, 84% of diagnoses were coded identically by all five doctors.

The problems that arose were that some disorders had no code listed in the ICD index, e.g., facial migraine, and some disorders had more than one applicable ICD code, e.g., all forms of headaches (excluding migraine) and non-haemorrhagic strokes were allocated to four specific codes only. Many neurology patients had no formal diagnosis. For example, a patient with "giddiness" was coded according to presenting symptoms. The study recommended that doctors accept personal responsibility for coding patients’ diagnoses at the time of consultation or discharge from the hospital.

1.9.19 Recent Canadian Studies

Although there have been several American studies and a few European studies, there have been very few Canadian ones. Until recently, the quality of the HMRI data in Canada has not been seriously examined. Prior to the data collection stage of this study, there was only one published investigation, a study of 1971 data of the Manitoba Health Services Commission, by Roos et al. (1982) (discussed in section 1.9.2). At that time, two other studies were pending publication. One was the coordinated efforts of the Ontario Health Association (OHA), the Ontario Ministry of Health (OMOH), and the Hospital Medical Records Institute (HMRI) (OHA, 1991). The other was through the Department of Paediatrics, in the Faculty of Medicine at the University of British Columbia (UBC) (Armstrong, 1992). Both studies have since been completed and are reviewed below.
1.9.19.1 *OHA, OMOH & HMRI - Ontario (Fiscal 1988/89)*

The purpose of the Ontario study (OHA, 1991) was to examine the reliability of the separation data submitted to HMRI for the fiscal year 1988/89 because of the possibility of reimbursement through the use of Case Mix Groupings (CMG’s\(^\text{18}\)) (Helyar, personal communication). Both ICD-9 and ICD-9CM nomenclatures were used. Medical records were re-abstracted by independent abstractors and compared to the HMRI database:

The results demonstrated that the agreement for the MRD was 81%, the PP was 88%, and the non-medical data ranged from 93 to 100%. There was no indication of the level of specificity used in the study, i.e., whether results were compared to the third or fourth digits. The authors found secondary diagnoses to be more susceptible to individual interpretation. In the secondary diagnoses category, after discretionary codes were removed, only 40% of the codes agreed. However, this only equated to a 14.7% difference in the assignment of CMG’s.

One source of discrepancy for the diagnostic category was that rules of the ICD classification system, Statistics Canada or HMRI had not been actively followed. The level of specificity was also found to be a problem. Coders could have either achieved a higher level of specificity by reviewing the entire record, or there was not enough detail in the record. Another source of discrepancy was hospital policy. When hospital policy was contrary to established rules, coders were forced to make an inappropriate choice of codes. Coders were faced with inconsistencies and discrepancies in the documentation within the records. In addition, discrepancies often resulted when coders tended to code by default,

\(^{18}\)CMG is a registered trademark of HMRI
i.e., choosing a code when not enough information is in the chart, or from memory (i.e., guessing the code number).

A source of discrepancy for the procedural category was that PPs were not recorded as PP's on the face sheet. It was concluded that the Canadian Classification of Diagnostic, Therapeutic and Surgical Procedures (CCP) manual needed more specific guidelines for procedural codes, especially in the selection of the principal procedure. The coding of non-operations was optional; however, some hospitals chose to code a non-operation as the principal procedure. Hospital policy also played a major role in the selection of the PP. The conclusion was that the quality of the Ontario data is consistent with the quality of the data in various jurisdictions within the United States.

1.9.19.2 Armstrong et al. (UBC) (1985)

The purpose of the Armstrong et al. (1992) study was to establish a provincial perinatal and newborn database that links Vital Statistics information with B.C. Hospital Programs information (Armstrong, personal communication). The objective was to evaluate the effectiveness of very specific ICD-9 perinatal diagnostic codes that are unique to a very limited population within very few specialized centres. The diagnosis of neonatal encephalopathy was not incorporated into the latest edition of the ICD-9 (Armstrong & Kube, 1992) and, therefore, the question arose about whether it was routinely coded by hospitals. To review this, the investigators examined the extent of agreement between the original health record coders and a second coder employed to recode 1,200 charts randomly selected from the total provincial birth population over a three year period.
The results demonstrated considerable variation among the raters (Armstrong et al., 1992). For the obstetrical diagnostic codes, the overall level of agreement was 65%. Agreement varied for individual codes, with a range across codes of 38-97%. Overall level of agreement for principal newborn codes was 58% with a discrepancy range for individual codes between 12-89%. Hospitals with Level I nurseries had lower levels of accuracy than hospitals with Level III nurseries.

The main source of discrepancy was the lack of representation of this specific diagnosis within the ICD-9 classification system. Hospitals frequently coded a chart as normal when there was a medical problem. The investigators concluded that the degree of agreement was insufficient for monitoring newborn morbidity. They stated that "specific caution should be taken to not compare the outcome across hospitals, levels of care or across provincial/state, national or international levels unless evidence is available for the accuracy of diagnosis".

1.9.20 Summary of Studies in General

As demonstrated by the above summaries, coded data and health care databases have not been proven to be very reliable. Reliability of the coded diagnostic data ranged from 18.0% to 90.9% of cases accurately coded. Procedural data ranged from 58% to 88.0% accurate coding. Error rates as high as 82.0% in diagnostic data and up to 42% for procedural data are significant when one considers what the data are used for.

The studies have determined several sources of discrepancy. The primary sources were three: coding, documentation, and system errors.
1.9.20.1 Coding Errors

Coding errors were due to miscoding (Lloyd & Rissing, 1985; Massanari et al., 1987; Hsia et al., 1988), transcription (Massanari et al., 1987), lack of understanding in medical terminology (George & Maddocks, 1979) and coding errors not specified (George & Maddocks, 1979; Barnard & Esmond, 1981; Doremus & Mickenzi, 1983). Lloyd & Rissing (1985) found that the miscoding errors were based on what to code, i.e. the decision on whether to code, rather than in choosing the actual code. The OHA et al. (1991) found that when codes were generated in error, the coders were either coding by default, i.e. determining a code when there was not enough information in the chart, or coding from memory. Specificity was a problem with several studies reporting higher accuracy with the three-digit code than the four-digit code (IOM, 1977a; 1977b; 1980; George & Maddocks, 1979; Roos et al., 1982). The OHA et al. (1991) claimed that coders could have achieved a higher level of specificity by gleaning more information from the chart.

Hendrickson & Myers (1973) were the first to suggest that coding skill and training needed updating and standardizing. Other authors have confirmed this finding (IOM, 1977a; 1977b; Demlo, Campbell & Brown, 197818; George & Maddocks, 1979; Barnard & Esmond, 1981; Doremus & Mickenzi, 1983; Lloyd & Rissing, 1985; Massanari et al., 1987; Hsia et al., 1988). Demlo & Campbell (1981)19 have stated that there is a real association between the general organization, orderliness and logic of the health record and accurate coding. Doremus & Mickenzi (1983) stated, however, that the number of coding errors was less than

18Demlo, Campbell & Brown, 1978 and Demlo & Campbell, 1981 did not conduct studies, but were active staff participants in the IOM's studies and published articles documenting its findings and recommendations after completion.
the number of discrepancies in the chart. A problem with documentation was that even if all
the information was available, failure to conduct a full review of the chart was a major
problem within the coding process (IOM, 1977a; 1977b; 1980; NJDH, 1978; Demlo,

1.9.20.2 Documentation Errors

Documentation was the second category of discrepancies. Coding from the face sheet,
(IOM, 1977a; 1977b; 1980; NJDH, 1978; Demlo, Campbell & Brown, 1978; Demlo &
Campbell, 1981; OHA et al., 1991) especially before completion of the chart by a physician
(NJDH, 1978) and the under-recording of information (George & Maddocks, 1979; Roos et
al., 1982) were the main charting problems. George & Maddocks (1979) state that the
under-recording of information was mainly due to physicians not completing the health
record. Several authors have stated that coding errors result from a failure by physicians to
report the necessary procedure or diagnosis (IOM, 1977a; 1977b; Roos et al., 1982; Lloyd &
Rissing, 1985; Massanari et al., 1987). The IOM (1977a; 1977b) recommended major
changes to the face sheets to enable diagnoses and procedures to be prioritized so that the first
diagnosis or procedure listed is the principal one.

Charting was a major problem with documentation. Several investigators have stated
that physicians are inconsistent in their description of the diagnosis throughout the chart; or
that the wording of the diagnosis can lead to errors in specificity of coding (Doremus &
Mickenzi, 1983; Lloyd & Rissing, 1985; Iezzoni et al., 1988; OHA et al., 1991); or that
there were discrepancies in the chart between physicians (OHA et al., 1991). Doremus &
Mickenzi (1983) have stated that the way the physician lists the diagnoses or procedures on the discharge summary is the order in which they will be coded. They also state that physicians do not even list procedures on the discharge summary, especially non-operative procedures and those procedures not done by the primary physician. The IOM (1977a; 1977b) recommended that education programs for physicians be set up at each hospital, but Demlo & Campbell (1981) take this further and state that training needs to occur in the medical schools. The U.S. national studies recommended that quality assurance programs were needed to increase the quality of documentation in the chart (IOM, 1977a; 1977b; Demlo, Campbell & Brown, 1978; Demlo & Campbell, 1981).

1.9.20.3 System Errors

Coding rules and conventions were the third category of discrepancy. The IOM study was the first to determine that some codes were more reliable than others (IOM, 1977a; 1977b; 1980; Demlo, Campbell & Brown, 1978; Demlo & Campbell, 1981). For instance, for diagnoses such as cataracts, there was consistent agreement over 94%, while for diabetes, agreement was consistently below 40%. Classification problems in general (NJDH, 1978; Connell et al., 1984), ambiguities in nomenclature (Iezzoni et al., 1988) as well as judgement, sequencing and selection errors (IOM, 1977a; 1977b; Demlo, Campbell & Brown, 1978; NJDH, 1978; Corn, 1981; Barnard & Esmond, 1981; Connell et al., 1984; Hsia et al., 1988) were among those most commonly referred to.

Hendrickson & Myers (1973) stated that the abstract itself was a problem. They also felt that abstract vendors could make more of an effort to do computer edits to verify the
information they receive. Massanari et al. (1987) mentioned the lack of space on the abstract, especially when the diagnoses and procedures exceed the allowable limit.

Roos et al. (1982), Wood et al. (1989), and Armstrong et al. (1992) point out that for some diagnoses, codes are not applicable. Wood et al., (1989) and Armstrong et al. (1992) claim that some diagnoses are not listed in the Alphabetical Index at all, while some disorders have more than one applicable code (Wood et al., 1989). ICD coding rules also incorporate the coding of phrases, such as "rule out", "suspect", "question", as if the disease actually took place, making it difficult to differentiate between the true incidence of disease, and the number of occurrences actually coded (Iezzoni et al., 1988).

The definition and selection of principal diagnosis are problems in situations involving multiple diagnoses and complex records when only one diagnosis is to be recorded (NJDH, 1978; Connell et al., 1984; Lovison & Bellini, 1989). Connell et al. (1984) advocated the need for a system that can handle multiple diagnoses.

The definition and selection of principal procedure (Barnard & Esmond, 1981; Connell et al., 1984; OHA et al., 1991) was also a problem, especially where more than one procedure was identified. The OHA et al. (1991) stated that more specific guidelines were needed for the CCP, especially in determining the principal procedure. The investigators mention that the coding of non-operative procedures is optional and hospitals can code these as the principal procedure if they so choose.

Hospital procedures that conflict with established rules force coders to make inappropriate choices (Hendrickson & Myers, 1973; IOM, 1977a; 1977b; Demlo, Campbell & Brown, 1978; NJDH, 1978; Demlo & Campbell, 1981; OHA et al., 1991). Hospitals also
choose to ignore or change the rules of the ICD, Statistics Canada or HMRI (OHA et al., 1991).

1.9.20.4 Reassignment to DRG/CMG

It has been found generally that reassignment of codes to DRG/CMG categories is fairly insensitive to coding errors (Corn, 1981; Connell et al., 1984; OHA et al., 1991). Since only the three-digit assignment of a code is necessary for DRG/CMG assignment, the use of one or two decimal places is irrelevant to the assignment (Barnard & Esmond, 1981). The Ontario study showed up to 40% errors in secondary diagnostic categories, but only a 15% error in assignment to CMG category (OHA et al., 1991).

1.9.21 Conclusion of Literature Review

It has been shown that there is variation between data sources. The overall conclusion from the studies reviewed is that most of the databases examined were not suitable for the activities of program planning and development, quality assurance monitoring, inter-system hospital comparisons, epidemiologic studies, disease demography, resource allocation, or patient-oriented medical research. There was a great deal of variation between the data sources. In addition, since most of the databases are organized by diagnosis (principal diagnosis for American data, and most responsible diagnosis for Canadian data), case finding can be imprecise (Connell et al., 1984), especially when utilizing pooled data. As stated by Beard et al. (1988), inter-rater reliability of data collection can seriously hamper the results of a study, thereby potentially biasing any study utilizing these databases.
In the Canadian studies, the Ontario agreement rates for the MRD were much higher than most American studies for principal diagnosis at the four-digit level, but within the same range at the three-digit level. For PP agreement, the Ontario study figures were higher than for any American study. However, as stated previously, since no level of specificity was clarified in the study, it is assumed that agreement to the three-digit level of coding for diagnoses only was compared, as only the three-digit codes are required to determine the CMG assignment. The Armstrong et al. (1992) study fell in the mid-range for diagnostic four-digit coding and on the low side for procedural code agreement.
1.10 SUMMARY

The unreliability of hospital-derived diagnostic and procedural databases in general has been consistent throughout the literature. As of April 1990, the Ministry of Health in British Columbia has mandated that all hospitals participate in the HMRI (CIHI) system. The Ministry is also considering the use of this database as a funding source for hospitals, as has been begun in other provinces.

This study of the reliability of coding in a subset of Vancouver hospitals is necessary for the reasons outlined throughout this chapter. Reliable, accurate and timely clinical data are necessary to enable health care professionals, administrators and planners to make decisions about the quality of care and researchers to have access to an accurate database. Therefore, such a database, if reliable, has considerable potential in helping to meet the health needs of the nation effectively and efficiently.

The importance of training health record coders and coding from completed health records, was a consistent message throughout the literature. Today’s demands for accurate data require that procedures be developed for assessing the level of accuracy, for analyzing the causes of error, and for recommending actions needed for improvement. Objective criteria must be designed so that they detect errors in individual codes, the totality of codes and in the sequencing of codes. Assessment of coding accuracy must address the reliability of the coding process rather than the consistency of individual codes.

The Ontario study (OHA et al., 1991) concluded that only 40% of secondary diagnostic codes agreed to three-digits; however, the investigators found that this translated to only a 15% mis-assignment to CMG categories. Although the investigators of the Ontario
study did not see this as a concern for the narrow CMG classification purpose, a 40% non-agreement at the three-digit level will compromise the accuracy of national and provincial aggregate databases which are aggregated from the HMRI (CIHI) data and used in independent research involving inter alia inter-hospital comparisons. With the recent trend to reviewing workload by case mix, it becomes essential that all the basic medical details are accurate. This includes individual elements within the record, not just the CMG category assignment.

Based on the wide utilization of the HMRI (CIHI) database in B.C. and the future potential of funding through the assignment of CMG's, it is essential that reliability of this database be examined. The quality of clinical data is critical to the future of the health care delivery system: the accuracy of the data significantly affects all concerned. This study proposes to examine the reliability of this database as outlined in the following chapters.
CHAPTER 2 - METHODOLOGY

2.0 RATIONALE

The literature review shows that, in general, classification systems have very low reliability. Within the province of B.C., the Hospital Medical Records Institute/Canadian Institute for Health Information (HMRI/CIHI) database is currently used for program planning, program evaluation, monitoring and analyzing programs, utilization reviews, quality assurance, administrative decision-making, allocation of resources, accreditation, research, medical education, inter-hospital comparisons, and policy formulation. With the present trend towards utilizing this information for funding of the health care system using the HMRI's Case Mix Groups (CMG's\textsuperscript{20}) and Resource Intensity Weights (RIW's\textsuperscript{20}), it is imperative that reliability of this system is examined.

As revealed in the previous section, competent coders are necessary (although not necessarily sufficient) for the creation of accurate computerized abstracted health records. Missed secondary diagnoses, complications and comorbidities play a decisive role in determining whether the correct diagnostic category is chosen. Such decisions could affect such things as hospital funding, missing or inaccurate data for research purposes and the miscounting of diagnoses in determining cumulative incidence. The importance of coding\textsuperscript{21} data accurately is increasing constantly in today's environment, and health care

\textsuperscript{20}CMG and RIW are registered trademarks of HMRI(CIHI)

\textsuperscript{21}See Appendix A for definition of the term "coding". 
facilities must therefore employ competent coders to ensure the veracity of not only their own, but of local, provincial and national databases (see Appendix A for definition of coding).

The main aim of this research is to determine the inter-rater reliability of coders utilizing the HMRI/CIHI database at six Vancouver hospitals in British Columbia (see Appendix E for listing of hospitals). Other purposes include reviewing the types of discrepancies and determining whether individual hospital policies and procedures differ.

As Feinstein (1988) stated:

"I believe this type of information is essential for monitoring the health of the population and the delivery of medical services. I believe that with recognition of the problems come solutions to overcome them and the challenge is worth pursuing."
2.1 OBJECTIVE AND PURPOSE

The main objective of this study is to compare selected data items within the HMRI/CIHI in-patient hospital discharge database against a comparable re-abstracted set of data items selected from the original health record in a sample of six Vancouver acute care hospitals for the fiscal year 31 March 1986 to 1 April 1987. The purpose is to determine the degree of inter-observer agreement between the original coder and the research coder. Where agreement does not exist, the frequency and type of discrepancy is examined. Agreement with the research coder is used as a proxy for accuracy and lack of agreement is at least an indication of some inconsistency between the database and the health care record.

On the basis of this analysis, this study suggests ways to improve the reliability of the database. It presents a preliminary look at the HMRI/CIHI database in B.C. to determine whether province-wide investigation on reliability of the database is warranted.

This study does not examine the extent to which the recorded information in the health record accurately reflects the patient's condition, or the quality of the data in the record. These are important questions, but lie outside the scope of the present research.

---

22The definition of a coding discrepancy is described in Appendix A.
2.2 THE QUESTIONS PROPOSED FOR THIS STUDY

The questions posed for this study are as follows:

- What is the level of agreement between selected variables in the HMRI/CIHI inpatient hospital discharge database and the re-abstracted variables from the original health record within the selected hospitals in Vancouver, B.C.?
- Are there differences in agreement scores among the selected hospitals?
- Are there differences in agreement scores among the selected diagnostic categories?
- Are there differences in agreement scores among the coders within and across hospitals?
- Are there particular dominant sources or types of discrepancy?
- Is there a need for improvement?
2.3 HYPOTHESES

The hypotheses developed for this study are:

- The level of agreement is less than 90% between the selected variables in the HMRI/CIHI inpatient hospital discharge database and the re-abstracted variables from the original health record within selected hospitals in Vancouver, B.C.
- There are less than 15% difference in agreement scores among the selected hospitals.
- There are less than 15% difference in agreement scores among the selected diagnostic categories.
- There is one dominant type of discrepancy.
- There are less than 15% difference in agreement scores among the coders within a hospital.
2.4 NULL HYPOTHESES

The null hypotheses developed for this study are:

• The level of agreement is greater than 90% between the selected variables in the HMRI/CIHI inpatient hospital discharge database and the re-abstracted data from the original health record within the selected hospitals in Vancouver, B.C.

• There are greater than 15% difference in agreement scores among the selected hospitals.

• There are greater than 15% difference in agreement scores among the selected diagnostic categories.

• There is more than one dominant type of discrepancy.

• There are greater than 15% difference in agreement scores among the coders within and across hospitals.
2.5 SAMPLE SIZE CALCULATIONS

Sample size calculations were determined prior to the selection of cases. Two subsamples were determined: a distinct diagnostic category subsample and a stratified random subsample. The sample size calculation is based on the non-central chi-squared distribution (Greenland, 1985). The non-centrality parameter (lambda) can be calculated in two independent ways, once from the expected alternate hypothesis data (which actually gives the ratio of lambda to the sample size: lambda/N) and once from the design specifications (alpha, beta, degrees of freedom). Equating these two expressions for lambda, one can solve for N. In fact, N is lambda (given by the second method) divided by lambda/n (from the first method). See Appendix F for details on sample size calculations and Appendix G for details on standard error calculations and confidence intervals.

2.5.1 Three Distinct Diagnostic Categories Selection

Three distinct diagnostic categories were selected from those most frequently occurring for hospitals in B.C. during the fiscal year 1 April 1986 to 31 March 1987. See Appendix H for a listing of the most frequently occurring diagnoses in B.C. The three diagnostic categories selected from this listing were diabetes mellitus (250-259), myocardial infarction (410-414), and bronchopneumonia (490-496). This selection was necessary to determine if there were differences in agreement scores among the selected diagnostic categories.

For this sample hypothesis testing situation, using a one-sided test, with level of significance alpha(α)=0.05, beta(β)=0.20, and degrees of freedom(df)=5,
lambda/N = 0.0505, Chi critical point \((\alpha = 0.05) = 5.99\), lambda \((\beta = 0.20) = 9.64\). The alternate hypothesis is that one diagnostic category has a discrepancy rate of 10%, \(p_0 = .10\), while two have 30%, \(p_1 = .30\). The null hypothesis is that the same number of discrepancies occur as specified in the alternate hypothesis, but equally distributed across groups. Total sample size per diagnostic category was determined to be 195 cases.

Oversampling was not necessary as alternate records were prepared to ensure adequate representation within the selected categories. A decision was made to round up each diagnostic category to 20 cases \((20 \times 3 \text{ categories} \times 6 \text{ hospitals} = 360 \text{ cases})\).

2.5.2 Stratified Random Categories Selection

To detect differences in agreement scores among hospitals, ten diagnostic categories were defined using the ICD-9, three-digit categories. Selection was based on a stratification of these categories. Not all hospitals in the study admitted psychiatric, maternity or cancer patients; therefore, these types of patients were excluded from the ICD-9 categories. Excluded three-digit ICD-9 categories were Neoplasms, Mental Disorders, Complications of Pregnancy, Childbirth and the Puerperium, and Certain Conditions Originating in the Perinatal Period. The total stratified random sample component includes those cases previously selected in the distinct diagnostic categories. See Appendix I for the details on the ICD-9 classification of categories defined for the stratified random sample.

This testing situation is the same as in the first subsample. The alternate hypothesis is that three hospitals have discrepancy rates of 10% and three have 30%. The
null hypothesis is that the same number of discrepancies occur as specified in the alternate hypothesis, but equally distributed across groups.

Direct standardization would increase the power of the study (Kahn et al., 1988). Therefore, cases were selected distributed by the frequency of the U.B.C. hospital's case mix by diagnostic distribution. The U.B.C. hospital was chosen as the base because the distribution of cases had no obstetrics, no paediatrics, and no open heart surgery. This pre-determined mix included all common conditions of diagnostic interest and excluded unusual or atypical cases. Calculations for the other five participating hospitals in the study were based on U.B.C.'s distribution of cases.

When the category proportions were determined, whole numbers were assigned and low categories were augmented. Partial percentages were increased up to the nearest round number.

It was calculated that a total of 206 cases was necessary to compare the hospitals. The total sample of cases to be used for hospital comparison purposes included the 360 cases selected within the three distinct diagnostic categories for a total of 566 cases stratified by diagnostic category.

Oversampling was not necessary as alternate records were prepared to ensure adequate representation within the selected categories. The total number of cases was rounded up to 600 (100 cases per hospital). Forty records per hospital were randomly selected after stratification within the ten diagnostic categories. The 360 cases previously selected plus 240 cases (40 cases per hospital) were selected for the random sampling component of the study. See Appendix F for details on the number of selected cases per stratified random sample.
2.5.3 Total Sample Size

Twenty (20) cases for each of the three distinct diagnostic categories (i.e., 60 cases for each hospital) and 40 stratified random cases were selected per hospital. One hundred cases per hospital, for a total of 600 cases were selected for the study. Because alternates were determined and tracking of records was difficult, total re-abstracted charts resulted in 606 cases. See alternate records in section 2.6.5 and tracking of records in section 2.15.3.
2.6 SAMPLING PLAN

A stratified sampling plan was used to determine the selection of cases for the study.

2.6.1 Selection of Hospitals

All hospitals in the Greater Vancouver Regional Hospital District were initially included in the selection process. The selection criteria included only acute care non-specialty facilities (i.e., not a rehabilitation or maternity hospital) utilizing the ICD-9 nomenclature during the entire fiscal year 1 April 1986 to 31 March 1987. A sample of six acute care hospitals was selected for the study. All hospitals had agreed to participate prior to drawing the sample of cases. The selection of hospitals included Lion’s Gate Hospital, University Hospital - U.B.C. site, University Hospital - Shaughnessy site, Vancouver General Hospital, St. Paul’s Hospital and Mount Saint Joseph’s Hospital. See Appendix E for a listing of the selected hospitals.

2.6.2 Sampling Frame

The sample of cases was selected from all acute medical and surgical separations from the six hospitals occurring in the fiscal year 1 April 1986 to 31 March 1987. There were 90,509 acute care separations from all six hospitals during that fiscal year (see Appendix F).
2.6.3 Selection of Cases for the Distinct Diagnostic Categories

Individual cases were selected in each of the three distinct diagnostic categories. Criteria for selection included that the case had the appropriate code in either the most responsible diagnostic position or the secondary diagnostic position, i.e. each case was chosen because the diagnosis could be found in either the MRD or the secondary diagnosis position within the MOH database.

Twenty cases per distinct diagnostic category (20 x 3 = 60 in all) were randomly selected for each hospital. For the six hospitals in the study, a total of 360 cases were selected.

2.6.4 Selection of Cases for Stratified Random Sample

The selection of 40 cases per hospital was based on a stratified (by diagnosis) random sample in an attempt to match the frequency of case mix for each facility to the U.B.C. Hospital's case mix.

Cases excluded in this stratified random selection were the cases previously selected for the first subsample, the three distinct diagnostic categories. That is, no case could be sampled twice: those cases that were included in the original three diagnostic categories, i.e. diabetes (250-259), myocardial infarction (410-414), and bronchopneumonia (490-496) were not re-selected in this stratification process.

Criteria for selection included that the case had the appropriate code either in the most responsible diagnostic position or the secondary diagnostic position for each individual category, i.e. each case was chosen because the diagnosis could be found in
either the MRD or the secondary diagnoses position within the MOH database. The selection of cases was limited to acute care discharges, both medical and surgical cases.

2.6.5 Alternate Records

Alternate records were selected to ensure that adequate representation within all the categories was maintained. If a record was not available, an alternate was chosen. The use of alternate records and the inability to track records increased the total number of cases from 600 to 606. See tracking of records in section 2.15.3.

2.6.6 Subsample for IRR Reliability

A 10% random subsample of the 600 previously selected cases, stratified by hospital, was selected for intra-rater reliability of the p.i.. No attempt was made to separate these cases into specific diagnostic categories, as the sample was too small. There were 10 charts per hospital, for a total of 60 cases.
2.7 RESEARCH DESIGN

The research design was a retrospective descriptive study for the year 1 April 1986 to 31 March 1987. This fiscal year was selected to coincide with the Health Database Linkage Project: A Feasibility Study, (Linkage Study) conducted by the Centre for Health Services and Policy Research at the University of British Columbia (U.B.C.). The data from three of the hospitals in this study were also utilized by the Linkage Study. One hundred health records were re-abstracted for each of the six selected hospitals for a total of 600 cases. These re-abstracted records were compared with the HMRI/CIHI database as provided by the B.C. Ministry of Health.
2.8 DATA SOURCE

The sources of data comparison are:

- the HMRI/CIHI in-patient hospital discharge database from the B.C. Ministry of Health (MOH), Hospital Programs Data File (See Appendix J);
- the re-abstracted data items obtained from the original health records; and
- the face sheet codes on the original health records.

Documentation of permission to access the data from the MOH and each participating hospital is provided in Appendices K through P.
2.9 RESEARCH INSTRUMENTS

There were three research instruments used in this study: a re-abstracting form, a hospital questionnaire, and the *International Classification of Diseases, Ninth Revision*, (ICD-9) classification system that consists of Volume 1, Tabular List and Volume 2, Alphabetical Index for diagnostic coding. Included as part of the classification system is the *Canadian Classification of Diagnostic, Therapeutic and Surgical Procedures* (CCP), developed by Statistics Canada, Health Division, Nosology Reference Centre for procedural coding.

2.9.1 Re-abstracting form

Input variables re-abstracted from the original health record onto a re-abstracting form include the date of hospital admission, date of hospital separation, discharge status, sex, birthdate, coder identification number, the Most Responsible Diagnosis (MRD), secondary diagnoses, the Principal Procedure (PP), and secondary procedures. See Appendix Q for an example of this form.

This form, developed by the principal investigator (p.i.) includes the necessary re-abstracted data items from the record and the data items from the MOH hospital in-patient discharge database.

2.9.2 Hospital Questionnaire

It was speculated that individual hospital policies, procedures and practices in determining the MRD could differ and that this might be one cause of variation across hospitals in apparent incidence of particular codes. The hospital questionnaire was
directed to the Head of the Medical Record Department and/or Coding Supervisor (supervisor). The p.i. consulted with Deanna Hickie, HMRI/CIHI liaison representative and Brenda Justason, then Chairperson of the Data Quality Committee of the B.C. Health Record Association before developing the contents of the questionnaire. The feasibility of collecting this information was discussed. This questionnaire was not piloted as there were only six hospitals participating in the study. Appendix R contains the questionnaire.

2.9.2.1 **Purpose of Questionnaire**

There were two purposes to this questionnaire. The first purpose was to gather information about coder training. It was utilized to determine the place of coder training, the credentials, the number of years' experience and the on-going training, such as workshops, of each coder employed by the hospital during the fiscal year of the study. There was a problem with recall so employee records were accessed. Complete details were made available to the p.i. a few weeks after leaving each site.

The second purpose was to gather information about specific policies and procedures within each hospital that could have an effect on the aggregate database. It was used to determine the process each hospital used when a problem concerning selection of a diagnosis or procedure was encountered. Questions were posed to determine the type of system used, the process used in the selection of the MRD, the timing of completion of records, individual hospital procedures, and in particular, those that differ from the HMRI/CIHI, the International Classification of Diseases (ICD) or the B.C. MOH guidelines. The questionnaire was intended to elicit information on overall hospital procedures and did not relate to individual records.
2.9.3 ICD-9 Classification System

The diagnostic categories used in this study were determined using *The International Classification of Diseases, Ninth Revision, 1977* (ICD-9) developed by the World Health Organization (WHO). The ICD-9 nomenclature is used within the HMRI/CIHI classification system and forms the basis of the coded diagnostic elements within this study. The ICD-9 system consists of two diagnostic volumes, the Listing of the Categories and the Alphabetical Index, and also one procedural volume. However, for procedural coding, Statistics Canada, Health Division, Nosology Reference Centre has directed that the *Canadian Classification of Diagnostic, Therapeutic and Surgical Procedures* (CCP) is to be used within all Canadian institutions (Statistics Canada, 1986). The combination of these three reference books forms the basis of the HMRI/CIHI classification system which is the gold standard within the B.C. health care database. Although some Canadian hospitals utilize the American version, ICD-9CM, for diagnostic coding, the CCP is the only option for procedural coding in Canada. See section 1.5.2 for details on Canadian usage of ICD-9CM. HMRI(CIHI) converts the ICD-9CM codes into ICD-9 codes for the appropriate Ministries/Departments of Health.
2.10 STUDY PLAN

The study plan involved a multi-step process that included proposals submitted to each of the six individual hospital research committees. Other procedures included updating all the pertinent diagnostic and procedural coding reference books, p.i. training, selection of cases and preparation of original records, hospital questionnaire interview, re-abstracting process, reconciliation process, compiling information and re-filing health records. A second re-abstracting subsample was developed for intra-rater reliability (IRR) of the p.i. before completing the computer input and statistical data analysis. See Appendix S for a summary of the study plan. Appendix T contains the approval for ethical review and Appendix U the approval from the thesis screening committee.

2.10.1 Individual Hospital Research Committees

Research proposals were sent to each of the six hospitals selected in the study for approval by the appropriate research committee or individual. Approvals from individual hospitals are documented in Appendices K through P.

2.10.2 Updating Diagnostic & Procedural Code Books

Before the process of re-coding and re-abstracting commenced, the three code books used in this study, *The International Classification of Diseases, Ninth Revision* (ICD-9), Volumes I and II, and the *Canadian Classification of Therapeutic, and Surgical Procedures* (CCP), were updated using the ICD-9 Consolidated Errata (Statistics Canada, 1986) up to March 31, 1986 to ensure that the coding nomenclature used in this study
contained similar information to that available to the original hospital coders during the fiscal year, 1986/87.

2.10.3 Principal Investigator Training

The principal investigator, who did the re-coding and re-abstracting, has a health record background. Since it had been a few years since she last utilized her coding skills, she spent a few months reviewing anatomy and physiology textbooks and reviewing current coding processes and procedures.

She supplemented her health record credentials, her previous practical experience, and a review of relevant material with a two-day training session in Basic Training I for ICD-9 and Diagnosis Typing conducted by a qualified HMRI/CIHI trainer. See Appendix V for details on training sessions.

To determine competency within a representative group, the p.i. was also tested against several current hospital coders located throughout the B.C. Lower Mainland who are actively utilizing the ICD-9 classification system. Selection and testing of these coders were conducted by Deanna Hickie, HMRI/CIHI Liaison Representative for B.C. The p.i. tested at the higher end of the scale. See Appendix W for testing results.

2.10.4 Selection of Cases and Preparation of Original Records

The study was conducted at each of the six hospital sites. See Appendix X for details on dates and contacts for each of the participating hospitals. Details on sample size and sampling plan were reviewed in sections 2.5 and 2.6 respectively.
The B.C. MOH listing of randomly selected cases by diagnostic category, prepared by the computer analyst and documented in section 2.11.3. was sent to each hospital before the arrival of the research team. This listing also included the randomly selected alternate cases for each category. The health record clerks, employed by the hospital, ensured that the selected records were readily available prior to the arrival of the p.i. and her assistant.

The assistant determined that the correct records were made available to the p.i.. She checked each record against the MOH listing of selected cases. If a wrong record was supplied, she requested the correct one. When the correct one was not available or if a record was missing, she requested an alternate record from those determined in the MOH listing.

Testing of reliability requires similar conditions within the environment between the coders; therefore, face sheets on the original health records were retained. It is a common procedure for hospital coders to document the diagnostic and procedural codes on the face sheet of the health record. Therefore, codes were removed for the p.i. The assistant photocopied the original face sheets ensuring that all previous codes were removed and replaced the original face sheet with the photocopied one. This enabled the p.i. to be blinded to the original codes, yet face similar conditions to those confronted by the original coder. The original face sheets were sorted in numerical order and given to the coding supervisor until the reconciliation stage of the study. Records were sorted and coded in numerical order to ensure the p.i. was blinded to diagnostic category.

Previous records were made available for the p.i.. It is noted, however, that the previous record was only accessed when a documentation problem in the record needed
verification. Coders also contact physicians for clarification, depending on hospital policy and individual coding practices. These are similar to the process currently utilized by hospital coders. At the request of each of the hospitals, no attempt was made to contact the attending physician for clarification of any documentation problem. If a previous clarification was not documented in the record, this could have produced a bias. There was no way to determine how much bias was introduced at this stage.

2.10.5 Hospital Questionnaire Interview

The p.i. interviewed the coding supervisor and/or the Director of Health Records at each individual hospital site to complete the questionnaire.

Information on workshops attended by coders prior to the year in question was obtained from Deanna Hickie, HMRI/CIHI liaison representative, who kept accurate records of every HMRI/CIHI workshop and its attendants. Permission was obtained from all the hospitals to access this information. In an attempt to maintain anonymity, once the individual hospital coders were connected with the correct workshops, the coders were assigned a unique coder number and identifying information was destroyed.

Each hospital uses a distinct numbering system for designating individual coders. These numbers are assigned by hospital and the coder has the responsibility to ensure this number is placed on the face sheet when the chart is abstracted. The p.i. was blinded to the identity of the coders until the reconciliation stage of the study. When all the hospitals' data were combined, it was discovered that several coders from different hospitals had the same coder numbers. Therefore, to distinguish one coder from another,
each coder was assigned a computer number from one to thirty-eight to maintain their individuality as well as to conceal their identity, both necessary for this study.

2.10.6 Re-Abstracting Process

Individual health records at each hospital site were then re-coded and re-abstracted by the p.i.. The order in which the records were re-abstracted was recorded on the re-abstracting form. There was no priority given to the secondary diagnoses or secondary procedures and there was no attempt to determine the importance of a diagnosis or a procedure beyond that of the MRD and the PP. The purpose of the study was to determine the level of agreement between variables such as the MRD and the PP and to compare agreement scores among hospitals.

Re-abstracting consisted of reviewing the entire record and determining the correct selection of codes using both the original record and the ICD code books. The process used was similar to actual coding within the hospital setting, with the exception that no attempt was made to contact attending physicians or other professionals for confirmation of information within the record. Coder numbers were recorded to distinguish among coders to compare agreement scores among them.

No HMRI/CIHI "typing" of the codes was necessary for this study. Typing is an addition to the code to determine whether the diagnostic condition is primary, i.e. a condition that contributed to the length of stay; a complication, i.e. a condition arising after beginning treatment; or a secondary condition, for which a patient may or may not have received treatment, but which is also an important condition that may have an influence on the length of stay. Since the purpose of this study was to determine
agreement scores between the ICD codes in the original and the re-abstracted datasets, it made no difference where the code was positioned, just that it was coded. Typing is a very subjective process and since the Ontario study demonstrated 40% inaccuracy with typing alone (OHA et al., 1991) it was felt that typing of the diagnoses would lead to a lesser agreement score if used in this study.

At the completion of the re-abstracting process at each hospital site, the assistant replaced the original face sheets on the records and ensured records were again sorted in numerical order. The photocopied face sheets were then destroyed.

2.10.7 Compiling Information and Re-Filing Health Records

After the p.i. had completed the re-abstracting and re-coding of all records from each hospital, the assistant copied the codes directly off the original face sheet onto the left margin of the re-abstracting form, presenting the second set of codes on this form. These codes were documented using a different coloured pen to be able to distinguish them from the re-abstracted codes.

After the reconciliation process, documented in the next section, and after the original health record codes were entered on the re-abstracting form, the assistant opened the sealed envelopes containing the individual MOH data items for each case previously extracted by the computer analyst (see section 2.11.3 for details). She then entered these data items on this re-abstracting form to the right side of the re-abstracted data, which produced three sets of data: A) the hospital face sheet codes, B) the re-abstracted codes and C) the MOH's individual data items.
2.10.8 Reconciliation Process

The initial reconciliation process involved a review of the MRD and the PP codes. Where a discrepancy was determined to exist, the p.i. re-checked the information in the original health record and determined the type of discrepancy. Type of discrepancy was recorded to determine if there was any particular dominant type. Determining the type of discrepancy initially represents the professional judgement of the p.i.. The type of discrepancy categories included Judgement, Classification, Procedural, Clerical, Incomplete Record, Discrepant, and Other. See section 2.12 for details.

The health records were re-examined to determine the type of each individual discrepant code. Notes were made on the re-abstracting form. Each record and re-abstracting form along with its accompanying notes was then reviewed with the coding supervisor to determine if there was general agreement between the p.i. and the supervisor. Together, the p.i. and the supervisor reviewed the health records to determine whether the original or the re-abstracted data codes were correct and to determine the type of each coding discrepancy. It is noted here that if one case had more than one discrepant code, then it follows that it may also have more than one type of discrepancy. When agreement between the p.i. and the supervisor was not reached, it was noted accordingly on the re-abstracting form. Although it would appear that the supervisor would tend to claim that the original coding was correct, there were very few records where agreement was not reached.

A cursory review only of secondary diagnostic and procedural codes was undertaken at this time. An in-depth review was performed later by the p.i..
2.10.9 Subsample for Intra-Rater Reliability

A second re-abstracting process, similar to the first, was conducted to determine the intra-rater reliability (IRR) of the p.i. A 10% random subsample of cases, stratified by hospital, was selected from the original sample by the computer analyst (10 cases per hospital x six hospitals = 60 cases in total). Because the subsample was small, it was not stratified by diagnostic categories. See section 2.6.6 on the sampling plan for the IRR subsample.

This second process was commenced at the first hospital site two weeks after the completion of the re-filing of the records at the last hospital site. There was approximately a three month gap between the two re-abstracting sessions at the first hospital and a four week gap between the sessions at the last hospital site. The same order in which the hospitals were originally visited was maintained for this stage of the study. See Appendix X for details on dates and contacts.

The process used to re-code and re-abstract the data in this step was similar to that used for the first step; however, because there were only ten records per facility, the assistant was not hired to assist in the preparation of records. The hospital clerks removed the codes and photocopied the face sheets and the investigator ensured that the sample was correct. The hospital questionnaire was not repeated and no reconciliation process was deemed necessary for this step.

The investigator then re-coded and re-abstracted the records. Yellow-coloured re-abstracting forms were used for this sample of cases to ensure this subsample was kept separate from the original sample and not analyzed twice during the data analysis.
Individual data items for each case, taken from the first set of re-abstracted data, were entered on the yellow forms by the assistant at the completion of all six hospital sites.

2.10.10 Computer Input

Before the data were ready for computer input, a more detailed review of the type of discrepancy category was required, especially regarding the secondary diagnoses and secondary procedures.

A microcomputer program, using *Epi Info, Version 5*, was developed for inputting of data and transformation into a language readable by SAS statistical analysis package. Epi Info is a program that handles epidemiologic data in questionnaire format and organizes study designs into text that can be converted into the SAS language. This process was done by the p.i.. The data were carefully input and checked for errors, and obvious errors were corrected. The data were then explored through the statistical analysis package for additional errors and these were also corrected before the preliminary analysis of the data occurred.

2.10.11 Statistical Data Analysis

Analysis using SAS statistical analysis package on a personal computer, was done to determine whether the re-abstracted codes agreed with the original hospital codes as represented in the MOH database. The analysis, done by the p.i., consisted mainly of frequency counts of discrepancies for the MRD, PP, Secondary Diagnoses, Secondary Procedures, Non-Diagnostics, and Combined Variables. Chi-squared and p-values were computed for the Most Responsible Diagnosis. Pearson correlation coefficient was
computed for coder scores. A ranking system was established for birthdate discrepancies, but because the data showed high agreement for admission dates, discharge dates, discharge status and sex, no ranking system was established for those variables.

Cohen's Kappa statistic was to be used on the MRD and PP variables to measure concordance between the coders. The Kappa statistic corrects for agreement expected by chance (Cohen, 1968). However, in the ICD Classification system, there are over ten thousand possible choices. Because of the large number of categories in the ICD, the possibility of an agreement occurring by chance is virtually zero; therefore Kappa score was not used.

Secondary diagnoses were examined to determine how frequently the MRD was coded not as most responsible, but as a secondary diagnosis. Secondary procedures were examined to determine how often a principal procedure was recorded as secondary, rather than as principal. Types of discrepancy and coder agreement scores were also examined over all the hospitals and between the hospitals for comparison purposes. A correlation was done between coders, credentials and years of experience.

Each hospital record has one and only one MRD. However, a case may have anywhere from zero to fifteen secondary diagnoses. For the purpose of analyzing secondary diagnoses, the denominators have been changed to reflect the actual number of occurrences of recorded secondary diagnoses. Not all patients had a procedure. Therefore, not all cases in the study have a PP. When analyzing the PP category, the denominator was changed to reflect only the number of cases that had procedures. If an individual case did have a PP, it could also have had anywhere from one to thirteen secondary procedures. Secondary procedures were analyzed by the number of coded
secondary procedures in that category and denominators have been changed from the number of cases with a PP to the number of occurrences of recorded secondary procedures.

A descriptive summary of the hospital questionnaire was undertaken to determine correlation with the discrepancies. Because the questionnaire was undertaken for each hospital, not each individual case, a regression analysis was not done. From the combined analysis, the p.i. attempted to answer the question, "is there a need for improvement?".
2.11 PROJECT TEAM

Although there were three key individuals involved in this study, the p.i., the assistant and the computer analyst, the hospital personnel from each of the individual hospital sites were also instrumental in helping to complete this project. This section presents a brief job description for personnel directly involved in the study.

2.11.1 Principal Investigator

The principal investigator was responsible for the planning, developing and coordinating of the entire project. She also conducted the actual coding and re-abstracting of the records, completed the final analysis, and recorded the results.

2.11.2 Assistant

The assistant was hired to ensure that the p.i. was blinded to the original diagnostic and procedural codes on the face sheet and in the computerized MOH database. She determined the correct selection of records by comparing the charts pulled by each individual hospital to the MOH list of selected cases prepared by the computer analyst. Records were sorted in numerical order so that the p.i. could not determine which diagnostic category the cases originated from. The assistant photocopied the original face sheet to ensure all previous codes were removed before the p.i. re-coded and re-abstracted the records.

The assistant filed the photocopied face sheet on the chart and gave the original face sheets to the coding supervisor until the completion of the re-abstracting stage of the
study. Before the re-abstracting was completed, the assistant ensured that all cases in each category were accounted for.

At the completion of the re-abstracting stage, the assistant re-filed the original face sheets on the records. She then copied the codes from the original face sheet onto the re-abstracting form.

She then opened the documents previously sealed by the computer analyst with the MOH listing of the individual data items (see section 2.11.3 for details). These items were recorded onto the re-abstracting form to enable comparison of the re-abstracted codes. The charts were again checked to ensure they were in numerical order. Each hospital elected to re-file its own charts.

For the IRR stage of the study and at the completion of all six hospital sites, the assistant entered the individual data items for each case, taken from the first set of re-abstracted data, onto the IRR re-abstracting forms for comparison with the first set of re-abstracted data.

2.11.3 Computer Analyst

A computer analyst extracted the random sample of records from the B.C. MOH computer tapes and produced six separate listings of selected cases, one for each individual hospital, by hospital record number stratified by diagnostic category. See Appendix E for details. Alternate records were produced within each of the diagnostic categories (see sampling plan, section 2.6).

The analyst extracted the individual data items from the MOH database for each case and placed this information in sealed envelopes for each hospital. The envelopes
were opened only after all the cases were re-coded and re-abstracted for each hospital. This process was necessary for future comparison with the re-abstracted data during the reconciliation process.

The analyst also produced an MOH listing of selected cases for the 10% subsample used for intra-rater reliability. A 10% random sample of each hospital’s cases was selected from the previous sample. This subsample was not stratified by diagnostic category.

2.11.4 Hospital Personnel

Hospital personnel assisted in many ways, but their main purpose was to supply the health records and re-file them at the completion of the study. For the intra-rater subsample, the health record clerks photocopied the face sheets to ensure that codes were removed for the p.i..
2.12 TYPE OF DISCREPANCY CATEGORIES

This study incorporates input and instrument reliability through the type of discrepancy category. Four major types of discrepancy categories were determined from the literature: System, Clerical, Documentation and Other. For this study, a type of discrepancy is attached to an individual code, not to a case. Each discrepant code is assigned only one type of discrepancy. Since each case can have several discrepant codes, it follows that each case may have several types of discrepancies determined.

When several types of discrepancy are deemed possible for one specific code, the one type that seemed the most likely was chosen. Discrepancy types were determined by the p.i.. Agreement with the supervisor at each hospital was sought to ensure this process was as accurate as possible.

2.12.1 System Discrepancies

These types of discrepancies are derived from incorrectly or improperly classifying coded information. These discrepancies incorporate Judgement, Classification and Procedural subcategories.

2.12.1.1 Judgement

A judgement discrepancy occurs when the classification system does not have adequate guidelines on which diagnosis or procedure to choose, leaving the choice to the judgement of the abstractor. The documentation in the health record does not clearly determine which diagnosis or procedure to choose as the MRD or PP respectively. Judgement discrepancies are derived from having to choose between two diagnoses being
the MRD or between two procedures being the PP. An example would be where a coder has to choose between two diagnoses: one diagnosis when the patient is admitted for surgery, or the complication that extends the length of stay by the same length of stay as the condition that necessitated surgery.

2.12.1.2 Classification

Classification discrepancies can occur when there is not a word-for-word correspondence between the Alphabetical Index of the ICD manuals and the terminology used in the documentation in the health record. An example is the diagnosis of recurrent tonsillitis when only acute or chronic tonsillitis is found in the manuals. The coder must determine which diagnosis to code. In this category, the coder may have an inadequate understanding of the nature of the condition to be classified, or the nature of the classification system itself, or both. There may be two codes in the manual equally able to capture the diagnosis or procedure.

2.12.1.3 Procedural

Procedural discrepancies occur when hospital policy is contrary to HMRI/CIHI guidelines or when mandatory coding is not exercised. Rules are not actively followed as set forth by Statistics Canada or HMRI/CIHI. Coding books are not kept up to date.

2.12.2 Clerical Discrepancies

The second major type of discrepancy is a 'clerical' discrepancy; a discrepancy that is the result of human carelessness. This category includes erroneously recorded
information or misinterpreting the information within the record. Clerical errors may be generated in several ways: data conversion, charting, and coding/abstracting errors.

2.12.2.1 Data Conversion

Examples of data conversion errors include misreading of a word, incorrectly transcribing a code, transposing a code number, or duplicating coding for the same procedure or diagnosis. A coder may code from memory, where over time the memory deteriorates and an incorrect code can often be used.

2.12.2.2 Charting Errors

A complete review of the entire health record may not have been done. All relevant information in the chart is not captured. An example within this category is that the coder either codes from an incomplete chart, i.e., the face sheet, or does not review the chart for specificity, i.e., in order to choose a fourth digit. An example can occur when the specificity of a code is required, the information is available in the record, but specificity is not determined by the coder.

2.12.2.3 Coding/Abstracting Discrepancies

Reliance on the Alphabetic Index with no reference to the Tabular Lists leads to failure to use the inclusions and exclusions noted in the Tabular List. Another can occur when code books are not used properly to determine specificity of the diagnosis, i.e., mandatory coding is not used.
2.12.3 Documentation Discrepancies

The third major type of discrepancy arises from the inaccuracy of the ‘documentation’ in the record itself. Documentation discrepancies are of two types: incomplete records and discrepant records.

2.12.3.1 Incomplete Record

Health records that are incomplete at the time the patient’s record is coded are one source of coding discrepancy. An incomplete record exists when all the pertinent information on a patient for the specific admission is not available. An incomplete record is also a result of inadequate or insufficient information in the record to determine a code.

2.12.3.2 Discrepant Records

These discrepancies involve incongruence in the health records where there is inadequate information to determine a code or where the information within the record is inconsistent or discrepant between recorders. There are discrepancies within the health record arising from inconsistent information among the recorders or inadequate information on which to determine a code. An example is a contradiction between the recorded diagnosis on the face sheet and the investigative reports. Another occurs when two consultants, each diagnosing the same disease entity, label a disease category differently. Who is correct? More than one description of a disease or procedure in the medical record can lead to more than one code in the classification system for the same disease or procedural entity. The physician has determined a diagnosis, but there is clear indication it is in error.
2.12.4 Other Discrepancies

Discrepancies that do not fall in any of the above categories fall into this category. Other discrepancies involve any explanation other than the previously defined six sub-categories. This category includes those discrepancies where a type could not be determined.

2.12.5 Extra .0 in MOH Data

There is an extra ‘.0’ in the MOH data where only a three-digit code is required. This was added for diagnostic codes only during the HMRI/CIHI conversion of the data for the MOH. HMRI/CIHI converts ICD-9CM codes to ICD-9 codes before sending to the appropriate Ministries/Departments of Health. In that process, fifth digits are severed and are not sent to the MOH. Fifth digits are only retained for individual hospital purposes. In that conversion process, some three-digit diagnostic codes that did not require fourth digits were changed to add a decimal point and a zero (.0). That is, to a diagnosis, such as acute myocardial infarction with an ICD-9 three-digit code of 410, the resultant four-digit code would be 410.0. This was done with diagnostic codes only.

Although this may seem insignificant, the accumulated totals in the B.C. MOH’s database in two separate categories were 410 (1,159 cases) and 410.0 (4,563) for the code of 410 (Province of B.C., 1987). It is possible that all cases would not be included in some research projects. All unmatched data, when this discrepancy was determined (99 cases for MRD and 160 cases for secondary diagnoses), were included in the data analysis as the four-digit, adjusted agreement scores.
2.13 CONFIDENTIALITY

Review of the health record necessitates protection of individual patient privacy. The HMRI/CIHI database identifies the patient, the hospital, and the coder by an identification number. Because individual identifying data were not required for this analysis, the patient's identification number was used only to identify the record for comparison purposes.

Confidentiality of the patient was assured by using only the data required for this comparison, such as admission date, separation date, sex, birthdate, diagnoses, and procedures, on the re-abstracting form. There was no reference to the name of the patient, the geographic location of the patient, or the attending physician. Hospitals are identified only by letters A through F.

The coder identification number was used to determine the credentials, qualifications and experience of the coder. This information was necessary to determine the frequency of coder discrepancy and whether coder training and/or years of experience had any influence on the number of discrepancies. Coder names had to be used initially to determine the pertinent information on the questionnaire, but individual identifying information was destroyed prior to the data analysis. The coder is identified in the final dataset only by the assigned computer number, one through thirty-eight.
2.14 SIGNIFICANCE OF FINDINGS

This study provides information on the reliability of the HMRI/CIHI database within selected hospitals in Vancouver. The study will also help to delineate the limitations of the database when utilized for clinical and research purposes beyond those of the clinical site.

There needs to be an awareness of individual hospital performance when collective data are utilized for research and planning purposes. Having an awareness only of national and provincial guidelines is not sufficient for utilizing aggregate data. A thorough knowledge of individual hospital policies and procedures is also necessary.

The recommendations from this study should help to improve the coding for individual hospitals, for the aggregate HMRI/CIHI database and ultimately for the B.C. MOH database. Improvements in coding practices should lead to a more accurate and consistent database.
2.15 DIFFICULTIES ENCOUNTERED

There were several areas within the study where difficulties were encountered. This section documents the problems with upgrading diagnostic and procedural code books, selection of cases, wrong selection of records, tracking of records, and the hospital questionnaire.

2.15.1 Updating Diagnostic and Procedural Code Books

There were several errata up to the fiscal year 1986/87. Numerous changes were necessary. This process was long and tedious and took several weeks to complete. The number of changes and the inserts into the code books produced reference books with several additions and deletions. Some of these additions and deletions were difficult to add to the manuals due to the extent of the changes. This process was considered an area where problems could be created for the coding process, which could ultimately affect the national database. Changes to the classification system were generated over a span of ten years (1977-1986).

2.15.2 Selection of Cases

A problem arose with the first hospital’s sample when it was determined that long term care and day surgical patients were selected along with the acute care cases. The sample was re-drawn to include only acute care patients.
2.15.3 Tracking of Records

Throughout the period of the study charts disappeared and reappeared as needed by the hospital for patient readmissions, thus making it difficult to keep accurate track of them. Missing records sometimes reappeared at night without the knowledge of the research team. Before the end of the re-abstracting process, the assistant ensured that all records in each category were accounted for. Alternate charts were requested until the required number of cases in each category were coded. It is noted that there are several hospitals that have more than one hundred re-abstracted cases. When alternate records were requested, they were coded along with the records that reappeared without noticing. Thus, some categories have more cases than necessary. There was no way to determine whether the content of the chart changed while the investigator was working on it. There was also no way to determine whether the chart changed between the initial re-abstracting and the 10% subsample re-abstracting. It was assumed that changes would not have been made to the particular hospital episode from 1986/87.

2.15.4 Wrong Selection of Records

Eighteen records at one hospital had to be re-coded because the wrong records had been selected by the assistant. The hospital selected the right patient, but the assistant selected the wrong chart for the study. This hospital's face sheets had the previous admission date recorded above the separation date - an unusual way of documenting these. This resulted in the wrong face sheet being photocopied and therefore, the wrong record being coded in the study. The error was found when it was determined that a separation date on one re-abstracted health record was external to the fiscal year of the study. The
eighteen cases external to the fiscal year of the study were removed from the study.

Eighteen new charts from the list of alternates were prepared by the assistant and these records were re-abstracted by the p.i..

2.15.5 Hospital Questionnaire

Obtaining details on coder training was a time-consuming process because although most of the information could be attained in one sitting, details of coders training, experience, and number of workshops attended needed to be determined for staff who were currently not employed with the hospital. The study was based on a fiscal year a few years earlier: some staffing had changed and old personnel records had to be reviewed. All necessary details were obtained. Details on training were obtained several weeks after the initial interview. This was consistent across hospitals.
2.16 LIMITATIONS OF THE STUDY

There were limitations arising from conditions between the original coder and the principal investigator and with the ICD-9 classification system.

2.16.1 Similar Conditions

Previous hospital records were made available to the principal investigator. If a patient had been previously admitted for the same condition, this code may have appeared on the previous record. It is noted, however, that the previous record was accessed when a documentation problem in the ‘current’ record needed verification. Every attempt was made to disregard codes from old records, a similar condition for hospital coders.

The time between that allocated by the p.i. and that allocated to code a record by the hospital staff differed. There were several records that took the principal investigator several hours to code. With quotas of 25 charts or more per day requested by each hospital in the study, the hospital staff could not have taken the same time to review these records in such detail. However, even though the p.i. took several hours to go through nursing notes, etc., the relevance of gathering data from those documents is debatable. Although there were very few codes gathered from these excessive documents, such as nursing notes, complications such as a patient injury resulting from a fall out of bed may only be found by examining these documents in detail.

2.16.2 The ICD-9 Classification System

There are several limitations associated with the ICD-9 Classification System.
2.16.2.1 Subjectivity

It is assumed that the database has face validity. However, it is known that there are inadequacies within the current ICD-9 nomenclature that often lead to differences in the exercise of individual judgement within individual diagnostic categories. The subjectivity of the ICD-9 coding system inevitably leads to inter-rater differences in the coding of data. An example is the interpretation of codes where a patient has asthmatic bronchitis: it is often difficult to determine from the health record whether the code should be obstructive chronic bronchitis, asthmatic (491.2) or asthma, unspecified, with bronchitis (493.9).

2.16.2.1.1 Subjectivity of the MRD

It is not always clear from the health record which diagnosis is the MRD. Care must be taken in determining the MRD and in selecting the associated alphanumeric code. The symptomatology and severity of conditions can often be interwoven, making it difficult, or even impossible, to choose one condition that must be regarded as that which describes the most significant condition causing the greatest length of stay (IOM, 1980). The MRD by HMRI/CIHI definition is that which describes the most significant condition of a patient which causes his stay in hospital (See HMRI Guidelines in Appendix Z). In a case where multiple diagnoses may be classified as most responsible, the diagnosis responsible for the greatest length of stay is the MRD.

The process for the selection of the MRD is not as straightforward as the definition makes it appear. Determining which diagnosis caused the greatest length of stay can be very subjective if the information in the chart is not documented clearly. The
interpretation of the chart does not always lend itself to a clear-cut selection of the MRD. For instance, one patient may have severe coronary artery disease complicated by an acute myocardial infarction while in hospital. It may be very difficult to determine which condition produced the greatest length of stay when both diagnoses can be equally instrumental in the overall hospital stay. Similarly, it may be unclear as to which diagnosis to code when a complication extending the patient’s stay is as important as the original reason for that hospital stay. That is, if the average length of stay for a hernia operation is three days, but a particular case is complicated by a wound infection that extends the stay another three days, should the infection or the hernia be coded as the MRD? If the conditions are of equal importance, should the condition relevant to diagnosis on admission be recorded as the MRD?

2.16.2.1.2 Subjectivity of Procedures

Not all the procedural categories in the classification system are self-explanatory. For instance, cancelled surgery and postponed surgery for a few days are not differentiated in the CCP code, 99.08. Does the patient have to be taken to the operating room before it is considered cancelled? If not, and the patient’s surgery was cancelled, but re-scheduled and carried out before discharge, would this code still be relevant? There are no guidelines to determine this and hospital practices differ in interpreting how this code is used.
2.16.2.2 Lack of Severity Measures

As stated in the literature, the ICD-9 classification system only indicates that a disease exists. It does not include the extent or the severity of that disease. One patient's disease may not be as severe as another, yet the same code is produced for each patient. HMRI/CIHI uses the RIW classification to address this issue.

2.16.3 Availability of Relevant Reports

It has been suggested that the coder ensure all relevant reports are available before proceeding to the coding process. Some hospitals were coding within a week or two after discharge, which according to most hospital policy is the desired time frame. However, many physicians do not dictate reports within this time frame. Several hospital supervisors suggested that several charts get coded without even having a discharge summary in the chart. Hospitals are currently coding from records that are missing relevant reports. All hospitals in the study reported that 10 to 90% of the coding is done from incomplete records.
CHAPTER 3 - RESULTS OF DATA ANALYSIS

3.0 INTRODUCTION

This chapter presents the results of the data analysis. It first examines the agreement between the principal investigator (p.i.) and the coding supervisor (supervisor). It then analyses coding discrepancies in relation to the following areas:

- Most Responsible Diagnosis (MRD),
- Principal Procedure (PP),
- Secondary Diagnoses,
- Secondary Procedures,
- Non-Diagnostic Variables, and
- Combined Variables.

Each of the above areas is subdivided to obtain frequencies of agreement scores over all the hospitals combined and for each individual hospital. Secondary diagnosis and secondary procedure categories are analyzed by the number of diagnoses and/or procedures recorded and by number of cases.

Diagnostic groups are examined by the frequency of agreement with the Most Responsible Diagnosis and by the Principal Procedure. Both the samples drawn for this study (section 2.5) have been combined to examine inter-diagnostic comparison.

Type of discrepancy is examined to obtain frequency of the type of discrepancy for each of the above coding areas for the hospitals combined and for individual hospitals.
Diagnostic variables and procedural variables are combined to determine the type of discrepancy by coding area.

"Coders" are examined for the frequency of agreement over all the hospitals and within the hospitals. The hospital questionnaire is analyzed to determine if there was an influence of hospital characteristics on the frequency of discrepancy and to determine the credentials and years of experience for each coder. Another area, "intra-rater reliability" is also assessed.
3.1 AGREEMENT BETWEEN PRINCIPAL INVESTIGATOR AND CODING SUPERVISOR

Agreement between the p.i. and the supervisor from each hospital was documented during the reconciliation process. Where a discrepancy between the re-abstracted and the Hospital Medical Records Institute/B.C. Ministry of Health (HMRI/MOH) data existed, both the p.i. and the supervisor reviewed the record to determine general agreement on which codes were correct and the type of each discrepancy. The objective was to determine whether the HMRI/MOH codes or the re-abstracted codes were correct. When agreement was not reached on at least one variable within the record, the entire case was deemed not to agree. An in-depth review of the MRD and the PP variables was completed, while only a cursory review of the secondary diagnoses and secondary procedures was undertaken.

TABLE 1 - AGREEMENT ON CORRECT CODES BETWEEN PRINCIPAL INVESTIGATOR AND CODING SUPERVISOR BY NUMBER OF CASES FOR INDIVIDUAL HOSPITALS
Re-abstracted Data Apr 1/86 - Mar 31/87

<table>
<thead>
<tr>
<th>HOSP #</th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>E</th>
<th>F</th>
<th>TOTAL</th>
</tr>
</thead>
<tbody>
<tr>
<td># CASES</td>
<td>103</td>
<td>100</td>
<td>101</td>
<td>102</td>
<td>100</td>
<td>100</td>
<td>606</td>
</tr>
<tr>
<td>DO NOT AGREE</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>0</td>
<td>4</td>
<td>1</td>
<td>7</td>
</tr>
<tr>
<td>AGREE</td>
<td>103</td>
<td>100</td>
<td>99</td>
<td>102</td>
<td>96</td>
<td>99</td>
<td>599</td>
</tr>
<tr>
<td>PERCENT</td>
<td>100.0</td>
<td>100.0</td>
<td>98.0</td>
<td>100.0</td>
<td>96.0</td>
<td>99.0</td>
<td>98.8</td>
</tr>
</tbody>
</table>
Agreement on codes between the p.i. and the supervisor was determined to be 98.8% of cases over all the hospitals, which supports the validity of the data used in the study. The seven cases without agreement were retained in the analysis of the data. See Table 1 for details.
3.2 MOST RESPONSIBLE DIAGNOSIS

The Most Responsible Diagnosis (MRD) is that one diagnosis which describes the condition of a patient which most contributes to his or her length of stay in hospital (HMRI, 1990). One and only one MRD is recorded for each patient at discharge.

A match between the two data sets is determined when the original code, as determined through the HMRI/MOH database matches the re-abstracted code. For this analysis, the codes must match in every digit to the four-digit level. However, the ICD-9 classification system incorporates complete three-digit codes where no fourth-digit is required, such as acute myocardial infarction, coded as 410. This code does not require further specificity within the confines of ICD-9. Therefore, a diagnostic match at the four-digit level includes those complete three-digit codes that do not require further specificity. A three-digit match is examined only for those codes where a fourth digit is necessary, but where only the first three digits of the codes match.

A discrepancy in the MRD category is important because this is the variable most often used for research purposes, planning and allocation of resources. It is the foundation for the assignment of CMGs. This is especially important for the future funding of health care utilizing the HMRI (CIHI) database.

3.2.1 Frequency of Match - All Hospitals

Table 2 demonstrates that over all the hospitals, the agreement score for the MRD at the four-digit level, raw score, and before any adjustments, is 45.1% of cases. It became apparent during the reconciliation stage that some of the original codes on the face sheets did not match the HMRI/MOH data codes. The original computerized data records
were checked and found to be the same as those on the face sheets. HMRI converts the data from ICD-9CM to ICD-9 format to present standardized data for the MOH. This means that the re-abstracted data would have agreed the data that had been submitted to HMRI, but the codes were changed during the conversion process. These codes were given a unique type of discrepancy code, i.e. type of discrepancy = eight, to distinguish them from other errors. When the data were adjusted to include the data (99 cases) changed by HMRI in the transformation of the data for the B.C. MOH, the overall agreement rate for the MRD at the four-digit, adjusted level is 61.4% of cases.

**TABLE 2 - FREQUENCY OF MATCH FOR MOST RESPONSIBLE DIAGNOSIS BY NUMBER OF CASES OVER ALL HOSPITALS**

Re-abstracted Data Apr 1/86 - Mar 31/87

<table>
<thead>
<tr>
<th>LEVEL OF MATCH</th>
<th>FREQUENCY</th>
<th>PERCENT</th>
<th>CUM PERCENT</th>
</tr>
</thead>
<tbody>
<tr>
<td>4-digit (Raw Score)</td>
<td>273</td>
<td>45.1</td>
<td>45.1</td>
</tr>
<tr>
<td>4-digit (Adj.*)</td>
<td>99</td>
<td>16.3</td>
<td>61.4</td>
</tr>
<tr>
<td>3-digit</td>
<td>65</td>
<td>10.7</td>
<td>72.1</td>
</tr>
<tr>
<td>Not Matched</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>- as MRD, 4-digit match, but listed as 'secondary'</td>
<td>51</td>
<td>8.4</td>
<td>80.5</td>
</tr>
<tr>
<td>- not matched at all</td>
<td>118</td>
<td>19.5</td>
<td>100.0</td>
</tr>
<tr>
<td>Valid Cases</td>
<td>606</td>
<td>100.0</td>
<td></td>
</tr>
</tbody>
</table>

* Adjusted for converted HMRI data (99 cases)

At the three-digit level over all the hospitals, the agreement increases 10.7% from the four-digit (adjusted) level for a cumulative match to 72.1% of cases. Thus, 72% of cases match to at least three-digits.
There are an additional 8.4% of cases where the codes agree to four-digits, but the four-digit code recorded by the p.i. as the MRD was listed as a secondary diagnosis by the original hospital coder. For these cases the diagnosis is recorded correctly, but a discrepancy exists in the selection of the MRD.

The remaining 19.5% of the MRDs coded by the original hospital coder disagree with codes in the re-abstracted data, both in the MRD and secondary diagnoses. In addition, there were 2,500 extra secondary diagnoses recorded by the p.i. which were not recorded by the original hospital coder. Codes were examined for agreement only if they were recorded by the original coder. Thus, diagnoses in the "not matched" category are those coded by the original coder where there was no match in the re-abstracted data. The term "overcoding" may be applied here (section 1.8.3.2 discusses overcoding). Most health record coders in practice do not code every diagnosis in the chart. However, even though the codes in this study were not identified as miscodes, they are not all "overcoded". Several of these codes should have been coded by the original hospital coder. The impact of not identifying these as miscodes could seriously understate the actual number of discrepancies within the database.

The level of coding agreement for the Most Responsible Diagnosis between the HMRI (CIHI) database and the re-abstracted data is less than 90%. Therefore, the null hypothesis is rejected.

3.2.2 Range of Match - Individual Hospitals

Table 3 demonstrates that agreement on the MRD for individual hospitals at the four-digit level, raw score, ranges between 32.7 and 55.9%. At the four-digit, adjusted
level, which includes the converted HMRI data, the rate increases in every hospital to between 52.0 and 69.6%. At the three-digit level, the match increases from the four-digit (adjusted) level in every hospital to range between 62.4% for Hospital C and 79.4% for Hospital D. The increase in agreement from the four-digit, adjusted to the three-digit level ranges from 6.9% for Hospital C to 14.0% for Hospital E. The chi-squares demonstrated a statistically significant difference among the observed hospital scores, both at the 4- and 3-digit level.

<table>
<thead>
<tr>
<th>HOSP</th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>E</th>
<th>F</th>
</tr>
</thead>
<tbody>
<tr>
<td># CASES</td>
<td>103</td>
<td>100</td>
<td>101</td>
<td>102</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>MATCHES</td>
<td>#</td>
<td>%</td>
<td>#</td>
<td>%</td>
<td>#</td>
<td>%</td>
</tr>
<tr>
<td>4-DIGIT, RAW SCORE</td>
<td>43</td>
<td>41.8</td>
<td>49</td>
<td>49.0</td>
<td>33</td>
<td>32.7</td>
</tr>
<tr>
<td>4-DIGIT, ADJ.*</td>
<td>70</td>
<td>68.0</td>
<td>58</td>
<td>58.0</td>
<td>56</td>
<td>55.5</td>
</tr>
<tr>
<td>3-DIGIT</td>
<td>81</td>
<td>78.6</td>
<td>71</td>
<td>71.0</td>
<td>63</td>
<td>62.4</td>
</tr>
</tbody>
</table>

*Adjusted for converted HMRI data (99 cases)
According to the hospital questionnaire, noted in section 2.9.2 and analyzed in section 3.11, Hospital B coded less than 10% from incomplete charts, Hospital C coded from 20%, Hospital D from 25%, Hospital A from 30%, Hospital E from 70% and Hospital F from over 90% during fiscal year 1986/87. However, no significance can be derived from this. All hospitals, except Hospital E, routinely coded the first listed diagnosis on the face sheet as the MRD. It would appear from this information that Hospital E should have the highest agreement in this category, however, the results show the opposite.
3.3 PRINCIPAL PROCEDURE

The denominator for the Principal Procedure (PP) category reflects only those cases where procedures are recorded. Not all hospital cases include a codable procedure. Since there are 109 cases in which no procedures are recorded, the total number of cases with procedures is 497. There were no changes to PP codes during the HMRI conversion of data; therefore, no adjustment for this process is necessary.

The PP category incorporates non-operative procedures, such as x-rays and ECG's. With procedures such as these, it is difficult to determine whether one procedure is more important than another. When non-operative procedures, such as ECG and chest x-rays, are recorded by both the original hospital coder and the p.i., and where it cannot be determined which procedure is the principal one, the records are assumed to be coded correctly, despite the order in which the procedure is recorded. Non-operative procedures were recorded in the PP position when no other procedure was relevant.

Procedures are coded using a two-digit category code followed by a third and sometimes a fourth digit to determine further specificity. In the MRD diagnostic categorization, three-digit codes such as acute myocardial infarction (code 410) appear as a complete match at the fourth-digit level. Unlike the diagnostic classification, the procedural classification of the CCP has no complete two-digit procedural codes. All procedural codes require at least one digit beyond the two-digit category. Therefore, unlike the MRD analysis, matches in the PP category all necessitate a match to the three-digit level. There are very few four-digit procedural codes. However, where a fourth code is required, to be in agreement that digit must also match. Since this study only analyzed matches to the third digit, all fourth-digit matches are captured within the three-
digit level of match. Because all procedural codes require three digits, within the agreement at the two-digit level, matches were only determined for codes where the first two digits agreed, but the third did not. Such 'matches' are considered close, but not perfect.

3.3.1 Frequency of Match - All Hospitals

Table 4 demonstrates the agreement score over all the hospitals in the PP category at the three-digit level (and if required, the fourth digit level) to be 64.8% of cases. At the two-digit level, the cumulative agreement increases by 8.1% to 72.9% of cases over all the hospitals. Those cases where the PP coded by the p.i. does not match the PP by the original coder, but is listed as a secondary procedure in the hospital coder's data,

<table>
<thead>
<tr>
<th>LEVEL OF MATCH</th>
<th>FREQUENCY</th>
<th>PERCENT</th>
<th>CUM PERCENT</th>
</tr>
</thead>
<tbody>
<tr>
<td>Full 3- or 4-digit match</td>
<td>322</td>
<td>64.8</td>
<td>64.8</td>
</tr>
<tr>
<td>2-digit match</td>
<td>40</td>
<td>8.1</td>
<td>72.9</td>
</tr>
<tr>
<td>Not matched</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>- as PP, 3-digit match, but</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>listed as 'secondary'</td>
<td>87</td>
<td>17.5</td>
<td>90.4</td>
</tr>
<tr>
<td>- not matched at all</td>
<td>48</td>
<td>9.7</td>
<td>100.0</td>
</tr>
<tr>
<td>Valid Cases*</td>
<td>497</td>
<td>100.0</td>
<td></td>
</tr>
</tbody>
</table>

* The number of cases is adjusted to reflect only those cases where a procedure was documented by the original hospital coder.
represented an additional 17.5% of cases. This demonstrates that PP's are frequently not recorded as the principal procedure. In 9.7% of cases, there is no match whatsoever.

The level of coding agreement for the Principal Procedure between the HMRI (CIHI) database and the re-abstracted data is less than 90%. Therefore, the null hypothesis is rejected.

3.3.2 Range of Match - Individual Hospitals

Table 5 shows the PP match rate at the three-digit level by individual hospital. This match rate ranged between 56.3% for Hospital C and 75.6% for Hospital D. At the two-digit level, the rate of match increased in every hospital to range between 61.4% for Hospital F and 85.4% for Hospital B. The increase in agreement per hospital from the three-digit level to the two-digit level ranges between 4.6% for Hospital F and 12.3% for Hospital A.

<table>
<thead>
<tr>
<th>HOSP</th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>E</th>
<th>F</th>
</tr>
</thead>
<tbody>
<tr>
<td>n=</td>
<td>65</td>
<td>89</td>
<td>87</td>
<td>78</td>
<td>90</td>
<td>88</td>
</tr>
<tr>
<td>MATCHED</td>
<td>#</td>
<td>%</td>
<td>#</td>
<td>%</td>
<td>#</td>
<td>%</td>
</tr>
<tr>
<td>3-DIGIT</td>
<td>44</td>
<td>67.7</td>
<td>67</td>
<td>75.3</td>
<td>49</td>
<td>56.3</td>
</tr>
<tr>
<td>2-DIGIT</td>
<td>52</td>
<td>80.0</td>
<td>76</td>
<td>85.4</td>
<td>57</td>
<td>65.5</td>
</tr>
</tbody>
</table>
All hospitals claim to code non-operative procedures. Hospitals A to E coded all non-operative procedures, while Hospital F coded only selected ones. Hospitals who code non-operative procedures are more likely to get them into the PP field.
3.4 SECONDARY DIAGNOSES

Although each case has one and only one MRD, each can also have from zero to fifteen recorded secondary diagnoses. Agreement for secondary diagnoses can be approached either using all recorded secondary diagnoses as the denominator, or using number of cases as the denominator. In the former, the denominator is the number of secondary diagnostic codes that appeared in the original record. In both cases, secondary diagnoses are matched to the four-digit level only. Secondary diagnoses are considered in agreement when the secondary diagnosis coded by the original hospital coder matches a re-abstracted code, irrespective of position. If a secondary diagnosis of the original hospital coder matches the MRD of the p.i., it is considered a match within this category.

There were 2,500 secondary diagnoses in the 606 cases, which were coded by the p.i. but which do not appear in the MOH records. On average, there are 4.1 more diagnoses per record recorded by the p.i. The match rate for secondary diagnoses does not include these diagnoses. There is a lot of missing diagnostic information. This under-recording of secondary diagnoses by the hospital coders is consistent over all the hospitals in the study. (See notes on overcoding in section 1.8.3.2.)

3.4.1 By Number of Secondary Diagnoses Recorded

The denominator for secondary diagnoses is determined by the number of secondary diagnoses actually recorded by the original hospital coder. The numerator reflects the number of secondary diagnoses actually recorded by the original coder that match the diagnostic codes to the four-digit level, whether coded as MRD or secondary, by the p.i..
3.4.1.1 Frequency of Match - Four-digit, Raw Score

<table>
<thead>
<tr>
<th>HOSP</th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>E</th>
<th>F</th>
<th>TOTAL</th>
</tr>
</thead>
<tbody>
<tr>
<td>MATCHED</td>
<td>64</td>
<td>147</td>
<td>137</td>
<td>175</td>
<td>184</td>
<td>124</td>
<td>831</td>
</tr>
<tr>
<td>TOTAL # CODES</td>
<td>147</td>
<td>260</td>
<td>227</td>
<td>285</td>
<td>320</td>
<td>231</td>
<td>1470</td>
</tr>
<tr>
<td>PERCENT</td>
<td>43.5</td>
<td>56.5</td>
<td>60.4</td>
<td>61.4</td>
<td>57.5</td>
<td>53.7</td>
<td>56.5</td>
</tr>
</tbody>
</table>

Table 6 demonstrates the match rate for secondary diagnoses at the four-digit level, raw score, over all the hospitals to be 56.5% by number of secondary diagnoses recorded by the original coder. Table 6 also demonstrates the range for the secondary diagnoses at the four-digit level, raw score, to fall between 43.5% for Hospital A and 61.4% for Hospital D. See Table 6 for details on individual hospital rates.

3.4.1.2 Frequency of Match - Four-digit, Adjusted Level

When the secondary diagnostic data are adjusted for the converted HMRI data (160 cases), the frequency of match at the four-digit level increases to 67.4%. At the four-digit adjusted level, the range falls between 52.4% for Hospital A and 72.3% for Hospital C. See Table 7 for details on the adjusted data.
3.4.2 By Number of Cases

When match rates for secondary diagnoses are computed using number of cases as the denominator, the match rate drops dramatically. Codes are matched to the four-digit, adjusted level for secondary diagnoses only. For a case to be discrepant, only one code needs to disagree. For a code to be discrepant, only one secondary diagnostic digit needs to disagree. If a case has no recorded secondary diagnoses, it is determined to agree.

Those additional secondary diagnoses recorded by the p.i. and not the original coder were not included in this analysis. Table 8 demonstrates that only 34.5% of cases agree in the secondary diagnoses category. The range of match by secondary diagnoses is from 20.0% for Hospital E to 47.6% of cases for Hospital A. According to the questionnaire, quotas
required per coder ranged from 16 to 40 charts per day. Hospital F required 16, Hospitals B & E required 25, Hospital D 30 and Hospitals A & C 40. No correlation between discrepancy rates and quotas could be determined.

TABLE 8 - FREQUENCY OF MATCH FOR SECONDARY DIAGNOSES 4-DIGIT LEVEL, ADJUSTED* BY NUMBER OF CASES
Re-abstracted Data Apr 1/86 - Mar 31/87

<table>
<thead>
<tr>
<th>HOSP</th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>E</th>
<th>F</th>
<th>TOTAL</th>
</tr>
</thead>
<tbody>
<tr>
<td>MATCHED</td>
<td>49</td>
<td>32</td>
<td>33</td>
<td>37</td>
<td>20</td>
<td>38</td>
<td>209</td>
</tr>
<tr>
<td># CASES</td>
<td>103</td>
<td>100</td>
<td>101</td>
<td>102</td>
<td>100</td>
<td>100</td>
<td>606</td>
</tr>
<tr>
<td>PERCENT</td>
<td>47.6</td>
<td>32.0</td>
<td>32.7</td>
<td>36.3</td>
<td>20.0</td>
<td>38.0</td>
<td>34.5</td>
</tr>
</tbody>
</table>

*Adjusted for converted HMRI data (160 cases)
3.5 SECONDARY PROCEDURES

Not every case has a procedure code. However, if a case does have at least one procedure code, there will be a procedure code in the principal procedure position. A record can have from zero to thirteen secondary procedures coded. Because the number of recorded secondary procedures varies with each case, two types of analysis were undertaken for this category: by number of secondary procedures recorded by the original coder and by number of cases. In both cases, secondary procedures are matched to the three-digit level only. Secondary procedures are considered a match when the secondary procedure coded by the original hospital coder matches the re-abstracted code of the p.i., irrespective of position, including the PP position.

There were 1,259 secondary procedures coded by the p.i. in the 606 cases that were not coded by the original hospital staff. On average, there are 2.1 more procedures per case recorded by the p.i.. The match rate for secondary procedures does not include these procedures. This under-recording of secondary procedures by the hospital coders was consistent over all the hospitals in the study.

3.5.1 By Number of Secondary Procedures Recorded

The match rate for secondary procedures takes the number of secondary procedures actually recorded by the original hospital coder as the denominator. The numerator reflects the number of secondary procedures actually recorded by the original coder that match the procedural codes to the three-digit level, whether coded as PP or secondary by the p.i..
3.5.1.1 Frequency of Match - Overall Hospitals

Table 9 demonstrates the rate for the secondary procedures agreed to the three-digit level over all the hospitals to be 80.7% by number of secondary procedures recorded. This score is higher than that for secondary diagnoses and may indicate that procedures may be more straightforward to code.

3.5.1.2 Range of Match - Individual Hospitals

Table 9 also demonstrates the range for secondary procedures at the three-digit level is from 69.4% for Hospital A to 84.1% for Hospital E by number of secondary procedures recorded. See Table 9 for individual hospital rates.

TABLE 9 - FREQUENCY OF MATCH FOR SECONDARY PROCEDURES 3-DIGIT LEVEL BY NUMBER OF PROCEDURES RECORDED
Re-abstracted Data Apr 1/86 - Mar 31/87

<table>
<thead>
<tr>
<th>HOSP</th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>E</th>
<th>F</th>
<th>TOTAL</th>
</tr>
</thead>
<tbody>
<tr>
<td>MATCHED</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>475</td>
</tr>
<tr>
<td>TOTAL</td>
<td>34</td>
<td>106</td>
<td>111</td>
<td>31</td>
<td>106</td>
<td>87</td>
<td>475</td>
</tr>
<tr>
<td># CODES</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>589</td>
</tr>
<tr>
<td></td>
<td>49</td>
<td>131</td>
<td>141</td>
<td>37</td>
<td>126</td>
<td>105</td>
<td>589</td>
</tr>
<tr>
<td>PERCENT</td>
<td>69.4</td>
<td>80.9</td>
<td>78.7</td>
<td>83.8</td>
<td>84.1</td>
<td>82.9</td>
<td>80.7</td>
</tr>
</tbody>
</table>
3.5.2 By Number of Cases

When secondary procedures are examined by the number of cases, the match rate drops significantly. Cases are matched to the three-digit level. Matched cases include those cases where no secondary procedures are recorded. For a case to be discrepant, only one code needs to disagree. For a code to be discrepant, only one digit needs to disagree. Table 10 demonstrates that only 59.7% of cases over all the hospitals match in the secondary procedures category. The range of match for secondary procedures by case varies from 48.5% for Hospital A to 69.0% for Hospital B.

<table>
<thead>
<tr>
<th></th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>E</th>
<th>F</th>
<th>TOTAL</th>
</tr>
</thead>
<tbody>
<tr>
<td>MATCHED</td>
<td>50</td>
<td>69</td>
<td>55</td>
<td>65</td>
<td>63</td>
<td>60</td>
<td>362</td>
</tr>
<tr>
<td># CASES</td>
<td>103</td>
<td>100</td>
<td>101</td>
<td>102</td>
<td>100</td>
<td>100</td>
<td>606</td>
</tr>
<tr>
<td>PERCENT</td>
<td>48.5</td>
<td>69.0</td>
<td>54.5</td>
<td>63.7</td>
<td>63.0</td>
<td>60.0</td>
<td>59.7</td>
</tr>
</tbody>
</table>
3.6 NON-DIAGNOSTIC VARIABLES

There is a high rate of agreement on all non-diagnostic variables. This category includes five variables: admission date, separation date, birthdate, discharge status and sex.

3.6.1 Frequency of Match - All Hospitals

Table 11 shows match rates for the five variables. For all hospitals combined, admission date agreement was found in 98.8% of cases. Discharge dates agree in 98.8% of cases, birthdate in 96.4%, discharge status in 98.0%, and sex in 99.8%. All non-diagnostic variables match in 92.4% of cases over all the hospitals.

The greatest number of errors is with the 'birthdate' variable, with an error rate of 3.6%, followed by 'discharge status' at 2.0%. There was only one case in the sample (0.2%) with a 'sex' discrepancy.

3.6.2 Range of Match - Individual Hospitals

Table 11 demonstrates that for individual hospitals, the agreement score for admission date ranges from 97.1 to 100%. The separation date match ranges from 96.1 to 100%, birthdate from 92.2 to 100%, discharge status from 96.0 to 99.0 and sex from 99.0 to 100%. The match rate for the combined non-diagnostic variables ranges from 87.3% for Hospital D to 98.0% for Hospital F. For details on individual hospital rates, see Table 11.
TABLE 11 - FREQUENCY OF MATCH FOR NON-DIAGNOSTICS BY NUMBER OF CASES INDIVIDUAL HOSPITALS RATES Re-Abstracted Data Apr 1/86 - Mar 31/87

<table>
<thead>
<tr>
<th>HOSP</th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>E</th>
<th>F</th>
<th>TOTAL</th>
</tr>
</thead>
<tbody>
<tr>
<td>n=</td>
<td>103</td>
<td>100</td>
<td>101</td>
<td>102</td>
<td>100</td>
<td>100</td>
<td>606</td>
</tr>
<tr>
<td>MATCHED</td>
<td>#</td>
<td>%</td>
<td>#</td>
<td>%</td>
<td>#</td>
<td>%</td>
<td>#</td>
</tr>
<tr>
<td>ADMISSION DATE</td>
<td>100</td>
<td>97.1</td>
<td>100</td>
<td>100.0</td>
<td>100</td>
<td>99.0</td>
<td>100</td>
</tr>
<tr>
<td>SEPARATION DATE</td>
<td>99</td>
<td>96.1</td>
<td>99</td>
<td>99.0</td>
<td>101</td>
<td>100.0</td>
<td>102</td>
</tr>
<tr>
<td>BIRTHDATE</td>
<td>102</td>
<td>99.0</td>
<td>96</td>
<td>96.0</td>
<td>99</td>
<td>98.0</td>
<td>94</td>
</tr>
<tr>
<td>DISCHARGE STATUS</td>
<td>101</td>
<td>98.1</td>
<td>96</td>
<td>96.0</td>
<td>100</td>
<td>99.0</td>
<td>100</td>
</tr>
<tr>
<td>SEX</td>
<td>103</td>
<td>100.0</td>
<td>100</td>
<td>100.0</td>
<td>101</td>
<td>100.0</td>
<td>101</td>
</tr>
<tr>
<td>TOTAL NON-DIAGNOSTICS</td>
<td>95</td>
<td>92.2</td>
<td>91</td>
<td>91.0</td>
<td>97</td>
<td>96.0</td>
<td>89</td>
</tr>
</tbody>
</table>
3.6.3 Ranking of Birthdate

A ranking system is used for demographic data, but is restricted to the birthdate variable because of the high agreement or rate of match for the other non-diagnostic variables. Table 12 demonstrates that approximately one-half the discrepancies with birthdate involved differences of less than 30 days, while approximately one-half are greater than one year.

<table>
<thead>
<tr>
<th>TABLE 12 - RANKING OF BIRTHDATE*</th>
</tr>
</thead>
<tbody>
<tr>
<td>BY NUMBER OF RECORDS WITH DISCREPANT BIRTHDATES</td>
</tr>
<tr>
<td>OVER ALL HOSPITALS</td>
</tr>
<tr>
<td>Re-abstracted Data Apr 1/86 - Mar 31/87</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>HOSP</th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>E</th>
<th>F</th>
<th>TOTAL</th>
</tr>
</thead>
<tbody>
<tr>
<td>n=</td>
<td>1</td>
<td>4</td>
<td>2</td>
<td>8</td>
<td>7</td>
<td>0</td>
<td>22</td>
</tr>
<tr>
<td>&lt;30 DAYS</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>6</td>
<td>3</td>
<td>0</td>
<td>10</td>
</tr>
<tr>
<td>&lt;30 DAYS - &lt; 1 YR</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>&gt;1 YR - &lt;5 YR</td>
<td>0</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>6</td>
</tr>
<tr>
<td>&gt;5 YR</td>
<td>0</td>
<td>2</td>
<td>0</td>
<td>1</td>
<td>3</td>
<td>0</td>
<td>6</td>
</tr>
</tbody>
</table>

*Total % of cases not matched = 3.63%
3.7 COMBINED VARIABLES

This category includes the MRD, the PP, secondary diagnoses, and secondary procedures. Because the number of variables recorded varies with each case, the analysis was done two ways: by number of diagnoses and procedures recorded and by number of cases. The MRD and the secondary diagnoses are required to match to the four-digit level (adjusted for conversion of HMRI data), and the PP and secondary procedures to the three-digit level.

3.7.1 By Number of Diagnoses and Procedures

The match rate for the variables analyzed by number of diagnoses and procedures recorded takes the number of secondary diagnoses and secondary procedures actually coded by the original hospital coder as the denominator. The numerator reflects the number of diagnoses and procedures actually recorded by the original coder that match the diagnostic codes to the four-digit level for diagnoses and the three-digit level for procedures, whether coded as MRD, PP, secondary diagnoses or secondary procedures by the p.i..

3.7.1.1 Combined Diagnoses

Table 13 demonstrates the frequency of match for the combined MRD and secondary diagnoses variables at the four-digit, adjusted level to be 65.7%. It also demonstrates the agreement scores for individual hospitals for the MRD and secondary diagnoses, at the four-digit adjusted level, to range between 58.8% for Hospital A and 71.6% for Hospital D by number of diagnoses recorded.
TABLE 13 - FREQUENCY OF MATCH
FOR COMBINED VARIABLES
BY NUMBER OF DIAGNOSES AND PROCEDURES RECORDED
Re-abstracted data Apr 1/86 - Mar 31/87

<table>
<thead>
<tr>
<th>HOSP</th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>E</th>
<th>F</th>
<th>TOTAL</th>
</tr>
</thead>
<tbody>
<tr>
<td>MRD</td>
<td>4-DIGIT</td>
<td>58/100</td>
<td>56/101</td>
<td>71/102</td>
<td>52/100</td>
<td>65/100</td>
<td>372/606</td>
</tr>
<tr>
<td></td>
<td>ADJ* %</td>
<td>68.0</td>
<td>58.0</td>
<td>55.5</td>
<td>69.6</td>
<td>52.0</td>
<td>65.0</td>
</tr>
<tr>
<td>SEC. DIAGS</td>
<td>4-DIGIT</td>
<td>177/260</td>
<td>164/227</td>
<td>206/285</td>
<td>211/320</td>
<td>156/231</td>
<td>991/1470</td>
</tr>
<tr>
<td></td>
<td>ADJ** %</td>
<td>52.4</td>
<td>68.1</td>
<td>72.3</td>
<td>72.3</td>
<td>65.9</td>
<td>67.5</td>
</tr>
<tr>
<td>TOTAL</td>
<td>4-DIGIT %</td>
<td>58.8</td>
<td>65.3</td>
<td>67.1</td>
<td>71.6</td>
<td>62.6</td>
<td>66.8</td>
</tr>
<tr>
<td>PP</td>
<td>3-DIGIT</td>
<td>44/65</td>
<td>49/87</td>
<td>59/78</td>
<td>53/90</td>
<td>50/88</td>
<td>322/497</td>
</tr>
<tr>
<td></td>
<td>%</td>
<td>67.7</td>
<td>75.3</td>
<td>56.3</td>
<td>75.6</td>
<td>58.6</td>
<td>56.8</td>
</tr>
<tr>
<td>SEC PROCS</td>
<td>3-DIGIT</td>
<td>106/131</td>
<td>111/141</td>
<td>31/37</td>
<td>106/126</td>
<td>87/105</td>
<td>475/589</td>
</tr>
<tr>
<td></td>
<td>%</td>
<td>69.4</td>
<td>80.9</td>
<td>78.7</td>
<td>83.8</td>
<td>84.1</td>
<td>82.9</td>
</tr>
<tr>
<td>TOTAL</td>
<td>3-DIGIT %</td>
<td>78/114</td>
<td>173/220</td>
<td>160/228</td>
<td>90/115</td>
<td>159/216</td>
<td>137/193</td>
</tr>
<tr>
<td>TOTAL</td>
<td>DIAGS &amp; PROCS</td>
<td>225/364</td>
<td>408/580</td>
<td>380/536</td>
<td>367/502</td>
<td>422/636</td>
<td>358/524</td>
</tr>
<tr>
<td></td>
<td>%</td>
<td>61.8</td>
<td>70.3</td>
<td>68.4</td>
<td>73.1</td>
<td>66.4</td>
<td>68.3</td>
</tr>
</tbody>
</table>

* Adjusted for converted HMRI data (99 cases)
** Adjusted for converted HMRI data (160 cases)
# Number of matched codes over total number of codes recorded
3.7.1.2 Combined Procedures

Table 13 demonstrates the frequency of match for the combined PP and secondary procedure variables at the three-digit level to be 73.4%. It also demonstrates the match for individual hospitals for the PP and secondary procedures, at the three-digit level, to range between 68.4% for Hospital A and 78.6% for Hospital B by number of procedures recorded.

3.7.1.3 Combined Diagnoses and Procedures

Table 13 demonstrates the frequency of match for the combined diagnostic variables to four-digits and the procedural variables to three-digits to be 68.3%. It also demonstrates the match for individual hospitals for the combined diagnostic (four-digit level, adjusted) and procedural (three-digit level) categories to range between 61.8% for Hospital A and 73.1% for Hospital D by number of procedures or diagnoses recorded.
3.7.2 By Number of Cases

For a case to match, the MRD and all secondary diagnostic codes to the four digits (adjusted level) recorded by the original hospital coder and all the PP and secondary procedural codes to the three-digit level, if applicable, must match the re-abstracted codes of the p.i. The MRD must maintain the MRD position in both sets of data and the PP must maintain the PP position in both sets of data, unless the record contains only non-operative procedures. If the record contains only non-operative procedures, then the position of the matched code is not relevant. In addition, the extra secondary codes recorded by the p.i. are not examined.

The data for Tables 8 and 10 are presented again in Table 14, which demonstrates that when agreement is assessed by case, the match rate drops significantly from the match rate by number of procedures or diagnoses recorded. When examined by case, 25.6% of total cases match in all diagnostic and procedural variables. The match rate for individual hospitals by case for diagnostic and procedural codes ranges from 17.0% for Hospital E to 39.8% for Hospital A.
TABLE 14 - FREQUENCY OF MATCH
FOR DIAGNOSTIC AND PROCEDURAL CODES
BY NUMBER OF CASES
Re-abstracted Data Apr 1/86 - Mar 31/87

<table>
<thead>
<tr>
<th>HOSP</th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>E</th>
<th>F</th>
<th>TOTAL</th>
</tr>
</thead>
<tbody>
<tr>
<td>n-</td>
<td>103</td>
<td>100</td>
<td>101</td>
<td>102</td>
<td>100</td>
<td>100</td>
<td>606</td>
</tr>
<tr>
<td>DIAGNOSES 4-DIGIT, ADJ* #</td>
<td>49</td>
<td>32</td>
<td>33</td>
<td>37</td>
<td>20</td>
<td>38</td>
<td>209</td>
</tr>
<tr>
<td>%</td>
<td>47.6</td>
<td>32.0</td>
<td>32.7</td>
<td>36.3</td>
<td>20.0</td>
<td>38.0</td>
<td>34.5</td>
</tr>
<tr>
<td>PROCEDURES 3-DIGIT #</td>
<td>50</td>
<td>69</td>
<td>55</td>
<td>65</td>
<td>63</td>
<td>60</td>
<td>362</td>
</tr>
<tr>
<td>%</td>
<td>48.5</td>
<td>69.0</td>
<td>54.5</td>
<td>63.7</td>
<td>63.0</td>
<td>60.0</td>
<td>59.7</td>
</tr>
<tr>
<td>TOTAL DIAGS &amp; PROC  #</td>
<td>41</td>
<td>22</td>
<td>23</td>
<td>31</td>
<td>17</td>
<td>21</td>
<td>155</td>
</tr>
<tr>
<td>%</td>
<td>39.8</td>
<td>22.0</td>
<td>22.8</td>
<td>30.4</td>
<td>17.0</td>
<td>21.0</td>
<td>25.6</td>
</tr>
</tbody>
</table>

*Adjusted for converted HMRI data (160 cases)
3.8 DIAGNOSTIC CATEGORIES

This section examines whether some diagnostic categories demonstrated a higher frequency of discrepancy than others. The analysis of this question is confined to the MRD, four-digit, adjusted, and the PP, three-digit codes. A complete listing of the categories appears in Appendix I. Criteria for initial random selection of cases were that a case had to have the appropriate code in either the MRD or secondary diagnostic position within each of the diagnostic categories. However, in this analysis, diagnostic position was determined by MRD only. Therefore, some categories had fewer than the expected number of cases. Groups 1, 8 and 9 were not included in this analysis because of small numbers.

3.8.1 Frequency of Match - Most Responsible Diagnosis

Table 15 demonstrates the least degree of match for category 5 (codes 460-519, Diseases of the Respiratory System) at 49.6% by MRD. The highest degree of match was for category 6 (Diseases of the Digestive System) at 76.1% by MRD. Categories 10 (Injury and poisoning), 2 (Endocrine, Nutritional and Metabolic Diseases, Immunity Disorders AND Diseases of the Blood and blood Forming Organs) and 5 (Respiratory system) were below the average across hospitals (61.4%).
TABLE 15 - FREQUENCY OF MATCH FOR MOST RESPONSIBLE DIAGNOSIS 4-DIGIT LEVEL, ADJUSTED* BY DIAGNOSTIC CATEGORIES** OVER ALL HOSPITALS
Re-Abstracted Data Apr 1/86 - Mar 31/87

<table>
<thead>
<tr>
<th>DIAGNOSTIC CATEGORIES</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td>3-DIGIT CODES</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>240-289</td>
<td>6</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>320-389</td>
<td>17</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>390-459</td>
<td>51</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>460-519</td>
<td>64</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>520-579</td>
<td>11</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DISCREPANT</td>
<td>61</td>
<td>47.7</td>
<td>17</td>
<td>35.4</td>
<td></td>
</tr>
<tr>
<td>MATCH</td>
<td>67</td>
<td>52.3</td>
<td>31</td>
<td>64.6</td>
<td></td>
</tr>
<tr>
<td>TOTAL</td>
<td>128</td>
<td>100.0</td>
<td>48</td>
<td>100.0</td>
<td>164</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>DIAGNOSTIC CATEGORIES</th>
<th>7</th>
<th>10</th>
<th>1-10</th>
</tr>
</thead>
<tbody>
<tr>
<td>3-DIGIT CODES</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>580-629</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>800-999</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TOTAL</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DISCREPANT</td>
<td>12</td>
<td>27.3</td>
<td></td>
</tr>
<tr>
<td>MATCH</td>
<td>32</td>
<td>72.7</td>
<td></td>
</tr>
<tr>
<td>TOTAL</td>
<td>44</td>
<td>100.0</td>
<td>30</td>
</tr>
</tbody>
</table>

*Adjusted for converted HMRI data (99 cases)
**Categories 1,8 and 9 have been dropped from analysis
3.8.2 Frequency of Match - Principal Procedure

For the PP, the initial categories for the "groups" were determined by diagnosis, not by procedure. Procedural codes use the Canadian Classification of Diagnostic, Therapeutic, and Surgical Procedures Index (CCP) and the procedural codes do not fall under any diagnostic grouping. Although some procedures can be directly related to a diagnostic category, others cannot. For example, by-pass heart surgery can be directly related to category 4 (Diseases of the circulatory system), while non-operative procedures, such as a chest x-ray may not be directly related to category 2 (diabetic category). Not all procedures are recorded on every case. As with the PP and secondary procedures categories, the denominator of 497 procedures represents only those cases with a recorded procedure.

For the PP, Table 16 demonstrates that for category 2 (codes 240-289, Endocrine, Nutritional and Metabolic Diseases), 60.2% of the PPs match. For category 4 (codes 390-459, Diseases of the circulatory system), there are 68.0% of PPs that match. Category 5 (codes 460-519, Diseases of the Respiratory System), demonstrates 60.0% of the PPs match.

Table 16 demonstrates that categories range in agreement from 60.0% for category 5 to 75.0% for category 10.
TABLE 16 - FREQUENCY OF MATCH FOR PRINCIPAL PROCEDURE 3-DIGIT LEVEL
BY DIAGNOSTIC CATEGORIES*
OVER ALL HOSPITALS
Re-Abstracted Data Apr 1/86 - Mar 31/87

<table>
<thead>
<tr>
<th>DIAGNOSTIC CATEGORIES</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td>3-DIGIT CODES</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>240-289</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>320-389</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>390-459</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>460-519</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>520-579</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>#</td>
<td>39</td>
<td>39.8</td>
<td>50</td>
<td>32.1</td>
<td>36</td>
</tr>
<tr>
<td>%</td>
<td>39.8</td>
<td>15</td>
<td>35.7</td>
<td>50</td>
<td>32.1</td>
</tr>
<tr>
<td>DISCREPANT</td>
<td>39</td>
<td>39.8</td>
<td>50</td>
<td>32.1</td>
<td>36</td>
</tr>
<tr>
<td>MATCH</td>
<td>59</td>
<td>60.2</td>
<td>106</td>
<td>68.0</td>
<td>54</td>
</tr>
<tr>
<td>TOTAL</td>
<td>98</td>
<td>100.0</td>
<td>156</td>
<td>100.0</td>
<td>90</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>DIAGNOSTIC CATEGORIES</th>
<th>7</th>
<th>10</th>
<th>1-10</th>
</tr>
</thead>
<tbody>
<tr>
<td>3-DIGIT CODES</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>580-629</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>800-999</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TOTAL</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>#</td>
<td>5</td>
<td>175</td>
<td>35.2</td>
</tr>
<tr>
<td>%</td>
<td>26.3</td>
<td>5</td>
<td>25.0</td>
</tr>
<tr>
<td>DISCREPANT</td>
<td>10</td>
<td>26.3</td>
<td>5</td>
</tr>
<tr>
<td>MATCH</td>
<td>28</td>
<td>73.7</td>
<td>15</td>
</tr>
<tr>
<td>TOTAL</td>
<td>38</td>
<td>100.0</td>
<td>20</td>
</tr>
</tbody>
</table>

*Categories 1, 8 and 9 have been dropped from the analysis
3.9 TYPE OF DISCREPANCY

Type of discrepancy was determined for those codes coded by the original hospital coder that did not match that of the p.i. Each discrepant code is assigned only one type of discrepancy. There can be several types of discrepancy for each case, depending on the number of discrepant codes per case. This category is examined by number of diagnoses and procedures recorded. For a detailed description of the type of discrepancy categories, see section 2.12. The discrepancy category for the converted HMRI data is used only to determine the changes in the conversion of data and is not examined in this section. There are variations in the implications of a discrepancy depending on whether it is in the MRD or PP position on the one hand, or for example, the thirteenth secondary diagnosis or procedure on the other.

3.9.1 Frequency of Type - All Hospitals

The type of discrepancy was determined for each discrepant variable within the MRD, PP, secondary diagnoses, secondary procedures, non-diagnostics, combined diagnoses, combined procedures, combined diagnoses and procedures categories.

3.9.1.1 Most Responsible Diagnosis

Table 17 demonstrates that 73.1% of discrepancies with the MRD can be attributed to ‘clerical’ errors, those arising as a result of human carelessness. Clerical errors include coding from the face sheet when relevant information, documented in the chart, was not captured. Reliance on the Alphabetic Index with no reference to the Tabular List was
TABLE 17 - FREQUENCY OF "TYPE OF DISCREPANCY"
AS A PERCENT OF DIAGNOSES AND PROCEDURES RECORDED
ALL HOSPITALS
Re-Abstracted Data Apr 1/86 - Mar 31/87

<table>
<thead>
<tr>
<th></th>
<th>JUDGE MENT</th>
<th>CLASSIFIC.</th>
<th>PROCEDURE</th>
<th>CLERICAL</th>
<th>INCOM RECOR</th>
<th>DISCREPANT</th>
<th>OTHER</th>
<th>TOTAL # DISC</th>
</tr>
</thead>
<tbody>
<tr>
<td>MRD #</td>
<td>6</td>
<td>13</td>
<td>14</td>
<td>171</td>
<td>4</td>
<td>16</td>
<td>10</td>
<td>234</td>
</tr>
<tr>
<td>%</td>
<td>2.6</td>
<td>5.6</td>
<td>6.0</td>
<td>73.1</td>
<td>1.7</td>
<td>6.8</td>
<td>4.3</td>
<td>100.1</td>
</tr>
<tr>
<td>SEC #</td>
<td>0</td>
<td>24</td>
<td>11</td>
<td>317</td>
<td>13</td>
<td>13</td>
<td>102</td>
<td>480</td>
</tr>
<tr>
<td>%</td>
<td>0.0</td>
<td>5.0</td>
<td>2.3</td>
<td>66.0</td>
<td>2.7</td>
<td>2.7</td>
<td>21.3</td>
<td>100.0</td>
</tr>
<tr>
<td>TOTAL #</td>
<td>6</td>
<td>37</td>
<td>25</td>
<td>488</td>
<td>17</td>
<td>29</td>
<td>112</td>
<td>714</td>
</tr>
<tr>
<td>%</td>
<td>0.8</td>
<td>5.2</td>
<td>3.5</td>
<td>68.3</td>
<td>2.4</td>
<td>4.1</td>
<td>15.7</td>
<td>100.0</td>
</tr>
<tr>
<td>PP #</td>
<td>82</td>
<td>2</td>
<td>5</td>
<td>75</td>
<td>2</td>
<td>3</td>
<td>6</td>
<td>175</td>
</tr>
<tr>
<td>%</td>
<td>46.9</td>
<td>1.1</td>
<td>2.9</td>
<td>42.9</td>
<td>1.1</td>
<td>1.7</td>
<td>3.4</td>
<td>100.0</td>
</tr>
<tr>
<td>SEC #</td>
<td>0</td>
<td>6</td>
<td>2</td>
<td>79</td>
<td>3</td>
<td>1</td>
<td>23</td>
<td>114</td>
</tr>
<tr>
<td>%</td>
<td>0.0</td>
<td>5.3</td>
<td>1.8</td>
<td>69.3</td>
<td>2.6</td>
<td>0.9</td>
<td>20.2</td>
<td>100.1</td>
</tr>
<tr>
<td>TOTAL #</td>
<td>82</td>
<td>8</td>
<td>7</td>
<td>154</td>
<td>5</td>
<td>4</td>
<td>29</td>
<td>289</td>
</tr>
<tr>
<td>%</td>
<td>28.4</td>
<td>2.8</td>
<td>2.4</td>
<td>53.3</td>
<td>1.7</td>
<td>1.4</td>
<td>10.0</td>
<td>100.0</td>
</tr>
<tr>
<td>TOTAL #</td>
<td>88</td>
<td>45</td>
<td>32</td>
<td>642</td>
<td>22</td>
<td>33</td>
<td>141</td>
<td>1003</td>
</tr>
<tr>
<td>%</td>
<td>8.8</td>
<td>4.5</td>
<td>3.2</td>
<td>64.0</td>
<td>2.2</td>
<td>3.3</td>
<td>14.1</td>
<td>100.1</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>NON- DIAGS</th>
<th></th>
<th></th>
<th>22</th>
<th>16</th>
<th>10</th>
<th>48</th>
</tr>
</thead>
<tbody>
<tr>
<td>%</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>45.8</td>
<td>33.3</td>
<td>20.8</td>
<td>99.9</td>
</tr>
</tbody>
</table>
quite evident (i.e., mandatory coding, exclusions and inclusions were not used). Singular diagnoses are easier to code than multiple diagnoses. Only 2.6% of the discrepancies could be considered ‘judgement’ errors, derived from having to choose between two or more diagnosis being the MRD. Hospital D and Hospital E had no judgement discrepancies with the MRD. Changes in hospital policy (procedural errors) accounted for 6.0% of discrepancies. Discrepant records, where there was inadequate information to determine a code or inconsistency among the recorders, accounted for only 6.8%. Hospital E had the most discrepant records with 14.6%. The null hypothesis that there is more than one dominant type of discrepancy is rejected.

3.9.1.2 Principal Procedure

In contrast, Table 17 demonstrates that most discrepancies in the PP category are in the ‘judgement’ category with 46.9%, followed by 42.9% ‘clerical’ errors. Determining which procedure to code as the principal procedure and the corresponding lack of CCP procedure guidelines are both confirmed as problems with the current HMRI (CIHI) system. When two major procedures on one patient are performed, it is not always clear which procedure is the principal one. Again, reliance on the Alphabetic Index with no reference to the Tabular List was quite evident. The PP is the only category where there is more than one dominant type of discrepancy.

3.9.1.3 Secondary Diagnoses

Table 17 demonstrates that the highest number of errors for secondary diagnoses were clerical errors (66.0%) followed by the ‘other’ category with 21.3%. The exact
nature of the 'other' category is uncertain; however, most of these can probably be attributed to clerical errors. It needs to be recognized that 2,500 secondary diagnoses recorded by the p.i. and not by the original hospital coders over the 606 cases in the study are not included in this analysis. These missed diagnoses could be attributed to any one of the discrepancy categories. Thus, this also under-represents the actual number of clerical errors.

3.9.1.4 Secondary Procedures

Table 17 reveals that the highest rate of discrepancy among secondary procedures was in the 'clerical' category with 69.3%, followed by 'other' at 20.2%. There were no 'judgement' discrepancies within this category. The dominant type of discrepancy for secondary procedures differs from that for the PP; however, 'clerical' discrepancies were pronounced in both the PP and the secondary procedures. Again, most of the 'other' errors could probably be attributed to clerical errors. There were very few discrepant or incomplete records, or changes to hospital policy.

3.9.1.5 Non-Diagnoses

The non-diagnostic types of discrepancy are restricted to 'clerical' errors, 'discrepant' records and 'other' categories. Table 17 demonstrates that the types of discrepancy for the non-diagnostic category were most commonly 'clerical' errors at 45.8%. 'Discrepant' records accounted for 33.3% of errors, and 'other' errors were 20.8%. All hospitals, except Hospital D, were downloading Admission-Discharge-Transfer (ADT) information. An ADT system has the potential to eliminate clerical
errors by eliminating the possibility of data conversion through erroneously recording the ADT information. However, this was not the case. These type of data conversion errors are attributed to the coder in this study. Even though errors may be generated in the Admissions Department where the downloaded ADT information originates, it is the coder's responsibility to ensure this information is correct for the HMRI/CIHI database.

3.9.1.6 Total Diagnoses

Table 17 demonstrates the discrepancies in the combined MRD and secondary diagnoses variables to be highest in the 'clerical' category with 68.3%, followed by 'other' at 15.7%. Clerical discrepancies were prominent in both the MRD and the secondary diagnoses.

For all hospitals, clerical errors, such as not reviewing the entire record and misuse of the classification system were the top discrepancies. However, each individual hospital varied as to the order in which these occurred. Hospitals A, B & C admitted to coding policies that were different from HMRI and/or ICD guidelines, but only Hospital C showed the same number of discrepancies from misuse of the classification system as from not reviewing the entire chart.

3.9.1.7 Total Procedures

Discrepancies for the combined PP and secondary procedure variables are greatest in the 'clerical' category with 53.3%, followed by 'judgement' at 28.4%. Discrepant and incomplete records account for only 1.4% and 1.7% respectively. There were very few hospital policies that affected the coding of procedures.
3.9.1.8 Total Diagnoses and Procedures

Table 17 demonstrated that when diagnostic and procedural data are combined, the greatest percentage of discrepancies is clearly with clerical errors. The errors attributed to a lack of information in the record, discrepant records, lack of word-for-word correspondence between the coding manuals and documentation within the record, and changes to hospital policy were very few. Total judgement categories were also fairly low over all. Considering the range which hospitals code from incomplete records, i.e., from 10 to 90%, there is not much difference in the range of discrepancies.

Hospitals D & E, with 47.37% and 48.65% respectively, had the highest number of judgement discrepancies, yet had no selection discrepancies for the MRD. Hospital B demonstrated the lowest judgement discrepancies in both the MRD and the PP categories.

There were 20.2% of discrepancies overall for the secondary procedures that could not be categorized. Hospital B had 44.0% of discrepancies in secondary procedures that could not be categorized.

3.9.2 Range of Frequency - Individual Hospitals

The range of frequency for type of discrepancy is determined for each individual hospital in each of the following categories: the MRD, the PP, secondary diagnoses, secondary procedures, non-diagnostics and combined diagnoses and procedures.
3.9.2.1 *Most Responsible Diagnosis*

Table 18 demonstrates the range for types of discrepancy by individual hospital for the MRD as follows:

<table>
<thead>
<tr>
<th>Category</th>
<th>Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>Judgement</td>
<td>0.0 - 6.1%</td>
</tr>
<tr>
<td>Classification</td>
<td>3.2 - 8.6%</td>
</tr>
<tr>
<td>Procedural</td>
<td>0.0 - 14.6%</td>
</tr>
<tr>
<td>Clerical</td>
<td>60.4 - 80.6%</td>
</tr>
<tr>
<td>Incomplete Record</td>
<td>0.0 - 6.3%</td>
</tr>
<tr>
<td>Discrepant</td>
<td>0.0 - 14.6%; and</td>
</tr>
<tr>
<td>Other</td>
<td>0.0 - 9.5%</td>
</tr>
</tbody>
</table>

The greatest range in discrepancy rates in the MRD category was in the clerical, with Hospital E having 60.4% and Hospital D having 80.6% of its errors in this category. Hospital E had no ‘judgement’ errors, but was highest for ‘procedural’ errors, i.e., changes to hospital policy, and discrepant records, both at 14.6%. Hospital F had no discrepancies attributable to changes in hospital policy, incomplete or discrepant records for the MRD. Hospital D had no discrepancies due to judgement or incomplete records. Although only Hospitals D, E & F claimed to have different policies from HMRI or the ICD guidelines, no major differences were noted between the hospitals.
<table>
<thead>
<tr>
<th>%</th>
<th>#</th>
<th>%</th>
<th>#</th>
<th>%</th>
<th>#</th>
<th>%</th>
<th>#</th>
<th>%</th>
<th>#</th>
<th>%</th>
<th>#</th>
<th>%</th>
<th>#</th>
<th>%</th>
<th>#</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hospital A</td>
<td>1.0%</td>
<td>0.0%</td>
<td>2.0%</td>
<td>1.0%</td>
<td>3.0%</td>
<td>1.0%</td>
<td>2.0%</td>
<td>1.0%</td>
<td>3.0%</td>
<td>1.0%</td>
<td>3.0%</td>
<td>1.0%</td>
<td>3.0%</td>
<td>1.0%</td>
<td>3.0%</td>
</tr>
<tr>
<td>Hospital B</td>
<td>1.4%</td>
<td>2.0%</td>
<td>4.0%</td>
<td>2.0%</td>
<td>4.7%</td>
<td>3.0%</td>
<td>7.1%</td>
<td>4.0%</td>
<td>9.5%</td>
<td>4.0%</td>
<td>9.5%</td>
<td>4.0%</td>
<td>9.5%</td>
<td>4.0%</td>
<td>9.5%</td>
</tr>
<tr>
<td>Hospital C</td>
<td>2.4%</td>
<td>3.0%</td>
<td>6.0%</td>
<td>3.0%</td>
<td>6.7%</td>
<td>1.0%</td>
<td>2.2%</td>
<td>3.0%</td>
<td>5.0%</td>
<td>1.0%</td>
<td>2.2%</td>
<td>3.0%</td>
<td>5.0%</td>
<td>1.0%</td>
<td>2.2%</td>
</tr>
<tr>
<td>Hospital D</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>2.0%</td>
<td>0.0%</td>
<td>2.0%</td>
<td>0.0%</td>
<td>2.0%</td>
<td>0.0%</td>
<td>2.0%</td>
<td>0.0%</td>
<td>2.0%</td>
<td>0.0%</td>
<td>2.0%</td>
<td>0.0%</td>
</tr>
<tr>
<td>Hospital E</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
</tr>
<tr>
<td>Hospital F</td>
<td>1.2%</td>
<td>3.0%</td>
<td>9.7%</td>
<td>3.0%</td>
<td>2.0%</td>
<td>0.0%</td>
<td>2.0%</td>
<td>0.0%</td>
<td>2.0%</td>
<td>0.0%</td>
<td>2.0%</td>
<td>0.0%</td>
<td>2.0%</td>
<td>0.0%</td>
<td>2.0%</td>
</tr>
<tr>
<td>Hospital G</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
</tr>
<tr>
<td>Hospital H</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
</tr>
<tr>
<td>Hospital I</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
</tr>
<tr>
<td>Hospital J</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
</tr>
<tr>
<td>Hospital K</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
</tr>
<tr>
<td>Hospital L</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
</tr>
<tr>
<td>Hospital M</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
</tr>
<tr>
<td>Hospital N</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
</tr>
<tr>
<td>Hospital O</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
</tr>
<tr>
<td>Hospital P</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
</tr>
<tr>
<td>Hospital Q</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
</tr>
<tr>
<td>Hospital R</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
</tr>
<tr>
<td>Hospital S</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
</tr>
<tr>
<td>Hospital T</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
</tr>
<tr>
<td>Hospital U</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
</tr>
<tr>
<td>Hospital V</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
</tr>
<tr>
<td>Hospital W</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
</tr>
<tr>
<td>Hospital X</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
</tr>
<tr>
<td>Hospital Y</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
</tr>
<tr>
<td>Hospital Z</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
</tr>
</tbody>
</table>
3.9.2.2 Principal Procedure

Table 19 demonstrates the range for types of discrepancy by individual hospital for the PP as follows:

<table>
<thead>
<tr>
<th>Classification</th>
<th>40.9% - 48.7%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Classification</td>
<td>0.0% - 4.6%</td>
</tr>
<tr>
<td>Procedural</td>
<td>0.0% - 10.5%</td>
</tr>
<tr>
<td>Clerical</td>
<td>36.8% - 45.5%</td>
</tr>
<tr>
<td>Incomplete Record</td>
<td>0.0% - 2.7%</td>
</tr>
<tr>
<td>Discrepant</td>
<td>0.0% - 5.3%</td>
</tr>
<tr>
<td>Other</td>
<td>0.0% - 9.1%</td>
</tr>
</tbody>
</table>

The greatest number of errors for the PP category was in the 'judgement' category. All hospitals were relatively high in this category, with the average being 46.9% across hospitals. Clerical errors were also relatively high with the average 42.9% across hospitals. There were very few incomplete or discrepant records, word-for-word correspondence or hospital policy changes, consistent across hospitals. Even though Hospitals D, E & F claimed to employ policies different from HMRI, again, no differences between the hospitals were noted.
### TABLE 19 - FREQUENCY OF DISCREPANCIES
BY TYPE OF DISCREPANCY, PRINCIPAL PROCEDURE
BY NUMBER OF PROCEDURES RECORDED BY ORIGINAL CODER
FOR INDIVIDUAL HOSPITALS
Re-Abstracted Data Apr 1/86 - Mar 31/87

<table>
<thead>
<tr>
<th>DISCREPANT CODES</th>
<th>JUDGEMENT</th>
<th>CLASSIFICATION</th>
<th>PROCEDURAL</th>
<th>CLERICAL</th>
<th>INCOMP RECORD</th>
<th>DISCREPANCY</th>
<th>OTHER</th>
<th>TOTAL DISCREPANT</th>
</tr>
</thead>
<tbody>
<tr>
<td>HOSP A</td>
<td>10</td>
<td>47.6</td>
<td>0</td>
<td>0.0</td>
<td>9</td>
<td>42.9</td>
<td>0</td>
<td>1 4.8</td>
</tr>
<tr>
<td>HOSP B</td>
<td>9</td>
<td>40.9</td>
<td>1</td>
<td>4.6</td>
<td>10</td>
<td>45.5</td>
<td>0</td>
<td>0 0.0</td>
</tr>
<tr>
<td>HOSP C</td>
<td>18</td>
<td>47.4</td>
<td>0</td>
<td>0.0</td>
<td>17</td>
<td>44.7</td>
<td>0</td>
<td>0 0.0</td>
</tr>
<tr>
<td>HOSP D</td>
<td>9</td>
<td>47.4</td>
<td>0</td>
<td>0.0</td>
<td>7</td>
<td>36.8</td>
<td>0</td>
<td>1 5.3</td>
</tr>
<tr>
<td>HOSP E</td>
<td>18</td>
<td>48.7</td>
<td>1</td>
<td>2.7</td>
<td>15</td>
<td>40.5</td>
<td>1</td>
<td>2.7 1 2.7</td>
</tr>
<tr>
<td>HOSP F</td>
<td>18</td>
<td>47.4</td>
<td>0</td>
<td>0.0</td>
<td>17</td>
<td>44.7</td>
<td>1</td>
<td>2.6 0 0.0</td>
</tr>
<tr>
<td>TOTAL</td>
<td>82</td>
<td>46.9</td>
<td>2</td>
<td>1.1</td>
<td>75</td>
<td>42.9</td>
<td>2</td>
<td>1.1 3 1.7</td>
</tr>
</tbody>
</table>
3.9.2.3 *Secondary Diagnoses*

Table 20 demonstrates the range for types of discrepancy by individual hospital for secondary diagnoses as follows:

- **Classification**: 0.0 - 9.5%;
- **Procedural**: 0.0 - 9.6%;
- **Clerical**: 54.9 - 80.7%;
- **Incomplete Record**: 0.0 - 8.0%;
- **Discrepant**: 0.0 - 4.8%; and
- **Other**: 11.0 - 28.2%.

The greatest range of discrepancy for secondary diagnoses was found in the clerical, which also demonstrated the greatest number of discrepancies for secondary diagnoses. There were no judgement discrepancies demonstrated in secondary diagnoses. Hospital B showed no discrepant records or classification errors at all but was one of the highest in procedural errors (changes in hospital policy) at 9.6%. Hospital E had the highest clerical errors at 80.7%. For secondary diagnoses, 21.3% of errors were in the "other" category where the exact nature of the coding error could not be determined. However, most of these errors can probably be attributed to clerical errors.
Table 20 - Frequency of Discrepancies
By Type of Discrepancy, Secondary Diagnoses
By Number of Secondary Diagnoses Recorded by Original Coder
For Individual Hospitals
Re-Abstracted Data Apr 1/86 - Mar 31/87

<table>
<thead>
<tr>
<th>DISCREPANT CODES</th>
<th>JUDGEMENT</th>
<th>CLASSIFICATION</th>
<th>PROCEDURAL</th>
<th>CLERICAL</th>
<th>INCOMP RECORD</th>
<th>DISCREPANCY</th>
<th>OTHER</th>
<th>TOTAL DISCREPANT</th>
</tr>
</thead>
<tbody>
<tr>
<td>HOSP A</td>
<td>0</td>
<td>0.0</td>
<td>3</td>
<td>4.2</td>
<td>39</td>
<td>54.9</td>
<td>3</td>
<td>4.2</td>
</tr>
<tr>
<td>HOSP B</td>
<td>0</td>
<td>0.0</td>
<td>0</td>
<td>0.0</td>
<td>8</td>
<td>9.6</td>
<td>54</td>
<td>65.1</td>
</tr>
<tr>
<td>HOSP C</td>
<td>0</td>
<td>0.0</td>
<td>6</td>
<td>9.5</td>
<td>42</td>
<td>66.7</td>
<td>0</td>
<td>0.0</td>
</tr>
<tr>
<td>HOSP D</td>
<td>0</td>
<td>0.0</td>
<td>4</td>
<td>5.1</td>
<td>47</td>
<td>59.5</td>
<td>4</td>
<td>5.1</td>
</tr>
<tr>
<td>HOSP E</td>
<td>0</td>
<td>0.0</td>
<td>7</td>
<td>6.4</td>
<td>88</td>
<td>80.7</td>
<td>0</td>
<td>0.0</td>
</tr>
<tr>
<td>HOSP F</td>
<td>0</td>
<td>0.0</td>
<td>4</td>
<td>5.3</td>
<td>47</td>
<td>62.7</td>
<td>6</td>
<td>8.0</td>
</tr>
<tr>
<td>TOTAL</td>
<td>0</td>
<td>0.0</td>
<td>24</td>
<td>5.0</td>
<td>317</td>
<td>66.0</td>
<td>13</td>
<td>2.7</td>
</tr>
</tbody>
</table>
3.9.2.4  Secondary Procedures

Table 21 demonstrates the range for types of discrepancy by individual hospital for secondary procedures as follows:

<table>
<thead>
<tr>
<th>Classification</th>
<th>0.0 - 16.7%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Procedural</td>
<td>0.0 - 5.6%</td>
</tr>
<tr>
<td>Clerical</td>
<td>44.0 - 90.0%</td>
</tr>
<tr>
<td>Incomplete Record</td>
<td>0.0 - 11.1</td>
</tr>
<tr>
<td>Discrepant</td>
<td>0.0 - 4.0%; and</td>
</tr>
<tr>
<td>Other</td>
<td>6.7 - 44.0%</td>
</tr>
</tbody>
</table>

There were no judgement discrepancies demonstrated for secondary procedures. The greatest number of errors were clerical which also had the largest range. The average in clerical errors across hospitals was 69.3%. 20.2% of errors were assigned ‘other’ type of discrepancy. Although the exact nature of coding error could not be determined for certain, it is assumed most of these errors were clerical. Hospital C’s discrepancies were almost all (90.0%) in the clerical category. Only one error over all the hospitals was attributed to a discrepancy within a record and only three errors were found for incomplete information within a record.
### TABLE 21 - FREQUENCY OF DISCREPANCIES
BY TYPE OF DISCREPANCY, SECONDARY PROCEDURES
BY NUMBER OF SECONDARY PROCEDURES RECORDED BY ORIGINAL CODE
FOR INDIVIDUAL HOSPITALS
Re-abstracted Data Apr 1/86 - Mar 31/87

<table>
<thead>
<tr>
<th>DISCREPANT CODES</th>
<th>JUDGEMENT</th>
<th>CLASSIFICATION</th>
<th>PROCEDURAL</th>
<th>CLERICAL</th>
<th>INCOMP. RECORD</th>
<th>DISCREPANCY</th>
<th>OTHER</th>
<th>TOTAL DISCREPANT</th>
</tr>
</thead>
<tbody>
<tr>
<td>HOSP A</td>
<td>0 0.0</td>
<td>1 6.7</td>
<td>0 0.0</td>
<td>11 73.3</td>
<td>0 0.0</td>
<td>0 0.0</td>
<td>3 20.0</td>
<td>15 100.0</td>
</tr>
<tr>
<td>HOSP B</td>
<td>0 0.0</td>
<td>1 4.0</td>
<td>1 4.0</td>
<td>11 44.0</td>
<td>0 0.0</td>
<td>1 4.0</td>
<td>11 44.0</td>
<td>25 100.0</td>
</tr>
<tr>
<td>HOSP C</td>
<td>0 0.0</td>
<td>1 3.3</td>
<td>0 0.0</td>
<td>27 90.0</td>
<td>0 0.0</td>
<td>0 0.0</td>
<td>2 6.7</td>
<td>30 100.0</td>
</tr>
<tr>
<td>HOSP D</td>
<td>0 0.0</td>
<td>1 16.7</td>
<td>0 0.0</td>
<td>4 66.7</td>
<td>0 0.0</td>
<td>0 0.0</td>
<td>1 16.7</td>
<td>6 100.0</td>
</tr>
<tr>
<td>HOSP E</td>
<td>0 0.0</td>
<td>2 10.0</td>
<td>0 0.0</td>
<td>13 65.0</td>
<td>1 5.0</td>
<td>0 0.0</td>
<td>4 20.0</td>
<td>20 100.0</td>
</tr>
<tr>
<td>HOSP F</td>
<td>0 0.0</td>
<td>0 0.0</td>
<td>1 5.6</td>
<td>13 72.2</td>
<td>2 11.1</td>
<td>0 0.0</td>
<td>2 11.1</td>
<td>18 100.0</td>
</tr>
<tr>
<td>TOTAL</td>
<td>0 0.0</td>
<td>6 5.3</td>
<td>2 1.8</td>
<td>79 69.3</td>
<td>3 2.6</td>
<td>1 0.9</td>
<td>23 20.2</td>
<td>114 100.0</td>
</tr>
</tbody>
</table>
3.9.2.5 Non-Diagnostics

Table 22 demonstrates the range for types of discrepancy for individual hospitals for the non-diagnostics category as follows:

- Clerical: 18.2 - 100%;
- Discrepant: 0.0 - 54.6%; and
- Other: 0.0 - 27.3%.

Only three types of discrepancy were found for the non-diagnostics category. The greatest range and number of errors were for clerical discrepancies. Records where the information was discrepant accounted for 33.3% of errors.

According to the hospital questionnaire, referred to in section 2.9.2, only Hospital D was not downloading Admission-Discharge-Transfer (ADT) information into the HMRI database. Downloading ADT information into the database can reduce the number of clerical errors in the recording of such information. Hospital D had the highest number of clerical errors; however, there were only six errors in total.
### TABLE 22 - FREQUENCY OF DISCREPANCIES
BY TYPE OF DISCREPANCY, BY NON-DIAGNOSTICS
FOR INDIVIDUAL HOSPITALS
Re-Abstracted Data Apr 1/86 - Mar 31/87

<table>
<thead>
<tr>
<th>DISCREPANT CODES</th>
<th>JUDGEMENT</th>
<th>CLASSIFICATION</th>
<th>PROCEDURAL</th>
<th>CLERICAL</th>
<th>INCOMP. RECORD</th>
<th>DISCREPANCY</th>
<th>OTHER</th>
<th>TOTAL DISCREPANT</th>
</tr>
</thead>
<tbody>
<tr>
<td>HOSP A</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>4.0</td>
<td>0.0</td>
<td>2.0</td>
<td>10.0</td>
</tr>
<tr>
<td>HOSP B</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>5.0</td>
<td>0.0</td>
<td>2.0</td>
<td>9.0</td>
</tr>
<tr>
<td>HOSP C</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>3.0</td>
<td>0.0</td>
<td>1.0</td>
<td>4.0</td>
</tr>
<tr>
<td>HOSP D</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>6.0</td>
<td>0.0</td>
<td>1.0</td>
<td>12.0</td>
</tr>
<tr>
<td>HOSP E</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>2.0</td>
<td>0.0</td>
<td>1.0</td>
<td>11.0</td>
</tr>
<tr>
<td>HOSP F</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>2.0</td>
<td>0.0</td>
<td>0.0</td>
<td>2.0</td>
</tr>
<tr>
<td>TOTAL</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>22.0</td>
<td>0.0</td>
<td>16.0</td>
<td>48.0</td>
</tr>
</tbody>
</table>
3.9.2.6 Total Diagnoses and Procedures

Table 23 demonstrates the range for types of discrepancy by individual hospital for the total diagnoses and procedures category as follows:

<table>
<thead>
<tr>
<th>Category</th>
<th>Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>Judgement</td>
<td>6.7 - 11.4%</td>
</tr>
<tr>
<td>Classification</td>
<td>2.3 - 5.7%</td>
</tr>
<tr>
<td>Procedural</td>
<td>1.2 - 7.0%</td>
</tr>
<tr>
<td>Clerical</td>
<td>60.0 - 68.8%</td>
</tr>
<tr>
<td>Incomplete Record</td>
<td>0.0 - 5.4%</td>
</tr>
<tr>
<td>Discrepant</td>
<td>1.8 - 4.7%;</td>
</tr>
<tr>
<td>Other</td>
<td>7.9 - 22.1%</td>
</tr>
</tbody>
</table>

When all the diagnostic and procedural codes were combined, the greatest number of discrepancies was in the clerical category with the average at 64.0% across all hospitals. Classification errors, those where there is an absence of word-for-word correspondence between the listed diagnosis or procedure in the medical record and wording in the coding manuals, represented only 4.5% of coding discrepancies over all the hospitals combined. Procedural errors, where hospital policy has changed accounted for only 3.2% of errors. Discrepant records, where information documented in the chart was not in agreement and incomplete records account for 3.3% and 2.2% respectively.
**TABLE 23 - FREQUENCY OF DISCREPANCIES**

**BY TYPE OF DISCREPANCY, COMBINED DIAGNOSES AND PROCEDURES**

**BY NUMBER OF DIAGNOSES AND PROCEDURESRecorded by Original Coder**

(Excludes Non-Diagnostics)

For Individual Hospitals

Re-Abstracted Data Apr 1/86 - Mar 31/87

<table>
<thead>
<tr>
<th>DISCREPANT CODES</th>
<th>JUDGEMENT</th>
<th>CLASSIFICATION</th>
<th>PROCEDURAL</th>
<th>CLERICAL</th>
<th>INCORP. RECORD</th>
<th>DISCREPANCY</th>
<th>OTHER</th>
<th>TOTAL DISCREPANT</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>#</td>
<td>%</td>
<td>#</td>
<td>%</td>
<td>#</td>
<td>%</td>
<td>#</td>
<td>%</td>
</tr>
<tr>
<td>HOSP A</td>
<td>12</td>
<td>8.6</td>
<td>6</td>
<td>4.3</td>
<td>3</td>
<td>2.1</td>
<td>84</td>
<td>60.0</td>
</tr>
<tr>
<td>HOSP B</td>
<td>10</td>
<td>5.8</td>
<td>4</td>
<td>2.3</td>
<td>12</td>
<td>7.0</td>
<td>104</td>
<td>60.5</td>
</tr>
<tr>
<td>HOSP C</td>
<td>20</td>
<td>11.4</td>
<td>10</td>
<td>5.7</td>
<td>3</td>
<td>1.7</td>
<td>121</td>
<td>68.8</td>
</tr>
<tr>
<td>HOSP D</td>
<td>9</td>
<td>6.7</td>
<td>6</td>
<td>4.4</td>
<td>5</td>
<td>3.7</td>
<td>83</td>
<td>61.5</td>
</tr>
<tr>
<td>HOSP E</td>
<td>18</td>
<td>8.4</td>
<td>12</td>
<td>5.6</td>
<td>7</td>
<td>3.3</td>
<td>145</td>
<td>67.8</td>
</tr>
<tr>
<td>HOSP F</td>
<td>19</td>
<td>11.4</td>
<td>7</td>
<td>4.2</td>
<td>2</td>
<td>1.2</td>
<td>105</td>
<td>63.3</td>
</tr>
<tr>
<td>TOTAL</td>
<td>88</td>
<td>8.8</td>
<td>45</td>
<td>4.5</td>
<td>32</td>
<td>3.2</td>
<td>642</td>
<td>64.0</td>
</tr>
</tbody>
</table>
3.10 CODERS

The frequency of match per coder is examined over all the hospitals and for individual hospitals. For each hospital, the match per coder is determined using a combination of MRD, PP, secondary diagnoses and secondary procedure variables. Diagnostic codes are matched to four-digits, adjusted level and procedural codes to three-digits. Caution must be exercised when interpreting these results. In analyzing coder performance, all discrepancies occurring are attributed to the coder. This is applied consistently across all hospitals. For instance, even when a discrepancy was demonstrated to be due to a discrepant record, this analysis attributes the discrepancy to the coder, regardless of the origin of error. Furthermore, it may not be the individual coder's choice to code only portions of the health record, such as the face sheet. This may be an administrative decision or a policy of the hospital carried out for one of several reasons, such as backlog of charts not coded, pressure from HMRI(CIHI) to complete the chart within the required 60 days after discharge, or inadequate funding from administration, which is necessary to provide the required personnel to do the job adequately. Differences in discrepancy rates across hospitals that could be attributable to different hospital rules were not examined.

3.10.1 Frequency of Match - All Hospitals

A Pearson correlation coefficient was done to determine whether a correlation existed between agreement scores of the individual coders by their credentials and agreement score by years of experience. The results demonstrated no statistical significance between agreement scores and by number of years of experience (0.001),
between agreement score and by credentials (-0.279) or by agreement score and by credentials and number of years experience (-0.057).

The coder with the highest agreement score, 81.8%, was an ART with a one-year correspondence course and greater than twenty years experience. Agreement scores for ARTs ranged from 60.9 to 81.8% based on a denominator of diagnoses and procedures recorded. An HRT with one-year college training and seven years experience at 80.0% was next highest score. The coder with the lowest score overall (excluding the two coders that only had one and four coding incidents) was an MRL, hospital-based, with an agreement score of 52.3%. The next two lowest were also MRLs with 55.3% (four-year graduate with greater than ten years experience) and 57.5% (two-year correspondence course with greater than twenty years experience). The next lowest was an HRA two-year college trained with greater than twenty years experience at 60.0%.

Whether trained through four-year baccalaureate degrees or two-year correspondence courses, MRL's consistently demonstrated lower scores. The range for MRLs was 52.3 to 70.5% by number of diagnoses and procedures recorded. HDTs were 66.7 to 80.0%, HRTs with one-year correspondence, 66.7 to 71.5%, and HRTs with one-year college 66.7 to 80.0%. The total range for HRAs among all the coders was 60.0 to 77.6% by number of diagnoses and procedures recorded. The negative results of the Pearson coefficient indicate that coders with the least amount of training are doing a better job of coding than those with more training.
3.11 QUANTITATIVE ANALYSIS OF HOSPITAL QUESTIONNAIRE

The hospital questionnaire was undertaken at the beginning of the data collection stage to determine if there was an influence of hospital characteristics on the frequency of discrepancy. The questionnaire helped to determine the process each hospital used when a problem concerning selection of a diagnosis or procedure was encountered. The questionnaire was intended to elicit information on overall hospital procedures and did not relate to individual records. The coding supervisor at each hospital site answered the questionnaire.

Hospital D was the only hospital not downloading Admission-Discharge-Transfer (ADT) information during the fiscal year of the study. Hospitals A, D & F did not have policy and procedure manuals in place. Hospital B had one, but it was incomplete. Hospitals C & E had manuals, but updated them only as problems arose. Only the hospitals that have since switched to ICD-9CM have had major policy revisions.

Only Hospital E had regular monthly control checks for coding and abstracting, where random charts were re-coded on every coder. The other hospitals inspected a random sample only occasionally or when problems arose.

Hospitals D & F had no hospital-wide quality assurance program in place. For those hospitals that had quality assurance programs, only one, Hospital A, made recommendations relating to coding policy.
All hospitals coded from incomplete charts, especially at month end to meet the HMRI (CIHI) deadlines. The proportion of incomplete charts coded ranged from less than 10% for Hospital B to over 90% for Hospital F.

When charts are re-coded for individual hospital purposes, most hospitals send corrections to HMRI (CIHI). Hospital B sent the most, but not all, corrections, and Hospital F left this decision up to the individual coder.

Hospital E was the only hospital not to routinely code the first listed diagnosis on the face sheet as the MRD. All hospitals claimed to code from the entire chart. Hospitals C & D actually change the face sheet when appropriate changes are necessary. Other hospitals make the changes for the database, but leave the face sheet unchanged. Physicians were routinely consulted on the choice of codes only in Hospitals C & D.

Hospital A was the only hospital to leave the decision of which classification code to use, when there are two or more codes that could be used for the same diagnosis, up to the individual coders. All other hospital coders consulted the coding supervisor. When subjectivity regarding multiple diagnoses arose, Hospital A and C left it up to the coder. Hospital B did not consult the attending physician on the choice of the MRD, when two or more diagnoses were equally sufficient. All other hospitals consulted the coding supervisor.
Hospitals B, C & D coded and abstracted in two separate stages, sometimes using two or more individuals. One individual coded the diagnoses and procedures, another individual entered this information into the computer.

All hospitals claimed to code non-operative procedures. Hospital A coded all the required ones, while Hospital F coded only selected ones.

Quotas of charts coded and abstracted per day per coder ranged from sixteen to forty charts per day. Hospital F required sixteen per day, Hospitals B & E twenty-five per day, Hospital D thirty per day, and Hospitals A & C forty per day. Since fiscal 86/87, Hospital B has increased the quota to thirty-five charts per day. All other hospitals have stayed the same.

Hospitals B & C policy was to change the coding of ‘late effect’ according to the individual physician. A ‘late effect’ indicates that the condition resulted from a late or residual effect of a given disease or injury rather than during the active phase. HMRI (CIHI) guidelines state the required guidelines are >1 year. Hospitals D, E & F all claim to have different policies than ICD-9 and HMRI (CIHI) guidelines.

Training and orientation for coders ranges from informal, depending on new coders’ needs for Hospital F, to twenty-four weeks for Hospital E. Hospital A had three weeks, Hospital B had eight weeks, Hospitals C & D had twelve weeks. Every hospital
re-checked and re-coded every chart of a new employee for the duration of the training period, except for Hospital E which rechecked every chart for two weeks and then randomly.

All hospital staff participated in in-service education and available workshops, but all hospitals stated there were very few appropriate sessions. Most hospitals stated that there were not enough workshops on coding through HMRI (CIHI). There were no workshops available through the national association, from Statistics Canada, or from the Ministry of Health.
3.12 INTRA-RATER RELIABILITY

Re-coding of a 10% subsample stratified by hospital was commenced four weeks after the initial study was completed. Agreement scores were determined by combining only the diagnostic and procedural coding incidents. Intra-rater reliability (IRR) for the p.i. is determined to be 91.0%. The agreement scores ranged from 89.3 to 93.2% for the individual hospitals. All non-diagnostic variables agreed between the first set of re-abstracted data and the subset for IRR. See Table 30 for details. Because the p.i. is a trained coder, the re-abstracting is a test of reliability in the customary sense and the test/re-test determines intra-rater reliability for this study.

<table>
<thead>
<tr>
<th>TABLE 30 - INTRA-RATER RELIABILITY FOR PRINCIPAL INVESTIGATOR FOR DIAGNOSTIC CODES TO 4-DIGITS AND FOR PROCEDURAL CODES TO 3-DIGITS BY TOTAL NUMBER OF DIAGNOSES AND PROCEDURES RECORDED</th>
<th>Re-Abstracted Data Apr 1/86 - Mar 31/87</th>
</tr>
</thead>
<tbody>
<tr>
<td>HOSP</td>
<td>A</td>
</tr>
<tr>
<td>DISC</td>
<td>5</td>
</tr>
<tr>
<td>MATCH</td>
<td>69</td>
</tr>
<tr>
<td>TOTAL</td>
<td>74</td>
</tr>
<tr>
<td># CODES</td>
<td></td>
</tr>
<tr>
<td>% MATCH</td>
<td>93.2</td>
</tr>
</tbody>
</table>
CHAPTER 4 - DISCUSSION AND CONCLUSION

4.0 INTRODUCTION

This chapter discusses the results of the analysis and relevant aspects of the study not covered elsewhere. It also provides the conclusion.

4.1 AGREEMENT SCORES FOR DIAGNOSES

The definition of the MRD in Canada and the principal diagnosis in the United States differ significantly in concept; however, the agreement scores are comparable. Although guidelines differ, the process for selection of the code is similar. Within this discussion, only the four-digit diagnostic scores at the adjusted level are considered, i.e., all the converted HMRI data are included in the analysis.

4.1.1 Most Responsible Diagnosis

The overall agreement score for the MRD to four-digits was 61.4%, with the individual hospitals ranging between 52.0 and 69.6%. The three Institute of Medicine studies (IOM, 1977a; 1977b; 1980) demonstrated agreement scores of 65.2%, 57.2%, and 63.4%, the Armstrong et al. (1992) study demonstrated an overall level of agreement in obstetrical
patients of 65%, and agreement scores from the Doremus & Mickenzi (1983) study ranged between 52.3 and 65.2%, all within a similar range to that found here. However, several studies demonstrated both higher and lower levels of agreement. The New Jersey Department of Health’s study (1978) and the George & Maddock’s study (1979) demonstrated higher rates for the principal diagnosis to four-digits at 72.6% and 80% respectively. However, the Armstrong study found much lower levels of agreement: newborn charts agreed to only 58% with individual hospital ranges between 12 and 89%.

In the studies reviewed, the Ontario study (OHA, 1991) demonstrated one of the higher scores for the MRD at 80.7%. However, the number of digits used in the agreement score was not stated in the study, i.e., whether to three- or four-digits. Because the CMG reassignment only demonstrates a 14.7% difference, and because only three-digit coding is necessary in CMG assignment, it is suspected that this agreement score was based on three-digits.

Agreement scores for the MRD increased by 10.7% overall when lesser levels of specificity were used. Scores for the MRD were more reliable at the three-digit level, which is consistent with other studies. The three-digit agreement score for the MRD in this study was 72.1%, with individual hospitals ranging between 62.4% and 79.4%. This is consistent with the IOM (1977a; 1977b; 1980) studies which showed increases of 9.4% overall when the three-digit level was assessed. The NJDH (1978) also stated that reliability increased at the three-digit level. Hsia et al. (1988) and Corn (1981) compared DRG’s to the three-digit level and demonstrated agreement scores of 72.9 and 81.5% respectively.
The p.i. chose not to code through CMG's as the Ontario study showed that up to 60% error in original coding of secondary diagnoses still resulted in the same CMG being chosen. There was a great discrepancy in the secondary diagnoses coding because the "type 3" discretionary codes were removed before agreement scores were reached. "Type 3" codes are those conditions for which a patient may, or may not, have received treatment, but which did not significantly contribute to the patient's length of stay (HMRI, 1990). It is the "typing" of codes that determines the assignment of CMGs. However "type 3" codes do not equate directly to a reassignment of a CMG. The CMG process rarely uses discretionary codes in the assignment of CMGs.

There were 8.4% of MRDs where there was agreement to the four-digit level between the p.i. and the original hospital coder, but the MRDs were coded as secondary diagnoses by the original hospital coder. In 19.5% of cases, the MRDs were not coded by the original hospital coder at all. Approximately 20% of the MRDs are not entering the database whatsoever, either in the MRD or secondary diagnostic positions. This should be a great concern for both the provincial and national databases.

4.1.2 Secondary Diagnoses

Agreement scores for secondary diagnoses to the four-digit level are higher overall than for the MRD at 67.4% of the secondary diagnoses recorded. The range by hospital was from 52.4 to 72.3%. However, this agreement score does not consider the 2,500 secondary
diagnostic codes for the 606 cases coded by the p.i. that were not coded by the original hospital coders. This under-recording of diagnoses may result in the agreement score of secondary diagnoses in this study being too high or upward biased.

The scores for the secondary diagnosis in this study are roughly consistent with, or slightly better than, other studies. The NJDH study (1978) demonstrated agreement scores for secondary diagnoses to the fourth digit between 52.9 and 66.5%. The differences in the denominators are twofold: eight diagnoses per case were collected in the NJDH study and in this study there were fifteen secondary diagnostic positions. A more important difference is that in this study, a diagnosis was included in the denominator only when the original coder specified the diagnostic code.

Agreement scores for secondary diagnoses in the OHA (1991) study were much lower than the MRD at 40%. Again, the difference in denominators in this study may account for this difference in results. If the "type 3" errors in the OHA study were maintained in the analysis, it may have produced an even lower agreement score for secondary diagnoses.

4.1.3 Total Diagnoses

The total diagnostic agreement score to four-digits for the combined MRD and secondary diagnoses is 65.7% by number of diagnoses recorded, with individual hospital ranges from 58.8 to 71.6%, consistent with the other studies reviewed in Chapter 1. These errors are significant when one considers the uses of the data.
4.2 AGREEMENT SCORES FOR PROCEDURES

The definition of the principal procedure is similar in both the United States and Canada and the agreement scores between the countries are comparable.

4.2.1 Principal Procedure

Agreement scores for the principal procedure at the three-digit level are slightly higher than for the MRD, at 64.8% of cases. The individual hospitals ranged from 56.3 to 75.6%. The score was within a similar range to that found in studies in other studies. The NJDH (1978) study found agreement scores of 66.4%, Doremus & Mckenzie (1983) found 58.4%, while the IOM (1977a; 1977b; 1980) studies reported 73.2%, 78.9% and 71.4%.

At the two-digit level, when less specificity was required, the overall rate of agreement increased to 72.8%. The range per hospital increased from 61.4 to 85.4%. Both Corn's (1981) and the OHA (1981) studies demonstrated higher scores for the PP at the two-digit level at 83.4 and 88% respectively. However, it needs to be mentioned here that all procedure codes require at least a third digit.

For 17.5% of PPs, there was agreement to the three-digit level, but the PP was coded as a secondary procedure. A further 9.7% of PPs were not coded by the original hospital coder at all. Thus, approximately 10% of principal procedures are not even reaching the provincial or national database.
4.2.2 Secondary Procedures

Agreement scores for secondary procedures at the three-digit level were much higher than for the PP, at 80.7% of procedures recorded. The individual hospitals ranged between 69.4 and 84.1%. There were 1,259 extra procedures coded by the p.i. that were not included in this analysis. Again, therefore, these agreement figures may be upward biased. Secondary procedure scores reported in the NJDH (1978) study ranged from 44.0 to 59.3%, much lower than the results of this study.

4.2.3 Total Procedures

Combined procedural data had an agreement score to three-digits of 73.4% by number of procedures recorded and individual hospitals ranged from 68.4 to 78.7%, quite consistent with similar studies reviewed.
4.3 TOTAL SCORES BY NUMBER OF CASES

When determining total scores by the number of cases, the agreement scores drop dramatically. Case agreement occurred only for those cases where every diagnostic and every procedural code agreed to the highest digit possible, i.e., four-digit diagnostic and three-digit procedural coding. (The extra diagnoses and procedures coded by the p.i. were not included in this analysis, so the results may still be upward biased.) This study demonstrated that only 25.6% of cases match in every code. The case agreement scores for individual hospitals ranged from 17.0 to 39.8%. Similarly, Barnard & Esmond (1981) found that in one hospital, overall agreement was 23.1% to the three-digit diagnostic level.
4.4 AGREEMENT SCORES FOR NON-DIAGNOSTIC DATA

The agreement scores for combined non-diagnostic data are high at 92.4% by case. For individual hospitals, the non-diagnostic variables range from 96.4 to 99.8% by case. These scores are consistent with similar studies, such as the NJDH (1978) which ranged between 96.9 and 99.7%, Beard et al. (1988) at 93.4%, the IOM studies (1977a; 1977b; 1980) with scores of greater than 90%, and the OHA (1991) between 93 and 100%.
4.5 AGREEMENT SCORES FOR DIAGNOSTIC CATEGORIES

Some diagnostic categories for the MRD demonstrated higher agreement scores than others at the four-digit level. Category 6 (codes 520-579, Diseases of the Digestive System) demonstrated agreement scores greater than 75%. Categories 2 (codes 240-289, Endocrine, Nutritional & Metabolic Diseases, Immunity Disorders) and 5 (codes 460-519, Diseases of the Respiratory System) demonstrated agreement scores close to or less than 50%. The Connell et al. study (1984) demonstrated that for diabetic codes, 23% could not be determined with certainty and that a choice of code could not be determined for more than 50% of diabetic cases. Similar choices had to be made in this study. Only a 50% accuracy within any category should be a concern for any use of the data.

As stated in section 2.6.3, initial categories were chosen by diagnosis, not procedure. In the PP category, Category 10 (codes 800-999, Injury and Poisoning) demonstrated agreement to 75%. Categories where procedures were missed because they were not documented on the face sheet may show a higher level of agreement because, again, codes were matched only to those procedures coded by the original hospital coder. For example, documentation for a procedure, such as stomach lavage, which could be associated with an overdose of drugs (Group 10) was often not recorded on the face sheet and therefore missed by hospital coders who code only from the face sheet. The results may have produced a higher level of agreement score than actual.
4.6 TYPE OF DISCREPANCY

The type of discrepancy was determined separately for the MRD, secondary diagnoses, the PP, secondary procedures, overall discrepancies and non-diagnostics.

4.6.1 MRD

Within the MRD category, 73.1% of discrepancies over all the hospitals were clerical. Clerical errors include misuse of the classification system by failing to use mandatory codes and/or mandated specificity and reliance on the Alphabetic Index with no cross reference to the Tabular List. This can lead to failure to apply the inclusion or exclusion notes. Clerical discrepancies also include transposing numbers, erroneously recording information, coding from memory, and mis-interpreting the information in the record, i.e. coding perineal instead of perianal. Coding from memory may produce a consistently incorrect code. Similar studies demonstrated these kinds of clerical errors (Barnard & Esmond, 1981; Massanari et al., 1987). Other clerical errors include not reviewing the entire record for specificity, i.e. the wrong selection of the fourth digit, or for not capturing all the relevant information in the chart. This discrepancy is associated with face sheet coding. Both misuse of classification system and face sheet coding are consistent with types of discrepancies as determined by other studies reviewed (IOM 1977a; 1977b; 1980; NJDH, 1978; Doremus & Mckenzie, 1983; OHA, 1991).

Within the MRD category, 5.6% of cases overall demonstrated classification discrepancies where there was a lack of word-for-word correspondence between the coding manuals and the written documentation in the charts. This is fairly significant. Physicians
and other health care professionals need to be educated on the ICD terminology and charting procedures.

Discrepant records were 6.8%, lower than in the OHA (1991) study. Only 2.6% discrepancies were judgement errors, lower than other studies. Similar studies also demonstrated judgement errors (IOM 1977a; 1977b; 1980; NJDH, 1978; Barnard & Esmond, 1981). Both Corn (1981) and Connell et al. (1984) studies reported sequencing errors.

4.6.2 Secondary Diagnoses

Secondary diagnoses demonstrated 66.0% clerical discrepancies. Again, misuse of the classification system represented the greatest percentage of this discrepancy. There were 21.3% of coding discrepancies for secondary diagnoses that could not be classified into any one category. Similar to the MRD category, 5.0% of secondary diagnoses had no word-for-word correspondence between the written diagnosis in the chart and the coding manuals. Again, this is fairly significant. Physicians and other health professionals need to be encouraged to use the ICD terminology.

4.6.3 Principal Procedures

The PP category demonstrated that the greatest percentage of discrepancies was in the selection of the principal procedure at 46.9%, consistent across all hospitals. These discrepancies are the result of subjective selections where it is not clearly determined which procedure to code. For instance, there may be more than one procedure appropriate for the principal procedure, but only one must be chosen.
The second greatest percentage of discrepancy for the PP was 42.9% over all hospitals for clerical errors, which include not reviewing the entire record. The information was documented in the chart, but it was not coded. This finding is similar to the IOM studies (1977a; 1977b; 1980). Other clerical errors included misuse of the classification system.

Only 1.1% of the PP discrepancies demonstrated a lack of word-for-word correspondence between the listed procedure in the record and the coding manual.

4.6.4 Secondary Procedures

Discrepancies for secondary procedures differed significantly from those determined for the PP. Secondary procedures demonstrated 69.3% clerical discrepancies involving a lack of review of the record as well as misuse of the classification system. Most hospitals did not have discrepancies in the chart noted for secondary procedures. That is, when the procedural information was documented, it was consistent throughout the record.

4.6.5 Overall Discrepancies

Similar to the findings of earlier studies in both the United States and Canada, the three most pronounced types of discrepancy involved clerical errors, which included an incomplete review of records and misuse of the classification system; erroneous judgement in the selection of the MRD and PP; and differences in hospital administrative procedures. Clerical errors account for 64.0% of discrepancies overall; 8% were judgement errors and 3.2% of discrepancies arose because of changes in hospital policy relative to the guidelines of HMRI, the ICD, or Statistics Canada. It is also important to note that in this study, 4.5% of
errors were due to a lack of word-for-word correspondence between the written diagnoses and procedures in the chart and the wording in the coding manuals. In such situations, the documentation in the chart did not conform to the standard terminology of the ICD.

4.6.6 Non-Diagnostics

Although the total number of discrepancies is small, the non-diagnostic category demonstrated clerical errors as the greatest type of discrepancy at 45.8%. This appears unusual given that the ADT transfer of information occurred in all but one hospital at that time. It appears that errors could be generated in other departments, such as Admitting, with data intended for the ADT system, before reaching the HMRI(CIHI) database. Discrepant records, where birthdates, admission and separation dates, and discharge status were inconsistent within the written documentation of the record, accounted for 33.33% of discrepancies.
4.7 CODERS

There are many variables that influence individual coder performance, ranging from individual differences in coder training, knowledge and skills to differences in hospital policies and procedures, to problems with the classification and database systems. Agreement scores, coder training, testing, and credentialing and continuing education are reviewed in this section.

4.7.1 Agreement Scores for Coders

Similar to the Armstrong et al. (1992) study, this study also demonstrated a variation among the coders. The total overall agreement score for coders ranged from 52.3 to 81.8%. When examining differences in agreement scores by years of experience, by credentials, and by years of experience and credentials, there was no statistical significance demonstrated. Although not statistically significant, this in itself is clinically significant, because the health record association, as well as the health service industry as a whole expects that, given the differences in level of training, there would be a difference. Whether trained by correspondence course, one- or two-year college diploma, or baccalaureate degree, there is no statistical difference in the level of coding. Furthermore, there are levels of inaccuracy within this database that are unacceptable for a provincial or national database.

In this study, the predominant type of discrepancy determined in all categories and across all categories was clerical. In particular, it was noted that misuse of the classification system was predominant. Clerical errors affect the data input and ultimately the accuracy of the entire database. Beard et al. (1988) claim that inter-observation variation is a source of
bias that could affect the outcome of any study and render it invalid. This study concurs with Beard et al. that agreement scores between 52.3 and 81.8% could possibly render any study invalid.

4.7.2 Coder Training

It is recognized that the accuracy of coding lies within the health record. Coders should be able to infer codes based on the documentation within the record. Health record coders need to be educated enough to challenge the diagnoses or documentation within the chart. When guidelines clearly state that the Most Responsible Diagnosis is that which determined the longest length of stay, the authority should lie with the documentation within the chart, not that which the individual physician documents on the face sheet. It must also be recognized that a coder does not have clinical evidence other than that which is documented in the health record. The entire database relies on health record coders to maintain accuracy and consistency of the collection of the data based on documentation in the health record.

There is a great deal of controversy over whether health record professionals are under-utilized or under-trained. Health record professionals consider themselves under-utilized; however, the perception of other health professionals is that they are under-trained. This study concurs with the position that health record coders, regardless of credentials, are under-trained. Coder training must reflect a level of responsibility relating to the clinical work site. Coding requires task-related duties and an increased understanding of medical terminology. George & Maddocks (1979) also claim that coders lack an understanding of
medical terminology. Coding and record management skills are different and need to be recognized as such. The status of a health record coder is low. The skills necessary to carry out the tasks are considered clerical, yet the coder requires the skills to challenge the documentation within the record. Health record coders need a revision in curriculum and standards to reflect the skills necessary for a coding professional.

Currently, there is no requirement specifically for the training of coders. There are, however, one- and two-year college-based training courses for health record technicians and health record administrators. The difference between the training for technicians and administrators is mainly administrative. The coding skills learned are identical in the two types of training. Coders with insufficient skills are being credentialed in an industry where standards are too low to meet the quality necessary to maintain an accurate and consistent database. Education preparation for the health record coder needs to be reviewed.

There are no statistics available on the national coding examinations for Canada, because the national association does not test specifically for coding skills. However, the average percent on the coding section of the American national examination is 68%. Scores this low are unacceptable as the basis of coding for a provincial or national database and were recognized as such by AHIMA. More stringent examinations with particular reference to coding sections were implemented a few years ago.

The only requirement of the CHRA/CCHRA is that the coder be a member of the CCHRA to at least the associate level. In addition, there is no mandate to require hospitals to conform to the CCHRA’s requirements. Hospitals are free to hire anyone to code. HMRI/CIHI has no requirements for coder certification, but assumes that coders have an
adequate understanding in medical terminology. No testing of skills or mandate for
certification is necessary. The standards for credentialling and testing coders is insufficient.
The current process of training, examining, and certifying health coders is not adequate.

Coding skills need to be reassessed and applicable courses, testing and credentialling
need to be developed and implemented. Another change of name title as has recently been
suggested by the CCHRA, is simply cosmetic, and will not solve the problem. Whatever the
name may be, involvement in the health care service industry of the 1990’s and beyond is
contingent upon the health record personnel’s ability to communicate, to understand and to
manage the data within a fast growing technological industry. Coder testing needs to occur at
least annually and the testing process should be connected to the credentialling process.
CCHRA needs to gain the power to remove credentials.

HMRI (CIHI) needs to be more actively involved in coder training programs, coder
certification, and the testing of coders. This can be achieved through affiliation with the
CHRA/CCHRA and its accreditation process for training programs for health record
personnel. Both HMRI (CIHI) and the CHRA/CCHRA are logical choices for developing
this process. Educators within the health record programs, the CHRA/CCHRA and HMRI
(CIHI) must jointly take the leading role in designing curriculum to meet the needs of the
industry and to ensure an accurate and reliable database.

A four-year degree program for health record managers could ensure more grounding
in computers with special attention to health data systems, systems analysis, and technical
aspects of information systems. Other core courses could include quality assurance,
evaluation, research methodology, epidemiology, strategic planning, forecasting, health
policy, biomedical ethics, labour relations, finance and legal issues. However, these skills, although necessary for health record managers, are not necessary for health record coders. Coding is a mundane, tedious, clerical and task-oriented job. An individual pursuing a management career would probably not be fulfilled in such a task-oriented position. Managers could become bored with such a task. The results of this study indicate this may be so.

The CHRA/CCHRA, along with HMRI (CIHI), needs to initiate educational change for coding requirements to ensure data accuracy within the entire health care industry. A task force on educational requirements needs to be initiated by the CCHRA and HMRI (CIHI). Membership on this task force should include key stakeholders within the healthcare industry from medicine, finance, administration, nursing and other professions. The mandate of this task force should be to identify and to forecast what the future needs of the industry are, and where the gap lies between the future needs and the current academic curricula at present. Coding skills in particular need to be assessed and curricula determined to best suit the needs of the industry. Individuals qualified to assume leadership roles in making the necessary changes need to be identified. These individuals should have the ability to view the healthcare system as a whole. Once the needs are determined, the resources available to implement academic program development need to be developed and a time frame within which to complete the process needs to be set. It is through such a task force that health records personnel maintain their focus on health records while at the same time move into the information technological explosion of the healthcare industry with credibility and legitimacy.
The INFOCUS task force of the CCHRA needs to complete the task originally set out for it, i.e., to enhance curriculum and to determine the academic facilities to deliver those education programs. They also need to focus on strategies to define practitioner competencies. However, this process should not involve health record personnel alone. INFOCUS needs to reach out and solicit advice from other health professionals, especially those who have an interest in either the documentation in the chart or the accuracy of the database, such as medicine, nursing, administration, finance, planners and researchers. They need to be included as active members of the educational advisory boards and the board of the professional association. Soliciting advice from others is one way to view the entire picture and broaden the scope of the health record profession. The addition of other health professionals on the board of directors of the CCHRA could provide an avenue for this necessary liaison to occur. Linking the association with other professionals could lead to credibility and legitimacy of their role within the health care industry. It is necessary so that other professions recognize the importance of the expertise held by this profession.

Once the education requirements for a coder "specialist" are determined from the task force, a training program specifically for ancillary health coders needs to be developed. Specific coder certification for credentialling of coders has been recently instituted in the United States. Yet the agreement scores between the two countries are similar. However, AHIMA has associated the poor quality data with the training and testing of coders and has recognized that a coder specialist is essential for maintaining an accurate and consistent database. Similarly, coder certification needs to be implemented for coder specialists in Canada. Currently, coder certification in the United States consists of testing and training of
coders. The CCHRA also needs to recognize that the quality of the data could be improved not only through better training, but also through more stringent testing and credentialling of coders. A specified designated category for coders should be established within the CHRA membership. The credential designation needs to reflect the identity of the professional, e.g., from a Health Record Administrator or Technician to a Coder or Coder Specialist. This would only be designated after successful completion of the national examination.

4.7.3 Testing of Coders

The role of the CCHRA in examining and certifying health record coders needs to be reviewed: more stringent national exams and requirements for eligibility need to be enhanced. National examinations need to reflect the accuracy and reliability expected of a coder. Sections on coding procedures need to be included as part of the regular testing of skills within the national exam. This needs to include coding from actual health records. A refresher course for upgrading coding and diagnosis skills offered through the CCHRA in association with the Canadian Hospital Association provides a certificate if the coder achieves an overall average of 60%. Can an industry endorse accuracy of a database to 90% or better if they are enabling passing grades within training or refresher courses to be as low as 60%? Passing grades for coders need to be much higher than 60%. Passing grades should reflect the minimum acceptable level of error that we expect within our database. Educational standards need to be tightened so that credentials cannot be attained without passing a stringent examination.
The CCHRA needs to ensure that the testing of coders goes beyond that established only at the time of credentialling. Examining coding skills is necessary to determine the acquired knowledge of coding skills. Ongoing examination enables the industry to ensure these skills are kept up. The standards need to be increased.

Both intra- and inter-rater reliability checks need to be implemented on a regular basis for all coders to ensure that the database actually reflects the real incidence of disease and a consistency and accuracy of the data. Reliability checks need to be maintained at as high a level as possible, such as between 90 and 95%. Neither HMRI (CIHI) nor the CHRA/CCHRA currently require the testing or ongoing testing of coders. HMRI (CIHI) needs to become more actively involved in monitoring the coding and abstracting process by incorporating regular and ongoing intra- and inter-rater reliability checks as mandatory procedures.

4.7.4 Credentialling and Continuing Education

Credentials provide entry into a profession and designate professional identity. However, this need not be a lifelong process. Systems change, skills change, and abilities change; therefore, credentialling should reflect an adeptness with these changes. Recognition for professional upgrading is key to maintaining a highly effective credentialling process.

The CCHRA must engage in the preparation and implementation of continuing education programs for its membership. The association must recognize that education is a lifelong learning process. The health care system is not static; therefore, ongoing mandatory education components need to be maintained. Ongoing mandatory continuing education for
health record personnel and health coders is essential in a health care industry where changes in technology, new breakthroughs in research and disease information, and changes in ICD classification are constantly occurring. Certification standards need to be tightened to ensure coders are maintaining a level of education necessary to do the job.

The association needs to attain self-regulation, i.e., to be able to give or take away credentials, dependent upon the member’s conforming to credentialling standards on an ongoing basis. Ongoing maintenance of credentials should be reviewed. The CCHRA needs to be able to remove the credentials of coders if the standards are not maintained. Coder upgrading and continuing lifetime testing with the implication that health coder credentials can be removed, if applicable, needs to be established into the current system. The costs associated with upgrading and monitoring must also be addressed within the industry, either by raising tuition fees or by raising membership fees.

Continuing education, such as workshops, in-service education, and access to other educational programs needs to be made more readily available for health record coders, especially for those who live in remote areas. The CCHRA needs to make every effort to provide alternate ways of delivering education for its membership. Just as society, in general, utilizes technology for educational purposes, so should the health record association. Education program options include distance education, the Knowledge Network, the Open Learning Institute, teleconferencing, video-conferencing, and local courses at community colleges. Lytle, Lytle & Youmans (1994) describe the consumer of today. The rationale for whether or not to engage in mandatory education is determined by choice, convenience, conservation of time, and cost.
HMRI (CIHI) training seminars and workshops on coding also need to be more accessible. Delivery of these training courses could be implemented through any one of the above options. It was also suggested by several hospitals in the study that information from these workshops should be disseminated to every coder in the province, not just to those attending the workshops. A logical method of delivering coder training programs would be to combine the efforts of HMRI (CIHI) and the CHRA/CCHRA.

The OHA study (OHA, 1993) has recognized that active membership is essential in both national and provincial associations. Membership in both associations in the same member category could provide a link between provincial and national associations. This could be provided through one application to the national level. Provincial and national associations can gain in efficiency in operations and sharing of common goals. Dual membership could also provide a more common ground for standards setting and continuing education efforts. Provincial associations could monitor membership on behalf of the national association. National education programs could be implemented through the provincial associations. Continuing competence of the health record coders could be accomplished through dual membership.
4.8 SUMMARY OF RESULTS

Guidelines for a national database should emphasize quality, effectiveness, and outcome measurement. Accuracy and consistency of individual coding elements may have a magnifying effect on the aggregate database. There is controversy in the literature about the acceptable level of specificity of coding elements. Coding may be more acceptable to only the three-digit level when CMG's are used. Barnard & Esmond (1981) state that three-digit coding only is necessary for DRG/[CMG] assignment, and therefore, the use of one or two decimal places is irrelevant. However, three- and four-digit coding is essential for some purposes where there is a need for greater specificity, such as planning, evaluation or other such purpose within the health care system. Coding strictly for CMG assignment may present adequate information for administrative and even overall planning needs, but does not maintain the specificity necessary for individual or small group patient research or even for incidence studies.

The purpose of coding needs to be stressed here: the coder's primary responsibility for coding resides with the hospital site and with individual hospital researchers, not with the provincial or national database where aggregate data are used for planning and research. However, when we discuss the accuracy and consistency of our provincial or national databases, we include the many uses of the data throughout the entire health care system. It needs to be stressed that it is the MRD variable that is most often used for research purposes and is key to the assignment of CMGs. Although three-digit coding may be acceptable for some purposes, it may not be for others. The database needs to be effective for all the purposes and activities previously outlined.
4.8.1 Overcoding

Overcoding is a controversial topic within the health coding field. However, some of the diagnoses and procedures coded by the p.i. that were not coded by the original coder, were not just minor details but major diagnoses and procedures relevant to the patient's stay. These codes were not included in the forgoing analyses, as agreement scores were only determined for those items coded by the original hospital coder.

Although overcoding may be subjective for diagnostic coding, it should not significantly affect procedural coding. If a procedure was done, it should be recorded. It is important to note here that non-operative procedures are discretionary for hospitals to code, but all the hospitals in the study acknowledged coding them.

4.8.2 Under-recording of Diagnoses and Procedures

Under-recording of these procedures may have significantly upward biased the agreement scores reported in this study. The procedures that were missed were not just non-operative procedures, but often major ones. As in the OHA study (1991), procedures that were performed in the operating room were often not recorded on the face sheet. Not all major procedures were recorded in the chart. When procedures were recorded, they were often not recorded to the proper level of specificity.

There is a great concern that under-recording of important information relevant to the patient's hospital stay is occurring and this ultimately affects the use of aggregate data for research purposes. Both George & Maddocks (1979) and Roos et al. (1982) demonstrated
under-recording of both diagnoses and procedures. Similar to Roos et al. (1982), there was more consistency in recording major diagnoses and procedures, than in recording minor ones.

4.8.3 Is This Acceptable?

There is a concern in the under-recording of major diagnoses and procedures. From this study, it appears that ten percent of significant procedures and twenty percent of MRDs do not even enter the database. As with the Connell et al. study (1984), the use of diagnostic codes for case finding in this study may be highly imperfect. Since it is the MRD that is used in case retrieval, and since this study shows that approximately 20% of MRD’s do not even make it into the database, the results could pose a serious problem. George & Maddocks (1979) refer to the possibility of inaccuracy for planning purposes when aggregate data are used. One of the more noticeable problems in the Armstrong et al. (1992) study was that coders were coding the charts as normal, when in fact, a medical problem not only existed but was also documented.

Similar to the findings of other studies, this study noted that several errors were due to physicians not reporting procedures (Lloyd & Rissing, 1985; Roos et al., 1982; George & Maddocks, 1979). Several procedures that were missed were done in the operating room, but not by the primary physician. The findings were similar here. Another common error, as noted by Doremus & Mickenzie (1983), was that if procedures were listed on the face sheet, the order recorded on the face sheet was abstracted, whether the major operation was listed first or not. Massanari et al. (1987) noted that most infections were not recorded on the face sheet but lab reports indicated infections were present. In this study, some were not even
recorded in the chart by the physicians. Physicians are encouraged to better document the patient's events both within the chart and on the face sheet. It is recommended that the terminology used by physicians conform to that of the International Classification of Diseases (ICD).

Even though there was very little variation among the diagnostic categories, the low agreement scores were still a concern. Although no statistical significance was noted for differences between the diagnostic categories, a 50% agreement score in any category is not within an acceptable range to produce reliable data for the numerous purposes discussed earlier.

Agreement for non-diagnostics was within an acceptable range; however, diagnostic and procedural data did not fare as well. Although agreement scores are similar to other studies, in general the results are not acceptable. Again, errors are significant when one considers the uses of the data.

The reliability of the data depends on the range of use and the degree of accuracy necessary for the purpose in question. Whether the three- or four-digit level of accuracy is acceptable, agreement scores in the 60 to 70% range are not acceptable. If individual diagnostic data are used to make inferences about the health status of populations, such low agreement scores will be problematic. This being the case, the real needs of the nation may not be being addressed. The findings in this study are similar to what Lloyd & Rissing (1985) found: that there is a substantial difference between the HMRI (CIHI) database and the health record. The applications of this study are similar to that reported by Roos et al. (1982), that with appropriate caution, the data can be meaningfully investigated using the
HMRI (CIHI) database. However, unless prudent discretion is undertaken, the data are unsuitable for most purposes outlined.

The majority of the discrepancies can be attributed to clerical errors; however, confounding variables, such as coder training, documentation within the record, hospital policy and classification and system inadequacies also have an effect on coder performance. However, the high level of clerical errors should be a concern to both the CHRA/CCHRA as well as the HMRI (CIHI). Both have a vested interest in ensuring that steps are taken to alleviate such low scores. Both should work to tighten the standards for coders and to ensure that these standards are maintained. Curricula revision, ongoing testing, and methods of continuing education should be priorities for both bodies. If standards are not met, credentials should be removed.
4.9 DOCUMENTATION IN THE CHART

A confounding variable that influences coder performance is the inadequate documentation in the chart. A coder cannot often derive the importance of a diagnosis merely by interpreting the physician's consultative or progress notes. There were several documentation problems with the health records in this study, consistent over all the hospitals. This section reviews the challenges in the chart relative to the function of the coder.

4.9.1 Discrepancies in Documentation

There were several blatant discrepancies found in the documentation of the charts reviewed during the study. Numerous physicians and other professionals, such as nurses, dieticians, occupational therapists and social workers, treat patients and record the activities or events of the patient. At times, these recorded events are inconsistent, which leads to discrepancies within the record itself. These discrepancies inevitably affect coder performance. For instance, medical consultants sometimes disagree on a diagnosis or on which diagnosis should take precedence over another. If the attending physician has not sorted this out in the discharge summary, the coder is forced to. Some of these discrepancies may not be significant enough to alter the code or coding sequence, but others may be.

Examples of documentation from charts in this study are as follows:

- 'Patient no longer drinks' was written in the history, but the patient was admitted for alcoholic gastritis.
- A physician has noted 'the patient has hypertension'. Another note by the same physician, stated 'no clear history of hypertension'.

A note at end of a cardiology report reads ‘It is terribly unfortunate that [this patient] infarcted and passed away prior to us being able to do anything for this patient’. However, the patient was discharged alive and returned for another admission!

• Insertion of a pacemaker was noted on the face sheet, but there was no follow-up evidence throughout the rest of the chart.

• Cancer of the bladder was the MRD on the face sheet, but history shows patient was ‘previously treated with radiation’. There was no treatment for cancer during this particular admission. The current admission was for chronic obstructive lung disease, not for cancer.

• Right decompression (17.39) was written on the face sheet, but right palmar fasciectomy (94.35) was the procedure recorded in the operative report. The same physician completed both forms.

• The physician’s notes state that ‘patient is amenorrheic X2 years’. However, the nurses’ notes state that ‘patient takes Provera to control menstruation’.

• Two consultants - one states ‘there is no evidence of diabetic retinopathy’, the second states ‘patient has diabetic retinopathy’.

• One consultation suggests ‘the patient had a silent myocardial infarction’; however, neither the ECG interpretation nor the biochemistry tests confirm this.

• A patient had been diagnosed with a chest condition but had an abdominal x-ray with no backup documentation to determine the reason.

Throughout the charts, terminology was often mis-used between the recording professionals, causing conflicting information within the chart. Either one professional uses one term, another uses the other, or the same professional uses both terms interchangeably. The coder has to decipher this information in order to determine the appropriate code. Some examples of discrepant terminology throughout the chart and common to all hospitals are:

• Hypertension and hypotension
• Perineal and perianal
• NIDDM and IDDM

In general, charting could be improved for all the hospitals examined.

4.9.2 Missing Information or Mis-filed Reports

Some consultants dictate very in-depth, lengthy reports, while others scribble only a few notes in the chart. Often the discharge summary or face sheet is written in haste,
sometimes not even reflecting the major events of the specific chart. Some physicians are consistently worse recorders than others. This study established that some operative reports did not give the specificity needed to complete a mandatory procedural code, i.e. whether the hernia repair was direct or indirect. In a case like this, the coder needs to determine the correct level of specificity. According to the questionnaire, not all hospitals contact physicians to verify the information before determining the correct code. It was also determined that medical essentials were not always given, such as whether hypertension is benign or malignant. Documenting every clinical event appears to be an impossible task. However, how and what to document is a matter of habit and choice of the documenting professional.

Several of the charts reviewed had no histories, no discharge summaries, or no procedure reports. These were for health records that should have been completed five years earlier! Minor procedures, such as sigmoids, especially those not done in operating rooms, were usually not recorded by the physician and often missed by the hospital coders.

In several charts, a note on the face sheet stated that the discharge summary was dictated, but the document was not found in the chart. This could be a result of a mis-filed report or a dictated report gone astray in the dictating system. However, one must wonder why it was not replaced by something else! There were several charts with misfiled documents. Some charts had the wrong history or the wrong discharge summary filed in the charts, others had misfiled x-ray and out-patient reports.

There was also frequently insufficient documentation on patient injuries to determine the extent and nature of the injury. For example, one chart stated that ointment was applied
to an abrasion on a patient's cheek in the post anaesthetic recovery room (PARR) notes, but there was no evidence of the incident elsewhere in the chart and no documented reason as to when and why it happened. This may have resulted from a surgical complication, an injury in the PARR, or some other reason unknown to the coder. Missing information does affect coder performance.

4.9.3 Repetition / Standardizing Forms / Order and Logic of the Record

There were several other documentation problems found throughout the hospital chart. In most of the hospitals visited, the repetition found in the charts was not only inappropriate, but also unnecessary. One problem with the length of the record is that almost all professions summarize the patient's history and diagnostic evaluations to determine treatment. On several occasions, professionals recorded the same historical events. No hospital had a standardized form for the patient's history. The coder was faced with a chart loaded with repetition, poorly documented events, and occasionally discrepant events. This was time-consuming for both the coder and for each professional documenting information already in the chart.

Standardization of health record forms is necessary to eliminate duplicate reporting of information. Professionals do not need to repeat the history of a patient. The history does not change. It would be appropriate for only one report to contain the historical information to cut down on the amount of duplication in the documentation within the record, the excess amount of information, and the amount of time in extracting the information. The form should give a clear picture of the necessary terminology needed for that patient. This could
be utilized by all professionals and each person can add or change information on this one form. When there is disagreement about interpretation of past events, it could be noted in the chart along with an explanation.

Reliability of the database appears to be associated with the order and logic of the individual health record. The findings of the IOM (1980) study stated that reliability could be associated with the general organization, orderliness and logic of the information recorded in the chart. It was found in this study that charts were easier and faster to code when the record was in order, the information was complete, there was no duplication and there were no discrepancies within the recorded events. Physicians need to accept more responsibility for coding diagnoses and procedures and for completing the documentation within the chart (George & Maddocks, 1979; Wood et al., 1989). However, hospitals in general and health record departments need to accept the responsibility for the organization and logic of the chart, including relevant forms for physicians and other professionals to utilize.

4.9.4 Better Documentation by Physician and Other Professionals

It is important to better train coders and maintain training programs specifically for coders. However, Hendrickson & Meyers (1973) noted that better training for coders may not guarantee a greater degree of accuracy, given the nature of the coding process. A coder must rely on the information in the chart. Medical schools and schools for the training of health professionals differ on charting procedures.

Education programs for physicians and other professionals that record in the record are necessary for better documentation within the record. All professions need to be made
aware of coding procedures and standard approved terminology. Discrepancies, missing information and repetition are all consistent with a poorly documented chart. It has been noted in this study that 4.5% of errors are due to a lack of word-for-word correspondence between the written terminology in the chart and the terminology in the coding manuals. It is necessary that physicians and other health professionals understand the ICD terminology and the importance of providing accurate and necessary information to determine the appropriate codes.

Health record departments need to promote the importance of utilizing the database on the local, provincial and national levels. Health record analysts need to advance the concept that these databases are useful, accurate and reliable. Positive aspects of the information need to be presented, not just the negative. Physicians and other professionals need to be encouraged to participate and recognize that documentation does make a difference for the final outcome of the coding process. Participation in education, workshops, in-service and public relations needs to go beyond what is currently being addressed within these departments.

The IOM studies (1977a; 1977b) stated that in order to improve coding, and ultimately the quality of the information, physicians need to be educated. Other professionals who document patients' events also need to be educated on the kind of information necessary for the database as well as for the coding process. The CHRA/CCHRA, along with the CIHI, should encourage health professional bodies to educate their members on documenting and charting procedures. The CHRA/CCHRA and the CIHI should encourage continuing education for all professionals who have access to charting using the ICD terminology.
Individual hospitals can also be encouraged to provide workshops to their own staff. Hospitals also need to ensure that physicians, in particular, report all diagnoses and procedures and that the wording of diagnoses and procedures relates to the terminology in the ICD.

Physicians need to be educated about, and involved in, activities that support the hospital, the provincial, and the national database. As stated previously, physicians need to accept more responsibility for coding patients' diagnoses and procedures and in completing the documentation within the chart. The CCHRA should ensure that avenues are available for them to do so and should lobby the physicians professional bodies to ensure adequate charting is taught in medical schools.

4.9.5 Continuous Charting

When patients are transferred within one institution from one level of care to another, i.e. from acute care to extended care, the patient is deemed to be a new admission, according to MOH guidelines. Some hospitals maintain continuous charting for these patients, i.e. the same record is used for that individual. Continuous charting does not separate the acute care portion of stay from the long term care portion. This type of charting makes it easier to care for the patient in the nursing unit, but is a nightmare to code. It makes it difficult to determine which part of the record was utilized for the acute care portion and which for the continuing care. Some hospitals prepared new charts when transferring patients from one level to another. Those charts were easier to code. Coding from records with continuous charting has a potential to produce incorrect data for the aggregate database. This study
recognizes that continuous charting for direct patient care is more important than coding and abstracting. With that in mind, this study recommends either discontinuing the use of continuous charting for intra-hospital transfer of patients or providing some kind of distinct sections, such as colour coding, to ensure that the right information is coded for the correct admission.

4.9.6 Face Sheets

Face sheets are very important to the coder as the first source of documentation of the diagnostic and procedural data. As is common practice in most health record departments, most face sheets have the alphanumeric code documented beside the written diagnoses and procedures. However, not all the written diagnoses on the face sheet match the alphanumeric codes. If the diagnoses (especially the MRD) have changed from what was originally written on the face sheet to what was actually determined as the MRD, the coder usually does not change the written information. Alphanumeric codes, although interpretable by the coder, may not be for most researchers or physicians. The MRD code represents the written description of the MRD, which may not be the first listed diagnosis on the face sheet. For instance, a physician may determine that inguinal hernia was the MRD because that was what the patient was admitted for and ultimately had surgery for. However, the coder may code the MRD as pneumonia because it extended the patient’s length of stay by ten days. The coder places the appropriate code for pneumonia on the chart as the MRD code. Unless the classification books are available to the researcher, it may be the written diagnosis of inguinal hernia that is interpreted as the MRD, not the alphanumeric code for pneumonia, depending
upon the qualifications and extent of knowledge of the researcher and the method in which
the diagnosis was retrieved.

The codes would be more meaningful to all concerned if the correct diagnosis were
actually written on the face sheet along with the correct alphanumeric code. This means
changing or adjusting the written diagnosis to match the alphanumeric code. Most health
record coders are reluctant to do this, and most physicians would not readily accept this
practice. Legally, this procedure could create problems if the physician has already
determined that the chart was accurate. However, this process may ensure better
documentation by the physicians if they could determine how the chart was interpreted by the
coder and, ultimately, how it will be classified and retrieved for future purposes. The coder
does not have the training or legitimacy to address this issue now.

A major concern is that coders often code only from the face sheet using the discharge
summary as back up information. Events need to "stand out" for the coder. Face sheets
need to be developed so that MRD and PP are predominantly displayed and there needs to be
better documentation by the physician. As stated previously, this relates to the training of
physicians in documenting procedures.

Beyond the face sheet; however, it is the physician's failure to document all relevant
information that is a problem. Major diagnoses and procedures are often missed because
relevant information is not documented or, if documented, they are not readily evident to the
coder. Within this study, too many procedures, both major and minor, were not adequately
documented on the face sheet.
One obvious observation is that if the attending physician completed the face sheet, but did not perform the surgery, the procedure is often missed on the face sheet and in the discharge summary. Another observation is that face sheets often contained scant information, the wrong information or no information at all. This was especially true for procedures. Some procedures were not recorded on the face sheet, but evidence was found in other documentation, such as the anaesthetic record. On the other hand, some procedures were recorded on the face sheet, with no evidence found throughout the rest of the chart. One wonders if these procedures were actually performed or if the physician mis-recorded the information.

Most hospital by-laws regulate that the documentation within a chart must be completed within 48 hours of discharge. Compliance with medical staff by-laws needs to be enforced at the local hospital level to ensure the accuracy of the database.
4.10 HOSPITAL POLICY

Another confounding variable influencing coder performance is differences in hospital policy. The purpose of a coded chart is to ensure that information is available for purposes other than the bedside care of the patient. As demonstrated, the method of collecting the data can determine the accuracy or inaccuracy of a database. It does make a difference whether the data are utilized for individual hospital sites, provincial or national databases. The fact that hospitals change policies to suit their own purposes may seriously compromise the national database.

It has been documented throughout the literature and demonstrated within this study that hospital administrative policies and procedures differ. Hospital policy contrary to HMRI (CIHI) or ICD-9 guidelines accounted for 6.0% of discrepancies with the MRD. It needs to be stressed that individual hospital policies and procedures that differ from the national guidelines do have an effect on the aggregate database and may have an impact on medical research. Hospital policy that is inconsistent with the HMRI (CIHI) guidelines forces coders to make inappropriate choices of codes. However, it is recognized that when guidelines or the classification system are inadequate for individual research, changing the guidelines is one way to enable individual hospitals to conduct their research. Both the OHA (1991) study and the NJDH (1978) study demonstrated that hospitals change policies to suit themselves.

Some hospitals in this study changed the HMRI and the ICD policies to suit the needs of the institution. For instance, one hospital coded unstable angina as 413.-1, when HMRI specifies it to be coded as 411. In the conversion of the data, the fourth and fifth digits, the '-' and the '1', are truncated for the aggregate database. Some hospitals coded angiograms
but not catheterizations: others coded both. Another coded 250.9 when an individual had two or more diabetic complications, instead of coding each complication separately, such as 250.2 and 250.3 (ICD-9CM has a combined complication code, but ICD-9 does not).

Individual policies and procedures can seriously bias information utilized in these databases. Hospitals should be discouraged from producing policies and procedures that differ from the guidelines of HMRI, the ICD and Statistics Canada. Data that may be useful to individual hospitals may be counterproductive to the database as a whole. The only way to avoid this conflict is to ensure that all rules and regulations are followed by each participating hospital. Hospitals should be encouraged to utilize the optional areas of the database for individual policies instead of changing national guidelines.

4.10.1 Coding From Incomplete Charts

The importance of training and coding from completed records is consistent throughout the literature. Coders should endeavour to review the entire health record to determine the greatest level of specificity possible and to gather all the relevant information necessary to select the appropriate code. 53.3% of total procedural discrepancies were clerical, mainly for not reviewing the entire record. The information was there, but not recorded. Coders need to be encouraged to code all relevant information, such as infections and complications. Relevant information is not "overcoding". Coding should be done only after all the reports are available. According to the hospital questionnaire, all hospitals admitted to coding (from 10 to 90%) from incomplete charts.
HMRI (CIHI) needs to ensure that every effort is made to ensure that health coders review the entire health record so that the greatest level of specificity can be determined for diagnostic and procedural codes. Responsibility for reviewing the entire record needs to be incorporated into the training programs for coders. Hospital policy also plays a major role and is discussed below.

4.10.2 Decreasing or Eliminating the Quota Required Per Day

Although not determined as a causal factor, the quota of records required per coder per day mandated by individual hospitals may decrease the opportunity for reviewing the entire record. But hospitals have staffing problems and decreasing quotas will either mean further delays in the coding and analysis or increasing the cost of coding in relation to patient care. In some instances, this may require an increase of funding to the health record department.

Given the demands on the data at provincial and national levels (and increasingly for regional purposes), it is essential that more accurate data be collected. Coders need more time to review the entire record. Health institutions need to assure this funding is allocated. CIHI needs to address the efficiency of coding if hospitals are to accept the increasing cost implications of reviewing the entire chart. This can be applied through efficiencies in database functioning, e.g., the development of an encoding system (discussed below) or assistance with computer-generated progressive coding.
4.11 CLASSIFICATION AND DATABASE SYSTEM PROBLEMS

Another confounding variable influencing coder performance is problems within the classification and database systems.

4.11.1 National Guidelines

Consistency in guidelines is required for a national database to produce accurate and reliable data. All levels of government need to be working together to produce a classification system understood and applied by everyone. HMRI (CIHI), Statistics Canada and the provincial Departments/Ministries of Health need to work together to ensure one set of rules, regulations and guidelines to maintain consistency and accuracy. National standards for all rules and guidelines need to be consolidated into one coding manual to eliminate the various sources of information currently available.

Dissemination of information on rules and guidelines needs to emanate from one central agency to ensure that the same guidelines are being used across the country. Information to health record coders should be disseminated only through one national body, such as the CIHI. In addition, HMRI (CIHI) needs to go beyond the editing of data into the actual auditing of data to ensure that what is coded is representative of the information in the chart itself.

National standardized terminology and operational criteria need to be determined on how to translate some codes that currently have varying interpretation. For instance, currently there are no clear guidelines for procedural coding. Procedural guidelines need to be developed, especially in the selection and usage of the principal procedure. Better
guidelines on non-operative procedures are needed, especially when a non-operative procedure is coded as the principal procedure. A separate category could be developed in the HMRI (CIHI) abstract for non-operative procedures.

Maintaining quality data in a fast changing health care environment is a challenge to the individual coder, the database company, as well as the provincial and national governments responsible for maintenance of the classification system. It is imperative that clear guidelines be developed and that all hospitals adhere to them. As stated earlier, individual hospitals should be discouraged from producing policies and procedures that differ from these guidelines, unless they are prepared to maintain parallel systems.

The role of the newly incorporated CIHI is to develop a comprehensive information system for Canada. It will develop and promote standards and guidelines for national health information. The challenge to the CIHI will be to develop procedures to ensure that the national databases are based on consistent guidelines, but that still enable individual researchers to investigate specific areas of concern. Common rules and guidelines from each provincial Ministry/Department of Health, Statistics Canada, HMRI (CIHI) and the ICD will be a challenge for CIHI.

4.11.2 Encoding System

An encoding system ensures that the conversion of data from written documentation to alphanumeric codes is made easier for the coder and, ultimately, more accurate for the national database. An encoding system needs to be developed and utilized as part of the HMRI (CIHI) software. The use of branching logic may eliminate the discrepancies caused
by the misuse of the classification system where inclusion and exclusion notes are often missed. An encoding system may also reduce the number of clerical errors caused by transposing digits and branching logic may decrease the selection problems associated with the MRD and the PP.

4.11.3 Incomplete Database

Ambulatory care patients represent a substantial workload for every hospital. Ambulatory care records and records from free-standing facilities in continuing care need to be mandated and incorporated into the national and provincial databases. It is essential that the database be as complete as possible. The recent trend to incorporate ambulatory care as part of an institution's experience should also be reflected in the provincial and national database. HMRI (CIHI) needs to mandate that ambulatory care records and records from free-standing continuing care facilities be coded, as these services are a large part of the nation's experience.

4.11.4 HMRI Reports to National and Provincial Governments

Hospitals are not completing all abstracting or corrections of the data within the time frame required by HMRI for submission to the provincial governments. Reports, such as autopsy and pathology reports are sometimes delayed beyond the time frame allowed by current guidelines. The original record is coded long before this event and sometimes the MRD differs between what the physician noted in his discharge summary and the actual
pathology report. All coders claim to address the changes, as per HMRI guidelines; however, this study shows the provincial and national databases are not being corrected.

Individual hospitals need to ensure that submission of data and corrections to be incorporated into the database are completed and all necessary corrections are provided to the appropriate provincial and national agencies. The Ontario Ministry of Health has instituted a penalty for the late submission of hospital abstracts. The B.C. Ministry of Health should also consider a penalty to hospitals not conforming to the timing requirements. HMRI needs to send preliminary reports to the Departments/Ministries of Health to ensure timely data, but also needs to follow up with final reports approximately six months later to ensure accuracy of the data.

The province is also in the process of creating new Regional Health Boards to manage health delivery within geographical regions. Since the allocation of resources will be based partly on data within the HMRI database, policies should be created to ensure procedures enable the timely submission of data. Funding needs to be provided to ensure this happens. Regional health boards need to take an active role in ensuring that data emanating from these regions are accurate and timely.

4.11.5 Conversion of HMRI (CIHI) Data

Prefixes, as well as fourth and fifth-digit suffixes, affixed to individual codes need to be retained so that the real incidence of disease requiring hospital care is accurately reflected in the database. In the conversion of data to meet the provincial and national governments needs, CIHI must ensure that the codes in the database actually represent the incidence of
disease and the workload of each hospital. Coding rules need to be designed so as to accurately reflect the real incidence of disease.

The concerns of Iezzoni (1988) are relevant here: some diagnoses that are suspect, or questionable are coded as real incidents and may lead to ambiguities and an inaccurate database. HMRI guidelines (see Appendix D) state that a diagnosis with "rule out" at the beginning of the term may be identified as the MRD and coded as if the disease existed. However, "ruled out" at the end of a term means that the MRD should be coded to V71.8, observation for a condition. "Rule out" diagnoses are being treated as real incidents. Another option is to place a "Q" in the prefix. However, if the "Q" is placed before the term for hospital purposes, HMRI drops it by the time it reaches the provincial or national databases. This process produces an upward bias in the real incidence of disease.

In the conversion of data from ICD-9CM codes to ICD-9 codes, care must be taken to ensure the real incidence of disease is not biased. Rule out, questionable and suspected diagnoses should not be translated as the occurrence of the disease. There needs to be some method to determine whether a diagnosis is suspect or questionable. Perhaps a fifth digit could determine such questionable disease entities. However, it is essential that these digits not be deleted in the data conversion from HMRI (CIHI) to the appropriate provincial Ministries/Departments of Health.

HMRI guidelines also need to enable fourth and fifth digit suffixes for procedure codes to be retained in the data conversion, such as when a patient had surgery in another hospital, e.g., a digit "8" is suffixed to the code. Conversion of the data sometimes deletes these suffixes (depending on the length of the procedural code) and the procedure appears as
if it was performed in that hospital. This may, or may not, be a bad thing, depending on the purpose of the application of the data. However, the role of the HMRI should be to ensure the actual events of the patient’s stay in hospital are recorded as accurately as possible.

Retention of prefixes and suffixes is important to ascertain such events.
4.12 ONGOING RESEARCH ON ACCURACY OF RECORD CONTENT

No attempt was made to examine the extent to which the recorded information in the health record accurately reflected the patient's condition or the quality of the documentation in the record. These were important questions, but lay outside the scope of this study. This study did not examine the extent to which the recorded information in the health record accurately reflects the patient’s condition or the quality of the data in the record but recommends that a study be undertaken to determine the accuracy of information in the health record. Although the accuracy of the health record has not been examined, the record is used as the gold standard for research studies. Based on the outcome of this study, it is imperative that reliability of the entire database is examined both provincially and nationally. Given the number of applications for this database, further study of its reliability is warranted.
Databases are invaluable tools for determining priorities in the health care system. They allow better utilization and understanding of the entire health care delivery system. Health care databases are utilized for purposes of budgeting, allocation of resources, comparing individual hospital practise to national practise, planning, and research. Information is essential for epidemiologic studies, for a better understanding of the health care system, to plan for the health needs of the nation, in describing, forecasting and analyzing the system, for evaluating program effectiveness and for clinical retrieval to provide access for research and teaching. Diagnoses and procedural codes are used to make decisions regarding the management of the entire health care delivery system. A uniform, comprehensive, and reliable database depends on the accuracy of coded data. Accurate information is essential for the future funding of the health care system, in making decisions regarding the management of the system, for inter-hospital comparisons, and in determining policy directions.

When information from health records is transformed into alphanumeric codes, something is inevitably lost in the translation. The use of health care databases to improve the quality of health care must be accompanied by precision in the collection of the data through the accuracy of the documentation in the chart and the transformation of that information into the database.

The analysis of the HMRI (CIHI) database in six Vancouver hospitals suggests that it is unsuitable for a quality of care study. A lack of inter-rater reliability is a source of variability that could seriously affect the outcome of or invalidate any study. Using the
HMRI (CIHI) database for some research purposes could bias the outcome unless details of the patient's record can be accessed. Care must be taken when using the HMRI (CIHI) database for some research purposes.

The reliability of the diagnostic and procedural information in the HMRI (CIHI) database is low and the contents may not be suitable for applications such as program planning, inter-hospital comparisons, patient-oriented medical research, or where the illness intensity of patients would underlie allocation of funding. The data need to be improved. The limitations of the database need to be delineated for clinical and research purposes beyond the clinical site. Clinicians should proceed with caution when utilizing this database. When data are collected, researchers need to be made aware of differences in individual hospital performance, not just differences in national or provincial guidelines. Each step in the collection of data can be a potential source of bias. Problems with the coding system are imparted directly into the database, whether from coder input or through the classification system itself. Accuracy of data input is essential to the HMRI (CIHI) database system.

The role of the health record coder and the role of health record departments is gaining status within the realm of health information processing, and the importance of the information emanating from these departments is on the increase. Although inadequate or poor training is not necessarily the cause of an unreliable database, there is a clear indication that health record coders need to upgrade and improve their skills.

Coders need to be trained better to meet the changing requirements in the health service delivery industry. However, training alone will not guarantee a greater degree of accuracy. Changes in variables associated with the coding process, such as the nature of the
coding process, hospital policies and procedures, documentation in charting requirements, education upgrading, testing procedures, credentialling, and inter-provincial requirements are all necessary to support the accuracy and reliability of the system. Both CIHI and CCHRA need to become more actively involved and take responsibility for the improvements necessary to make the system better. Rigorous educational and certification standards need to be implemented to ensure coders with sufficient skills are being utilized. The CCHRA needs to gain the power to remove credentials when testing standards are not met.

Through the structure of the newly developed CIHI, the system has the capacity to standardize the definitions, guidelines and interpretations which can provide uniformity and consistency nationally. It is up to the guardians of the system to ensure that these standards are maintained to ensure the viability of the contents. Educating other professionals may also produce a better quality of document within the record. Changing only one component of this guardianship will not effectively change the reliability of the system. Accuracy and reliability of any database takes teamwork of all the players involved.
BIBLIOGRAPHY


Abramson, J.H. Survey Methods in Community Medicine, Churchill Livingstone, 1990


Ambutas, Shirley. "A Comparison of Two Patient Classification Systems for an MICU" in Nursing Management, 1988, 19(9):64A-64H


Armstrong, Dr. Robert. Personal communication


Barnard, C. & Esmond T. "DRG -based reimbursement: the use of concurrent and retrospective clinical data" in Medical Care, 1981, 19:1071-82


Beard, C. Mary, Erik J. Bergstralh & George G. Klee. "Interobserver Variability in collecting data from medical records" in Archives of Pathology and Laboratory Medicine, 1988, 112:594-6


Berki, S.E. "Grading DRGs" in Medical Care, 1984, 22(12):1065-6

Beyt, B. Eugene, Susan Troxler & Joel Cavaness. "Prospective Payment and Infection Control" in Infection Control, 1985, 6(4):161-4


Blois, Marsden S. "Records and clinical databases" in MD Computing, 1984, 1(3):24-8


Botz, Charles K. & Chetram Singh. "Calculating the cost of CMGs in a Canadian Hospital" in Health Care, February 1985, 53-6


Broadway, Anne L. & Mary Ellen Brown. "Kingston General reviews the use of CMG's as a management tool" in Dimensions, 1987, 64(3):24-6


Buchanan, A. (Ed.) "B.C. Ministry of Health Data" in HMRI Newsletter, 1983, 8(2):3

Buchanan, A. (Ed.) "Case Mix and the HMRI Database" in HMRI Newsletter, 1982, 7(1):3-4

Bulatow, Nancy L. "Addressing concurrent review coding problems" in Journal of the American Medical Record Association, 1988, 59(5):36-8

Burford, Richard Baldwin, Jr. The relationship Between Diagnostic Information Available at Admission and Discharge for Patients in One PSRO Setting: Implications for Concurrent Review, Unpublished Manuscript In Candidacy for the Degree of Master of Public Health, 1977

Burnum, John F. "The Misinformation Era: The Fall of the Medical Record" in Annals of Internal Medicine, 1989, 110(6):482-4


Canadian Health Record Association. Constitution and Bylaws, 1990


Casper, Edward S. "A management system to maximize compliance with standards for medical records" in *Hospital and Community Psychiatry*, 1987, 38(11):1191-4

Chandler, Cheryl D. "Acute care burnout and the medical record practitioner" in *Journal of the American Medical Record Association*, 1987, 58(1):17-18


Clark, Donna L. "Writing - An Overlooked Competency in Medical Record Education" in *Journal of American Medical Record Association*, 1984, 55(7):39-42


Commission on Professional and Hospital Activities. *Professional Activity Study and Medical Audit Program*, Ann Arbor, Michigan, 1972
Commission on Professional and Hospital Activities, Inc. Professional Activity Study, Information Brochure, Ann Arbor, Michigan, 1959


Corn, Richard F. "Quality control of hospital discharge data" in Medical Care, 1980, 18(4):416-26


Curtis, Elizabeth H. "Quality assessment: monitoring and evaluating medical record documentation and services" in Topics in Health Records Management, 1989, 10(1):1-10


Davis, Karen, Gerard Anderson & Earl Steinberg. "Diagnosis related group prospective payment: Implications for health care and medical technology" in Health Policy, 1984, 4:139-47


De Rosis, F. "Regional and national hospital statistics" in Hospital Statistics in Europe, P.M. Lambert and F.H. Roger, (Eds.), Amsterdam: North Holland, 1982:137-45


Demlo, Linda K. & Paul M. Campbell. "Improving Hospital Discharge Data: Lessons from the National Hospital Discharge Survey" in Medical Care, 1981, 19(10):1030-40

Demlo, Linda K., Paul M. Campbell & Sarah Spaght Brown. "Reliability of Information Abstracted from Patients' Medical Records" in Medical Care, 1978, 16(12):995-1005

Demuth, Anne I. "Beyond first-generation coding systems" in Journal of the American Medical Record Association, 1986, 57(9):50-1


Dormeus, H.D. & E.M. Michenzie. "Data quality: An illustration of its potential impact upon a diagnosis related group’s case mix index and reimbursement" in Medical Care, 1983, 21:1001-11

Doremus, Harvey D. "DRGs may be raising false expectations" in Hospitals, August 1, 1980:47-51

Dowell, Susan P. "Clinical data systems provide clues for health care management in the 1990's" in Topics in Health Records Management, 1988, 8(3):10-6

Eliasoph, H. & E. Hassen, "VIP streamlines care and reduced length of stay" in Dimensions, October 1986:43-5


Feinstein, Alvan R. "ICD, POR, and DRG. Unsolved scientific problems in the nosology of clinical medicine" in Archives of Internal Medicine, 1988, 148(10):2269-74

Feinstein, Alvan R. Clinimetrics, Yale University Press, 1987


Finnegan, Rita. "Coding notes... guidelines for ambulatory coding" in *Journal of the American Medical Record Association*, 1988b, 59(10):26-7

Finnegan, Rita. "Coding notes - coded data on ambulatory care" in *Journal of the American Medical Record Association*, 1987a, 58(4):6, 38


Finnegan, Rita. "Essentials of accurate coding" in *Journal of the American Medical Record Association*, 1985, 56(9):9-16


Garrison, Linda S. "Personal computer applications for the medical record department" in *Topics in Health Records Management*, 1988, 8(4):54-63


Green, Deborah K. "A focused approach to coding quality control" in Topics in Health Records Management, 1986, 7(1):1-10


Grostick, Sara & Donna Slovensky. "THRM forum: From MRA to CIO -- are MRAs ready for the challenge?" in Topics in Health Records Management, 1989, 10(1):65-70


Harrington, Sherry. "Canadian Hospitals have ICD-9-CM option" in *HMRI Newsletter*, 1987, 12(3):1

Health Record Association of British Columbia. *Constitution & Bylaws* and other documents, 1990


Helyar, Chris. Vice-President, Research and Development, Hospital Medical Records Institute, personal communication


Hickie, Deanna. B.C. Representative, Hospital Medical Records Institute, personal communication - several occasions


Horn, Susan D. & Roger A. Horn. "Reliability and validity of the Severity of Illness Index" in *Medical Care*, 1986b, 24(2):159-78


Hospital Medical Records Institute. *Case Mix Methodology*, Unpublished document, April, 1990


Inglehart, John K. "Canada’s Health Care System (First of Three Parts)" in *The New England Journal of Medicine*, 1986a, 315(3):202-8

Inglehart, John K. "Canada’s Health Care System (Second of Three Parts)" in *The New England Journal of Medicine*, 1986b, 315(12):778-84


Israel, Robert A. "Planning for the Tenth Revision of ICD" in *Journal of the American Medical Records Association*, December 1984:26-7


Kingsley, David I. "The importance of documentation in peer review" in New Jersey Medicine, 1989, 86(3):185-6


Larks, Marlene. "Panel confirms data uniformity is key to improving healthcare quality" in Patient access, 1989, 12(6):2-3


McLean, Barbara. "Can you Defend the Accuracy of your Coded Data?" in HMRI Newsletter, 1984, 9(2):4


Macies, Joseph S. "Strategic information Planning for the 1990's" in Health Care, February 1984:22-3


Mansfield, Brian G. "How bad are medical records? A review of the notes received by a practice" in Journal of the Royal College of General Practitioners, 1986, 36:405-6


Martin, David L. "Computerizing medical records: the future is here, Part 1" in Dimensions, November 1987:50-1


Mendenhall, Stanley. "DRGs must be changed to take patient's illness severity into account" in *Modern Healthcare*, November 15, 1984:86-7


Miller, Susan C. "Medical record-related practices in hospices, Part II- Survey Results" in *Journal of the American Medical Record Association*, 54(12):23-9


Mullin, Robert L. "Diagnosis-Related Groups and Severity. ICD-9-CM, the Real Problem" in *Journal of American Medical Association*, 1985, 254(9):1208-10


New Jersey Department of Health. Division of Health Planning and Resources Development. Reliability of New Jersey Hospital Discharge Abstracts, December 1978


O'Gorman, J. "Data accuracy & reliability" in Hospital Statistics in Europe, P.M. Lambert and F.H. Roger, (Eds.), Amsterdam; North Holland, 1982:113-7


Paccaud, F. "Conversion of an administrative record system for reliability with the ICIDH" in International Rehabilitation Medicine, 1986, 8(1):15-7


Picard, Ellen I. Legal Liability of Doctors and Hospitals in Canada, 1978


Rabkin, Mitchell T. "Will the DRG Decimate Clinical Research?" in Clinical Research, 1985, 32(3):345-7


Redelmeier, Donald A. "Public Access to Death Certificates" in Annals of Internal Medicine, 1989, 111(1):92


Richards, Tom, Nicole Lurie, William H. Rogers, & Robert H. Brook. "Measuring differences between teaching and nonteaching hospitals" in Medical Care, May 1988, 26(5) Supplement:S1-141


Rosko, M.D. "DRGs and Severity of Illness measures: An analysis of patient classification systems" in *Journal of Medical Systems*, 1988, 12(4):257-74


Smith, Murray W. "Hospital discharge diagnoses: How accurate are they and their international classification of diseases (ICD) codes?" in *New Zealand Medical Journal*, 1989, 102: 507-8
Solomkin, Joseph S. "Who should do the coding?" in *Journal of American Medical Association*, 1986, 255(5):613-14


Stergachis, Andy S. "Record linkage studies for postmarketing drug surveillance: Data quality and validity considerations" in *Drug Intelligence and Clinical Pharmacy*, 1988, 22(2):157-61

Stern, Robert S. "Record Linkage. A powerful tool for epidemiologic analysis" in *Archives of Dermatology*, 1986, 122(12):1383-4


Taylor, Elizabeth. "International Report. 11th International Health Record Congress" in Progress Notes, February 1992, No.44:5


Taylor, Mrs. E., "How Much (or How Little) To Code?" in HMRI Newsletter, 1984, 9(4):3

Thompson, Barbara J. & Vergil N. Slee. "Accuracy of Diagnosis and Operation Coding", in Medical Record News, October 1978:43-60

University of British Columbia. Division of Health Services Research and Development. Office of Coordinator of Health Sciences. Production 81, 1982


Vogel, Hans-Peter. "Influence of additional information on interrater reliability in the neurologic examination" in Neurology, 1992, 42:2076-81


Williams, S.E. & Latessa P. "Improving the quality of discharge data" in Topics in Health Record Management, 1982, 2:41-8

Winchell, J.R. "Reduced costs can improve care" in Health Care, October 1985:15-20


Winger, B. & Peggy Wu. (Eds.) "Just a Reminder" in HMRI Newsletter, 1988b, 13(3):1

Wingert, F. "Automatic Indexing Based on SNOMED" in Methods of Information in Medicine, 1985, 24(1):27-34


Wright, Dr. Charles J. "Who cares about accurate coding?" in *HMRI Newsletter*, 1983, 8(2):1-2


Yamauchi, Kazunobo, Takayuki Miura, Yasutomi Ogawa & Masao Miyajima. "Utilization of medical records in Nagoya University Hospital" in *Japan Hospital*, Jul 1988, 7:67-70


Young, Wanda W. "ICD-9-CM code and coding practice changes for case mix management" in *Journal of the American Medical Record Association*, 1986, 57(3):29-33

Young, Wanda W. "Incorporating severity of illness and comorbidity in case-mix measurement" in *Health Care Financing Review*, November 1984, Annual Supplement:22-31


APPENDIX A - DEFINITIONS OF TERMS

Some terminology used throughout this study has specific meaning and requires further explanation.

ABSTRACTING Abstracting is a term used by the health record profession to denote the process used when converting the documentation in the health record into the computerized format required by the database company.

CODING Coding is a term used by the health record profession to denote the process used when changing the diagnostic and procedural terminology into alphanumeric codes of the specific classification system (such as ICD-9).

CODING DISCREPANCY A coding discrepancy is defined as any deviation between the computerized HMRI database and re-abstracted data elements where there is not a one-to-one correspondence between all the digits. A discrepancy exists if any of the significant digits in the diagnostic, procedural or non-diagnostic codes do not agree. All cases have an MRD and anywhere from zero to fifteen secondary diagnoses. Some cases have PP’s. Those that have PPs, have anywhere from zero to thirteen secondary procedures. Therefore, one individual case could have several coding discrepancies, e.g. each discrepant code could have a different “type of discrepancy”. For a case to be discrepant, only one code needs to be discrepant.

Non-operative procedural codes such as x-ray’s and ECG’s were considered to agree if both data sets documented the correct code, despite its position. These were not noted as discrepancies when the accompanying code presented itself in both data sets, the original and the re-abstracted.

DIAGNOSTIC CODES Diagnostic codes using the ICD-9 classification system are categorized by a three-digit code. These three digits determine the disease category. Sometimes this code is followed by a decimal and one or two other digits. The fourth- and fifth-digits, after the decimals more clearly and precisely specify the disease entity. However, not all codes have fourth and fifth digits. If a code does has a fourth digit, it does not necessarily mean it has a fifth digit. If it has a fifth digit, it almost always has a fourth digit: although a rare occurrence, this can happen.

Two examples are:

• Three-digits only: acute myocardial infarction, 410, where the three-digit code determines the disease category. In ICD-9, there are no fourth or fifth digits associated with this code.

• Fourth and fifth digits: diabetes mellitus, 250, where the three digits determine the disease category; diabetes mellitus without complication, 250.0 (4-digits), and diabetes mellitus without mention of complication, Type I, juvenile type onset, 250.01 (5-digits). The fourth digit indicates the complications and the fifth digit determines the type of onset.
APPENDIX A - DEFINITIONS OF TERMS (Page 2)

DIAGNOSTIC CODES (Cont.) Each code needs to be checked for inclusions or exclusions, for optional or mandatory subclassifications, and for whether it is a dagger (†) or an asterisk (*) code. Dagger and asterisk codes are alternate positions in the classification for relevant conditions (dual classification), such as the underlying condition or manifestation or complication. An example is diabetes retinitis which has a mandatory dagger (†) code, 250.4 under endocrine and metabolic disorders, and an optional asterisk (*) code, 362.0, under retinal disorders. Diabetic retinitis cannot be coded as most responsible under retinal disorders: the system mandates that the diabetic category be used as the most responsible code. The optional 362.0 code would be used subject to individual hospital policy, or when no policy exists, subject to the individual coder.

DISCREPANCY A discrepancy is defined as any deviation from the data elements where there is not a one-to-one correspondence. A discrepancy exists if all four digits in the diagnosis or procedure codes do not agree.

DISCREPANT RECORD When there is one or more discrepant data elements in a record.

FIFTH-DIGIT CODES At the time of this study in the province of British Columbia, the fifth digit, if the code has this sub-classification, was not mandatory. The classification system directs that the fifth-digit of a code is optional and, if utilized, is only retained for the individual hospital's records. HMRI severs this digit in the preparation of the data for the B.C. MOH. As noted earlier, hospital personnel send the abstracted data directly to HMRI (CIHI) for processing. CIHI then distributes this information to the appropriate Ministry/Department of Health, and provides the reports for the individual hospitals.

MOST RESPONSIBLE DIAGNOSIS The Most Responsible Diagnosis (MRD) by HMRI definition is that "one diagnosis which describes the most significant condition of a patient which causes his stay in hospital. In a case where multiple diagnoses may be classified as Most Responsible, the diagnosis that is responsible for the greatest length of stay is chosen" (HMRI Guidelines, Appendix Z).

NON-OPERATIVE PROCEDURES All remaining ICD codes not considered by definition as an operating room procedure.

OPERATING ROOM PROCEDURE Procedures, based on clinical judgement would, in most hospitals, be performed in the operating room. This procedural code does not need to make any connection to the Most Responsible Diagnosis.
APPENDIX A - DEFINITIONS OF TERMS (Page 3)

PRINCIPAL PROCEDURE  The Principal Procedure (PP) by HMRI definition is "that procedure considered most significant during a patient's stay" (HMRI's Terms Associated with CMG's in Appendix C). The PP position incorporates non-operative procedures such as x-rays and ECG's.

PROCEDURAL CODES  Procedural codes consist of two digits followed by a decimal point and one, two or sometimes three other digits. This study included both operative and non-operative procedures, although the coding of non-operative procedures by HMRI guidelines is optional. It is left up to the individual hospital whether to code non-operative procedures.

Unlike for the diagnostic categories, there are no clear-cut guidelines available to code procedures. Not all hospital patients have a procedure; therefore, not all hospital records have procedures. Denominators for this category reflected the actual number of coded procedures.
LIST OF VENDORS INSTALLING HMRI ABSTRACTING SERVICES

Savoire Faire
B. C. Info Health (BCHA)
Meditech
Prism Hospital Software
Sphere Holdings
1st Healthcomp
Network Data Systems Ltd.
Shared Medical Systems
GSC
3M
Health Vision Corp

also numerous in-house systems
<table>
<thead>
<tr>
<th>Field</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Admission Date</td>
<td>12/23/89</td>
</tr>
<tr>
<td>Discharge Date</td>
<td>12/30/89</td>
</tr>
<tr>
<td>Name</td>
<td>John Doe</td>
</tr>
<tr>
<td>Address</td>
<td>123 Main St, Anytown, USA</td>
</tr>
<tr>
<td>Age</td>
<td>30</td>
</tr>
<tr>
<td>Gender</td>
<td>Male</td>
</tr>
<tr>
<td>Ethnicity</td>
<td>White</td>
</tr>
<tr>
<td>Weight</td>
<td>5'10&quot;</td>
</tr>
<tr>
<td>Height</td>
<td>170 cm</td>
</tr>
<tr>
<td>Admission Category</td>
<td>Elective</td>
</tr>
<tr>
<td>Reason for Admission</td>
<td>Surgery</td>
</tr>
<tr>
<td>Diagnosis</td>
<td>Cancer</td>
</tr>
<tr>
<td>Procedure</td>
<td>Surgery</td>
</tr>
<tr>
<td>Date of Procedure</td>
<td>12/23/89</td>
</tr>
<tr>
<td>Time</td>
<td>10:00 AM</td>
</tr>
<tr>
<td>Type of Anesthesia</td>
<td>General</td>
</tr>
<tr>
<td>Date of Anesthesia</td>
<td>12/23/89</td>
</tr>
<tr>
<td>Time of Anesthesia</td>
<td>10:00 AM</td>
</tr>
<tr>
<td>Discharge Plan</td>
<td>Home</td>
</tr>
<tr>
<td>Social Services</td>
<td>Supportive</td>
</tr>
<tr>
<td>Special Care/Other Services</td>
<td>Palliative Care</td>
</tr>
<tr>
<td>Pre-Admission Workup</td>
<td>Complete</td>
</tr>
</tbody>
</table>
Definition of Most Responsible Diagnosis:

"The one diagnosis which describes the most significant condition of a patient which causes his stay in hospital. In a case where multiple diagnoses may be classified as Most Responsible, code the diagnosis responsible for the greatest length of stay."

1. CHOOSING BETWEEN TWO OR MORE DIAGNOSES OF EQUAL IMPORTANCE:

When two or more diagnoses of equal importance are listed with no clear indication in the medical record as to which one is the Most Responsible Diagnosis, select the condition for which a definitive (as opposed to diagnostic) surgical or non-surgical procedure has been performed. If no surgery was performed select the first-listed diagnosis as the Most Responsible Diagnosis.

example:

<table>
<thead>
<tr>
<th>Diagnosis</th>
<th>Code</th>
</tr>
</thead>
<tbody>
<tr>
<td>Acute bronchitis</td>
<td>466.0  (M)</td>
</tr>
<tr>
<td>Acute gastritis</td>
<td>535.0</td>
</tr>
<tr>
<td>Acute duodenitis</td>
<td>535.6</td>
</tr>
<tr>
<td>No procedures</td>
<td></td>
</tr>
</tbody>
</table>

2. SPECIFICITY:

When the "main" diagnosis describes a condition in general terms, but a more descriptive term providing more precise information about the site or the nature of the condition is reported among the "other" diagnoses, select the latter condition.

example:

<table>
<thead>
<tr>
<th>Main Condition</th>
<th>Code</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cerebrovascular accident</td>
<td>431</td>
</tr>
<tr>
<td>Other Condition: Cerebral haemorrhage</td>
<td></td>
</tr>
</tbody>
</table>

Code to 431 - Cerebral haemorrhage (M)

3. CONTRASTING OR COMPARATIVE DIAGNOSES:

An inconclusive diagnosis stated in terms of:

- comparative and contrasting diseases or conditions should be coded as both conditions being present and the first listed condition may be assigned as the Most Responsible Diagnosis.
3. example: Pneumonia versus Pulmonary infarction

Code Pneumonia (M)

- comparative and contrasting etiologies should be coded to unspecified cause of disease or condition.

example: Hypoplastic anaemia versus aplastic anaemia. Code to unspecified cause.

- symptom followed by contrasting or comparative diagnoses should be coded so that the symptom is the Most Responsible Diagnosis. Code also all of the contrasting diagnoses as suspected conditions.

example: Acute abdominal pain, NYD Secondary to diabetic neuropathy or cholecystitis

Code all three in the order given

4. DAGGER AND ASTERISK:

Diagnoses which appear with an asterisk in ICD-9 may never be recorded as the Most Responsible Diagnosis. Codes accompanied with an asterisk must always follow the code for the underlying cause of disease, (those appearing with a dagger sign), thus can never be recorded first.

example: Diabetic Cataracts

Code: 250.49 (M) Diabetes with ophthalmic manifestations
366.4 Diabetic cataract

5. MULTIPLE INJURIES:

When a patient has multiple injuries, the Most Responsible Diagnosis selected will be the most severe injury.

example: Subdural haematoma 852.0 (M)
Fracture 8th and 9th ribs 807.0
Fracture clavicle 810.0

See also item #22
6. **PRESENTING SYMPTOMS OF A DIAGNOSED CONDITION:**

Symptoms may never be assigned as the Most Responsible Diagnosis whenever the underlying cause has been diagnosed.

**Example:** Abdominal Pain due to 532.1 (M)
- Acute perforated duodenal ulcer

7. **SUSPECTED DIAGNOSIS:**

If the diagnosis established at the time of discharge is stated as "suspected", "questionable", "likely" or in similar terms, code the condition as if it existed. In such cases, the code for the suspected condition may be assigned as the Most Responsible Diagnosis.

**Note:** For hospitals wishing to identify "suspected" diagnoses, an optional prefix of "Q" may be used to designate "questionable" conditions.

**Example:** ? Multiple sclerosis Q340 (M)

8. **RULE OUT, RULED OUT, R/O:**

- When the terms "rule out" or R/O appear at the beginning of a diagnostic statement, it is treated as a suspected condition that is to be ruled out. This code may used as Most Responsible Diagnosis if appropriate.

**Example:** R/O duodenal ulcer Code 532.9 (M)

- When the term "ruled out" or R/O is not confirmed, not proven or not found appears at the end of a diagnostic statement, it is treated as no evidence of a specific condition found and the appropriate V-code is assigned. The V-code may be used as the Most Responsible Diagnosis code if appropriate.

**Example:** duodenal ulcer, ruled out Code V71.8 (M)
9. STATUS POST, POST:

When a post-surgical or post-diagnosis status is listed at the time of discharge, read the discharge summary or note to obtain a better understanding of the physician's intent. Categories V10-V15 indicate a personal history of potential health hazards with no recurrence. V40-V49 indicate conditions influencing the patient care management but are not in themselves a current illness or injury. V50-V59 indicate reasons for care. It may be necessary to translate a post-surgical status as a complication of a procedure. Also, refer to "absence of organ or anatomical site", acquired in the Alphabetic Index if it is relevant to the diagnosis. A history of past conditions not relevant to the current episode of care may be omitted for coding purposes unless there is a stated need for the information.

example:  Cystoscopy for bladder tumour recheck, no evidence of carcinoma  Code V67.0

Status post-tympanoplasty, transferred from another hospital for recuperation and surgical aftercare.  Code V58.4

10. LATE EFFECTS:

The Most Responsible Diagnosis for conditions resulting from late effects of injury or illness is that for the residual effect.

example:  Unequal leg length (acquired) 736.8 (M)
(late effect of poliomyelitis) 138

An additional code is used to identify the cause of the late effect. For example, if the diagnosis is stated as "scoliosis due to old infantile paralysis" the code for scoliosis is designated as the Most Responsible Diagnosis and the additional code for late effect of infantile paralysis is also recorded.

If the source document does not provide any information to identify the type of residual effect, the appropriate code for the cause of the late effect may be used as the Most Responsible Diagnosis.

For example, if the diagnosis at discharge is stated as "old infantile paralysis" and no further information can be found in the medical record to identify the specific type of residual effect, code 138 "late effects of acute poliomyelitis" as the Most Responsible Diagnosis.
11. **POISONING:**

When a patient is hospitalized due to poisoning, the code for poisoning (960.0 - 979.9) is recorded as the Most Responsible Diagnosis. The nature of the adverse reaction is recorded as an additional diagnosis. The poisoning E-code (accidental, suicide attempt, assault, undetermined) is also recorded.

example: Coma due to accidental overdose of Librium Code 969.4 (M) 780.0  E853.2

12. **ADVERSE EFFECT OF CORRECT THERAPEUTIC SUBSTANCE:**

The Most Responsible Diagnosis code is that for the manifestation of the drug reaction and the companion code is the pertinent E-code identifying the drug.

example: Acute aspirin gastritis resulting from correct dose properly administered.

Code Acute gastritis 535.0 E935.1 describes the drug - Salicylates

13. **ACUTE AND CHRONIC CONDITIONS:**

When a specific condition is stated to be both acute (or subacute) and chronic, and the Alphabetic Index and/or the Tabular List provides separate codes at either the three or four digit level for acute and chronic, use both codes, positioning the acute condition first.

example: Acute on chronic bronchitis Code: 466.0 (M) 491.9

14. **POST-OPERATIVE COMPLICATIONS:**

Use an additional code - 998.8 "other specified complications of procedures, not elsewhere classified" - to identify post-operative complications classified to Chapters 1 - 16 when they are not identified as such by the category or subcategory title.

example: Post-operative pulmonary edema Code 518.4 (M) 998.8 E878.9

**NOTE:** The code from Chapter 1 - 16 would take precedence over 998.8 in assigning the Most Responsible Diagnosis.
Only one code is necessary to identify post-operative complications classified to Chapters 1 - 16 where the category of subcategory title identifies them as such.

example: Tracheoesophageal fistula following tracheostomy Code 5190 (M)

**QUESTION:** When certain post-operative complication codes are not site-specific is multiple coding possible for purposes of better analysis?

**ANSWER:** When possible, use a code in addition to those in 996-998 to identify the site involved if the additional code does not alter the stated diagnosis.

Refer to the Alphabetical Index for codes to provide the site for certain post-operative complications only in-so-far as the additional code meets the criterion of the guideline stated in the previous paragraph. For instance, in some cases, it is possible to use an additional code to identify the site of a persistent post-operative fistula, which is coded 998.6.

example: Persistent post-operative rectovaginal fistula Code 998.6
(digestive-genital tract fistulae, female) 619.1
E878.9

On the other hand, persistent post-operative bronchopleural fistula would be coded 998.6 only. It is not possible to use an additional code to identify the site in this case because the code for "bronchopleural fistula NOS" is the same code for "empyema with fistula" and empyema is not necessarily present with a persistent post-operative bronchopleural fistula.

In many cases, multiple coding to identify the site would not be necessary because the operative procedure code would identify the site.

The code in 996-998 takes precedence in assigning the Most Responsible Diagnosis for readmissions.

15. **STERILIZATION** V25.2

If the chief reason for sterilization is for contraceptive purposes, used code V25.2 "sterilization" as the Most Responsible Diagnosis code. If there is an underlying, incapacitating medical or psychiatric condition, code the condition(s) as associated diagnoses. If the sterilization is purely elective, V25.2 will be sufficient.

example: Completed family for elective sterilization Code V25.2 (M)
If the sterilization procedure was performed for contraceptive purposes during a current admission for obstetrical delivery, use code V25.2 as an additional code. However, if the sterilization resulted from a hysterectomy performed because of injury or damage to the uterus during delivery, do not code V25.2.

example: Cesarean section delivery, viable female infant (oblique presentation) Subsequent tubal ligation (completed family).

Code: 652.31 (M) V25.2 V27.0

Emergency c-section for ruptured uterus, removal of uterus resulting in sterilization.

Code: 665.1 (M)

If the chief reason for hospitalization proved to be the treatment or evaluation of a given disease or disorder and sterilization was performed as part of the treatment or as a result of the evaluation, code only the given condition(s). Do not use code V25.2.

example: Hysterectomy due to uterine fibroids (intramural leiomyoma)

Code: 218 (M)

IN SUMMARY: Code V25.2 is to be used only when the sterilization procedure was performed for the major purpose of contraception rather than for treatment of a disease.

16. V10-V19.8 PERSONAL HISTORY AND FAMILY HISTORY:

The codes V10-V19.8 must never be used as a Most Responsible Diagnosis code. They are supplemental codes.
17. **V-CODES, FOLLOW-UP, MAINTENANCE, AFTERCARE:**

When the patient is hospitalized for follow-up, maintenance therapy, aftercare or related reasons that are generally classified in the Supplemental Classification of Factors Influencing Health Status and Contact with Health Service (V-Codes) the appropriate V-code is generally recorded as the Most Responsible Diagnosis. In these cases the diagnosis was established and the definitive treatment was rendered on a previous admission or the patient is without a diagnosis.

example: Change of dressing - amputation site Code V58.3 (M)

18. **DISCHARGE DIAGNOSIS STATED AS A PROCEDURE:**

If the diagnosis at discharge is stated in terms of a procedure, examine the operative report and the rest of the discharge summary for information about the condition for which the procedure was performed. (If the source information is insufficient consult the responsible physician for clarification).

example: Cholecystectomy (code to Cholecystitis) Code 575.1 (M)

19. **DISAGREEMENT OF DIAGNOSIS WITHIN THE CHART:**

If a disagreement is encountered between the statement of diagnosis or procedure on the face sheet and in the rest of the chart, refer the medical record to the supervisor who may find it necessary to refer to the attending physician for clarification.

20. **LATE EFFECTS OF POISONINGS:**

Code in the same manner as other later effects.

21. **LATE EFFECT OR ADVERSE REACTION TO THERAPEUTIC USE OF DRUGS:**

Code in the same manner as an acute reaction to the drug/chemical substance.
22. MULTIPLE INJURIES:

When a patient has multiple injuries, select the most severe as the Most Responsible Diagnosis. When two or more diagnoses are of equal severity, choose that for which definitive treatment was given as the Most Responsible Diagnosis. If there is still a question concerning which diagnosis is most severe, then code the first listed diagnosis as Most Responsible.
See also Item #5
ACUTE GENERAL HOSPITALS IN THE GREATER VANCOUVER REGIONAL HOSPITAL DISTRICT ON HMRI UTILIZING ICD-9 CLASSIFICATION SYSTEM SELECTED FOR FISCAL YEAR 1986/1987

Lion’s Gate Hospital (369*)
St. Paul’s Hospital (590*)
Vancouver General Hospital (885*)
Mt. St. Joseph’s Hospital (121*)
**University Hospital - U.B.C. site (300*)
**University Hospital - Shaughnessy Site (375*)

All hospitals have been contacted and have agreed to participate in the study. Individual hospital research applications have been made and approval obtained.

See Appendices K-P for documentation on individual hospital approval.

*The number in brackets behind the Hospital name indicates the number of acute care beds as indicated by the GVRHD as at January 1990.

**University Hospital is considered two separate sites because the Medical Record Departments are operating as if separate entities. The policies and procedures for each department are different.
APPENDIX F - SAMPLE SIZE CALCULATIONS, NUMBER OF SELECTED CASES PER CATEGORY AND SAMPLE FRAME

The sample size calculation is based on the non-central chi-squared distribution.

**Distinct Diagnostic Categories**
One diagnostic category has a discrepancy rate of 10%, while two have 30%

**Inter-Hospital Comparison (Random Sample)**
Three hospitals have discrepancy rates of 10% and three have 30%.

<table>
<thead>
<tr>
<th>Alternate hypothesis</th>
<th>df</th>
<th>lambda/N</th>
<th>chi critical point (α=0.05)</th>
<th>lambda (β=0.20)</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>Distinct Diagnostic Categories</td>
<td>2</td>
<td>0.0505</td>
<td>5.99</td>
<td>9.64</td>
<td>194.5</td>
</tr>
<tr>
<td>Inter-Hospital Comparison (Random Sample)</td>
<td>5</td>
<td>0.0624</td>
<td>11.07</td>
<td>12.83</td>
<td>205.6</td>
</tr>
</tbody>
</table>

**Total cases**
Distinct Diagnostic Category sample = 60 cases per hospital and
Random sample = 40 cases per hospital, for a total of 100 cases per hospital * 6 hospitals = 600 cases overall

**SAMPLE FRAME**
Vancouver General Hospital 28,013
St. Paul’s Hospital 20,458
Mount St. Joseph’s Hospital 4,704
Lions Gate Hospital 15,997
University Hospital - Shaughnessy Site 9,301
University Hospital - U.B.C. Site 12,036
Total Sample Frame 90,509
APPENDIX F - SAMPLE SIZE CALCULATIONS, NUMBER OF SELECTED CASES PER CATEGORY AND SAMPLE FRAME (Page 2)

BY DISTINCT DIAGNOSTIC CATEGORIES

<table>
<thead>
<tr>
<th>ICD-9 3-DIGIT CATEGORY</th>
<th># CASES</th>
<th># ALTERNATES</th>
</tr>
</thead>
<tbody>
<tr>
<td>250-259 DIABETES</td>
<td>20</td>
<td>4</td>
</tr>
<tr>
<td>410-414 CARDIOVASCULAR</td>
<td>20</td>
<td>4</td>
</tr>
<tr>
<td>490-496 BRONCHOPNEUMONIA</td>
<td>20</td>
<td>4</td>
</tr>
<tr>
<td>TOTAL # CASES</td>
<td>60</td>
<td>12</td>
</tr>
</tbody>
</table>

BY STRATIFIED RANDOM CATEGORIES

To ensure random population was similar in all hospitals, one hospital (UBC’s distribution of cases, 1986/87) was chosen as the standard. This standardized case mix was used for all hospitals. Percentages were taken on the number of cases in UBC’s population. These percentages were used to determine the numbers necessary for each category. Proportions determine whole number distributions in each of the categories.

<table>
<thead>
<tr>
<th>ICD-9 3-DIGIT CLASSES</th>
<th>STRATIFIED BY CATEGORY</th>
<th>UBC DISTRIBUTION</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>ESTIMATED # CASES</td>
<td>ACTUAL # CASES</td>
</tr>
<tr>
<td>001-139 INFECTIONS</td>
<td>2</td>
<td>5.00</td>
</tr>
<tr>
<td>240-289 ENDOCRINE</td>
<td>2</td>
<td>5.00</td>
</tr>
<tr>
<td>320-389 NERVOUS</td>
<td>5</td>
<td>12.50</td>
</tr>
<tr>
<td>390-459 CIRCULATORY</td>
<td>8</td>
<td>20.00</td>
</tr>
<tr>
<td>460-519 RESPIRATORY</td>
<td>1</td>
<td>2.50</td>
</tr>
<tr>
<td>520-579 DIGESTIVE</td>
<td>9</td>
<td>22.50</td>
</tr>
<tr>
<td>580-629 GENITOURINARY</td>
<td>8</td>
<td>20.00</td>
</tr>
<tr>
<td>680-709 SKIN/MUSCULO-SKELETAL</td>
<td>1</td>
<td>2.50</td>
</tr>
<tr>
<td>740-759 CONGENITAL ANOMALIES</td>
<td>2</td>
<td>5.00</td>
</tr>
<tr>
<td>800-999 INJURIES/POISONING</td>
<td>2</td>
<td>5.00</td>
</tr>
<tr>
<td>TOTAL # CASES</td>
<td>40</td>
<td>100%</td>
</tr>
</tbody>
</table>
APPENDIX G - STANDARD ERROR CALCULATIONS AND CONFIDENCE INTERVALS

Given sample size, n, and estimated proportion of $p_0 = 0.10$ and $p_1 = 0.90$, the Confidence Intervals (CI) are estimated each CI upper and CI lower limit. $\alpha = 0.05$, one-sided test, $Z_\alpha = 1.645$.

FOR DISTINCT DIAGNOSTIC CATEGORIES

$$\text{Std. Error} = \sqrt{\frac{(p_0)(1-p_0)}{n}}$$

where $p_0 = 0.10$ and $n = 360$

$$\text{Std. Error} = \sqrt{\frac{(0.10)(1-0.10)}{360}}$$

$$= \sqrt{\frac{0.09}{360}}$$

$$= 0.0158$$

Confidence Interval = $p_0 \pm Z_\alpha(\text{S.E.}) = 0.10 \pm 1.645(0.0158) = 0.10 \pm 0.03$

UPPER LIMIT = 0.13
LOWER LIMIT = 0.07

FOR INTER-HOSPITAL COMPARISONS

$$\text{Std. Error} = \sqrt{\frac{(p_1)(1-p_1)}{n}}$$

where $p_1 = 0.90$ and $n = 600$

$$\text{Std. Error} = \sqrt{\frac{(0.90)(1-0.90)}{600}}$$

$$= \sqrt{\frac{0.09}{600}}$$

$$= 0.012$$

CI = $p_1 \pm Z_\alpha(\text{S.E.}) = 0.90 \pm 1.645(0.012) = 0.90 \pm 0.02$

UPPER LIMIT = 0.92
LOWER LIMIT = 0.88
APPENDIX H - MOST FREQUENTLY OCCURRING DIAGNOSES IN B.C.

FISCAL 86/87

199.1 Malignant neoplasm, unspecified
221.8 Benign neoplasm, female genital organs, unspecified
*250.0 Diabetes mellitus without mention of complication
366.1 & 366.9 Cataract
*410 & 410.0 Acute myocardial infarction
*411 & 411.0 Other acute or subacute ischemic heart disease
*414.0 Other chronic ischemic heart disease
428.0 Congestive heart failure
428.1 Left heart failure
436.0 Acute cerebrovascular disease
465.9 Acute upper respiratory infection, unspecified site
486.0 Pneumonia, organism unspecified
*493.9 Asthma, unspecified
550.9 Inguinal hernia, without mention of obstruction or gangrene
592.1 Calculus of ureter
600. & 600.0 Hyperplasia of prostate
722.1 Displacement of thoracic or lumbar intervertebral disc without myelopathy
786.5 Chest pain
789.0 Abdominal pain
820.0 Fracture neck of femur, transcervical, closed
.2 Fracture neck of femur, pertrochanteric, closed

*These diagnoses are chosen for the study because they represent a large number of hospital separations and they are fairly representative of the ICD-9 categories.

Cases are selected by randomly selecting first, by individual hospital and then, by diagnostic category.

DISTINCT DIAGNOSTIC CATEGORIES SELECTED FOR THE STUDY

Diseases of Endocrine Glands (250-259)
Ischaemic Heart Disease (410-414)
Chronic Obstructive Pulmonary Disease (490-496)
### APPENDIX I - ICD-9 CATEGORIES FOR STRATIFIED RANDOM SAMPLE

**ICD-9 3-DIGIT CLASS BY GROUP**

<table>
<thead>
<tr>
<th>#</th>
<th>Category</th>
<th>ICD-9 3-DIGIT CODES</th>
</tr>
</thead>
<tbody>
<tr>
<td>#1</td>
<td>Infectious and Parasitic Diseases</td>
<td>(001-139)</td>
</tr>
<tr>
<td>#2</td>
<td>*Endocrine, Nutritional and Metabolic Diseases and Immunity Disorders AND Diseases of the Blood and Blood Forming Organs</td>
<td>(240-289)</td>
</tr>
<tr>
<td>#3</td>
<td>Diseases of the Nervous System and Sense Organs</td>
<td>(320-389)</td>
</tr>
<tr>
<td>#4</td>
<td>Diseases of the Circulatory System</td>
<td>(390-459)</td>
</tr>
<tr>
<td>#5</td>
<td>Diseases of the Respiratory System</td>
<td>(460-519)</td>
</tr>
<tr>
<td>#6</td>
<td>Diseases of the Digestive System</td>
<td>(520-579)</td>
</tr>
<tr>
<td>#7</td>
<td>Diseases of the Genitourinary System</td>
<td>(580-629)</td>
</tr>
<tr>
<td>#8</td>
<td>*Diseases of the Skin and Subcutaneous Tissue AND Diseases of the Musculoskeletal System and Connective Tissue</td>
<td>(680-739)</td>
</tr>
<tr>
<td>#9</td>
<td>Congenital Anomalies</td>
<td>(740-759)</td>
</tr>
<tr>
<td>#10</td>
<td>Injury and poisoning</td>
<td>(800-999)</td>
</tr>
</tbody>
</table>

Obstetrics, Perinatal period, Neoplasms, Signs & Symptoms, and Mental Diseases have not been incorporated.

*Categories have been combined.*
**Hospital Programs Data File**  
(from B.C. Ministry of Health, Hospital Programs)

<table>
<thead>
<tr>
<th>Field Name</th>
<th>Code</th>
</tr>
</thead>
<tbody>
<tr>
<td>Admission Date</td>
<td>(AD-Date)</td>
</tr>
<tr>
<td>Admission Hour</td>
<td>(AD-Hour)</td>
</tr>
<tr>
<td>Patient Number</td>
<td>(P-Num)</td>
</tr>
<tr>
<td>Medical Services Plan Number</td>
<td>(MSP-Num)</td>
</tr>
<tr>
<td>Medical Services Plan Dependent Number</td>
<td>(MSP-DEP)</td>
</tr>
<tr>
<td>Admission Category</td>
<td>(A-CAT)</td>
</tr>
<tr>
<td>Date of Birth</td>
<td>(BIRTHDATE)</td>
</tr>
<tr>
<td>Exit Alive Code</td>
<td>(E-Alive)</td>
</tr>
<tr>
<td>Hospital Code</td>
<td>(HOSP)</td>
</tr>
<tr>
<td>Hospital Transferred From</td>
<td>(HOSP-TF)</td>
</tr>
<tr>
<td>Hospital Transferred To</td>
<td>(HOSP-TT)</td>
</tr>
<tr>
<td>Hospital Level of Care</td>
<td>(Level)</td>
</tr>
<tr>
<td>Hospital Level of Care - Transferred From</td>
<td></td>
</tr>
<tr>
<td>Hospital Level of Care - Transferred To</td>
<td></td>
</tr>
<tr>
<td>ICU Days</td>
<td>(IC-DAYS)</td>
</tr>
<tr>
<td>Rehabilitation Days</td>
<td>(R-DAYS)</td>
</tr>
<tr>
<td>Procedure Codes 1-12</td>
<td>(PROCED 1-12)</td>
</tr>
<tr>
<td>Procedure Date 1-12</td>
<td>(PROC-DATE 1-12)</td>
</tr>
<tr>
<td>Procedure Anesthetic 1-12</td>
<td>(PROC-ANESTH 1-12)</td>
</tr>
<tr>
<td>Procedure Surgeon 1-12</td>
<td>(PROC-SURG 1-12)</td>
</tr>
<tr>
<td>Procedure Anesthetist 1-12</td>
<td>(PROC-ANMD 1-12)</td>
</tr>
<tr>
<td>Most Responsible Doctor</td>
<td>(PHYSIC)</td>
</tr>
<tr>
<td>Postal Code</td>
<td>(POSTAL)</td>
</tr>
<tr>
<td>Principle Diagnosis</td>
<td>(DIAG 1)</td>
</tr>
<tr>
<td>Other Diagnoses 2-16</td>
<td>(DIAG 2-16)</td>
</tr>
<tr>
<td>Diagnosis Type 1-16</td>
<td>(DIAG-TYPE 1-16)</td>
</tr>
<tr>
<td>Responsibility for Payment Code</td>
<td>(RFP)</td>
</tr>
<tr>
<td>Separation Date</td>
<td>(SEP-DATE)</td>
</tr>
<tr>
<td>Separation Hour</td>
<td>(SEP-HOUR)</td>
</tr>
<tr>
<td>Service Code</td>
<td>(SERVICE)</td>
</tr>
<tr>
<td>Sex Code</td>
<td>(SEX)</td>
</tr>
<tr>
<td>Birthweight</td>
<td>()</td>
</tr>
</tbody>
</table>
## Re-abstracting Form

### APPENDIX Q

<table>
<thead>
<tr>
<th>Re-abstracted Code</th>
<th>Ministry Code</th>
<th>Discrepancy</th>
<th>Which is correct?</th>
<th>Source of error</th>
</tr>
</thead>
<tbody>
<tr>
<td>Adm Date YYYY MM DD</td>
<td>YYYY MM DD</td>
<td>0 or 1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Disch Date YYYY MM DD</td>
<td>YYYY MM DD</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Birthdate YYYY MM DD</td>
<td>YYYY MM DD</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Disch Status A SO D</td>
<td>A SO D</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sex M F U</td>
<td>M F U</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**MOST RESPONSIBLE DIAGNOSIS**


**SECONDARY DIAGNOSES**


---

<table>
<thead>
<tr>
<th>a. DISCREPANCY</th>
<th>b. WHICH IS CORRECT?</th>
<th>c. SOURCE OF ERROR</th>
</tr>
</thead>
<tbody>
<tr>
<td>0. Yes</td>
<td>1. Original Abstract</td>
<td>1. Judgement</td>
</tr>
<tr>
<td>1. No</td>
<td>2. Re-abstract</td>
<td>2. Classification</td>
</tr>
<tr>
<td></td>
<td>3. Not determined</td>
<td>3. Procedural</td>
</tr>
<tr>
<td></td>
<td></td>
<td>4. Clerical</td>
</tr>
<tr>
<td></td>
<td></td>
<td>5. Ordering</td>
</tr>
<tr>
<td></td>
<td></td>
<td>6. Incomplete</td>
</tr>
<tr>
<td></td>
<td></td>
<td>7. Other</td>
</tr>
</tbody>
</table>

...cont
APPENDIX R - HOSPITAL QUESTIONNAIRE

NOTE: All information which would permit identification of individual hospitals or personnel will be referred to anonymously, used only for statistical analysis, and will not be used for any other purpose. The questionnaire will be destroyed at the completion of the study. All information in this questionnaire specifically refers to the fiscal year 1986/1987.

Hospital Number ______________ Date ________________________

System
What kind of system did you use:
1. paper abstracting ______________ 2. HMRI computerized abstracting ______________
3. HMRI computerized abstracting connected with ADT system ______________
4. HMRI computerized abstracting connected with Patient Index system ______________
5. other __________________________________

Was the non-medical data downloaded from the ADT or patient index system?
1. Yes __ 2. No __

How long had you been using this system? _______ Start Date ______________

Did your hospital maintain a history file? 1. Yes ____ 2. No __

Which nomenclature did you use? 1. ICD-9 ____ 2. ICD-9CM ____ 3. Other ______
Date ______________________

Which do you use now? 1. ICD-9 ____ 2. ICD-9CM ____ 3. Other ______ Date ______

Procedure Manual
Was there a Coding and Abstracting Policy & Procedure Manual? 1. Yes ____ 2. No __

How often was it updated?
1. annually _______ 2. as problems arose _______
3. other (explain) __________________________________

Were coding books kept up to date? 1. Yes ____ 2. No __

Have coding and abstracting policies and procedures changed since fiscal year 1986/87?
1. Yes _____ 2. No _____
Explain ____________________________________________
Control procedures
How often were control checks done to ensure accuracy of recorded information?
1. Never
2. Occasionally - Random sample
3. Occasionally - Problems only
4. Regular - Daily - Random sample
5. Regular - Daily - Problems Only
6. Regular - Weekly - Random sample
7. Regular - Weekly - Problems Only
8. Regular - Monthly - Random sample
9. Regular - Monthly - Problems only
10. Other (Specify)

Comments

Quality Assurance Program
Did your hospital engage in a QA program for fiscal 86/87? 1. Yes 2. No
When did the program start? Date
What was the extent of the program? Describe

As a result of this QA program, were recommendations made for coding policy or procedures? If so, explain

Timing of Abstracting
Did you code the MRD from incomplete charts? 1. Yes 2. No
Explain

If this is still the process, approximately what percentage of monthly discharges do these represent?

When coding from incomplete charts, were the charts re-coded and errors determined on completion of the record? 1. Yes 2. No

Were these errors corrected for the HMRI database? 1. Yes 2. No

Were correction forms regularly completed and sent to HMRI when requested by them? 1. Yes 2. No
Selection of Most Responsible Diagnosis
In the selection of the MRD, did you routinely use first listed diagnosis on the face sheet?  
1. Yes  
2. No

Did you routinely go through the entire chart?  
1. Yes  
2. No

Was the information changed on the face sheet, if necessary?  
1. Yes  
2. No

If so, was the attending physician consulted each time?  
1. Yes  
2. No

If not, explain

When there was subjectivity regarding the classification codes, who was the ultimate decision maker in determining the MRD?
1. the coder
2. the coding supervisor
3. the Director of the Department
4. the Consultant at HMRI
5. the attending physician
6. other (specify)

Comments:

When there was subjectivity regarding multiple diagnoses, who was the ultimate decision maker in determining the MRD?
1. the coder
2. the coding supervisor
3. the Director of the Department
4. the Consultant at HMRI
5. the attending physician
6. other (specify)

Comments:

Was the medical record maintained in chronological order for coding purposes?
1. Yes  
2. No

Was the entire coding and abstracting procedure done in one stage by one individual?
1. Yes  
2. No

Explain
Hospital Policies
Did you code non-operative procedures? 1. Yes __ 2. No __

Did you have a quota of charts for each coder to code each day? 1. Yes __ 2. No __
If so, what was it? _____ Has this changed? 1. Yes __ 2. No __ To what? _____

Was this quota determined through workload measurement studies? 1. Yes __ 2. No __
If not, how was it determined?

Did you keep listings available for coders, such as ranges of normal investigative reports or incubation periods for communicable diseases? 1. Yes __ 2. No __
If not, how were these determined?

Was your definition of late effect > 1 year, unless specified by a physician? 1. Yes __ 2. No __
If not, what was your definition of late effect?

Did your hospital guidelines recommend coding homosexuality and drug dependence? 1. Yes __ 2. No __

If the pathology report differs from principal diagnosis, what was your procedure? (e.g., acute appendicitis with appendectomy, but normal appendix shown on path report)

What was your guideline on complications?

Did you code optional data? 1. Yes __ 2. No __ If so, explain

Did you code "*" codes? 1. Yes __ 2. No __

Were there guidelines different from the code book (i.e., hospital-based decisions to code a specific way? 1. Yes __ 2. No __ Explain

APPENDIX R - HOSPITAL QUESTIONNAIRE - Page 5

HOSPITAL #

Credentials of Supervisor
Were coders supervised? 1. yes  2. no

What were the credentials, place of training and length of time coding since graduation?
1. ART/HRT  , 2. HRA  , 3. HDT  , 4. MRL  , 5. other (specify)
Coder # , place of training , length of time coding
Comments

Credentials of Coders - Full-time Employees for Full Year of Study
How many coders in terms of full-time equivalents were employed for the entire year of the study (exclude supervisor)?

What were the credentials of each full-time employee employed for the full year, place of training and length of time coding since graduation?

1. ART/HRT  , 2. HRA  , 3. HDT  , 4. MRL  , 5. other
Coder # , place of training , length of time coding

1. ART/HRT  , 2. HRA  , 3. HDT  , 4. MRL  , 5. other
Coder # , place of training , length of time coding

1. ART/HRT  , 2. HRA  , 3. HDT  , 4. MRL  , 5. other
Coder # , place of training , length of time coding

1. ART/HRT  , 2. HRA  , 3. HDT  , 4. MRL  , 5. other
Coder # , place of training , length of time coding

1. ART/HRT  , 2. HRA  , 3. HDT  , 4. MRL  , 5. other
Coder # , place of training , length of time coding

1. ART/HRT  , 2. HRA  , 3. HDT  , 4. MRL  , 5. other
Coder # , place of training , length of time coding

1. ART/HRT  , 2. HRA  , 3. HDT  , 4. MRL  , 5. other
Coder # , place of training , length of time coding

1. ART/HRT  , 2. HRA  , 3. HDT  , 4. MRL  , 5. other
Coder # , place of training , length of time coding

1. ART/HRT  , 2. HRA  , 3. HDT  , 4. MRL  , 5. other
Coder # , place of training , length of time coding

1. ART/HRT  , 2. HRA  , 3. HDT  , 4. MRL  , 5. other
Coder # , place of training , length of time coding

(If more space is needed, use back of page)
APPENDIX R - HOSPITAL QUESTIONNAIRE - Page 6

HOSPITAL #    

Credentials of coders - Full-time Employees Employed Part of the Year
How many full-time coders (FTE's) were employed for part of the year. __________

What was credentials of each full-time coder, how long were they employed during fiscal 1986/87, their place of training, and length of time coding since graduation?

1. ART/HRT __, 2. HRA ___, 3. HDT ___, 4. MRL ___, 5. other (specify) ______
Coder # _____, how long employed _____, place of training ________________________,
length of time coding ________________________

1. ART/HRT __, 2. HRA ___, 3. HDT ___, 4. MRL ___, 5. other (specify) ______
Coder # _____, how long employed _____, place of training ________________________,
length of time coding ________________________

1. ART/HRT __, 2. HRA ___, 3. HDT ___, 4. MRL ___, 5. other (specify) ______
Coder # _____, how long employed _____, place of training ________________________,
length of time coding ________________________
(If more space needed, use back of page)

Credentials of coders - Part-time and Casual Employees
How many part-time coders (FTE's) were employed for the full year ______ and how many (FTE’s) were employed for part of the year. ______

What was credentials of each part-time or casual coder, how long were they employed during fiscal 1986/87, FTE, their place of training, and length of time coding since graduation?

1. ART/HRT ___, 2. HRA ___, 3. HDT ___, 4. MRL ___, 5. other (specify) ______
Coder # _____, how long employed _____, place of training ________________________,
length of time coding ________________________, FTE ______

1. ART/HRT ___, 2. HRA ___, 3. HDT ___, 4. MRL ___, 5. other (specify) ______
Coder # _____, how long employed _____, place of training ________________________,
length of time coding ________________________, FTE ______

1. ART/HRT ___, 2. HRA ___, 3. HDT ___, 4. MRL ___, 5. other (specify) ______
Coder # _____, how long employed _____, place of training ________________________,
length of time coding ________________________, FTE ______
(If more space needed, use back of page)
Training and Orientation
Was there an on-site training program for newly hired employees? 1 Yes 2. No Explain process or attach copy of policy

Was there in-service education available for coding staff? 1. Yes 2. No Specify

HMRI Training Programs
How frequently did your staff participate in HMRI training programs, seminars, teleconferences, or workshops in 24 months prior to April, 1986 and including the 12 month period of study (full 3 year period)?
1. ART/HRT __, 2. HRA __, 3. HDT __, 4. MRL __, 5. other (specify) __________
Coder # _____, Describe the programs

1. ART/HRT __, 2. HRA __, 3. HDT __, 4. MRL __, 5. other (specify) __________
Coder # _____, Describe the programs

1. ART/HRT __, 2. HRA __, 3. HDT __, 4. MRL __, 5. other (specify) __________
Coder # _____, Describe the programs

1. ART/HRT __, 2. HRA __, 3. HDT __, 4. MRL __, 5. other (specify) __________
Coder # _____, Describe the programs

1. ART/HRT __, 2. HRA __, 3. HDT __, 4. MRL __, 5. other (specify) __________
Coder # _____, Describe the programs

1. ART/HRT __, 2. HRA __, 3. HDT __, 4. MRL __, 5. other (specify) __________
Coder # _____, Describe the programs

1. ART/HRT __, 2. HRA __, 3. HDT __, 4. MRL __, 5. other (specify) __________
Coder # _____, Describe the programs

1. ART/HRT __, 2. HRA __, 3. HDT __, 4. MRL __, 5. other (specify) __________
Coder # _____, Describe the programs

1. ART/HRT __, 2. HRA __, 3. HDT __, 4. MRL __, 5. other (specify) __________
Coder # _____, Describe the programs

1. ART/HRT __, 2. HRA __, 3. HDT __, 4. MRL __, 5. other (specify) __________
Coder # _____, Describe the programs

(If more space needed, use back of page)
Coding Sessions Other than Through HMRI
How frequently did coders attend coding training other than through HMRI over the same 3 year period?

1. ART/HRT __, 2. HRA __, 3. HDT __, 4. MRL __, 5. other (specify)____________
Coder # ____, Describe the programs _________________________________________

__________________________________________________________

1. ART/HRT__, 2. HRA __, 3. HDT __, 4. MRL __, 5. other (specify)____________
Coder # ____ , Describe the programs _________________________________________

__________________________________________________________

1. ART/HRT __, 2. HRA __, 3. HDT __, 4. MRL __, 5. other (specify)____________
Coder # ____ , Describe the programs _________________________________________

__________________________________________________________

1. ART/HRT __, 2. HRA __, 3. HDT __, 4. MRL __, 5. other (specify)____________
Coder # ____ , Describe the programs _________________________________________

__________________________________________________________

(If more space needed, use back of page)

FOR ADDITIONAL COMMENTS - USE BACK OF PAGE
APPENDIX S - SUMMARY OF STUDY PLAN

1. Proposal Preparation & timetable determined
2. Letter to MOH regarding use of data
3. Letter to HMRI re training session
4. Letter to each individual hospital formally requesting participation
5. Approval by Thesis Committee Members
6. Approval by HCEP Thesis Screening Panel
7. Approval by Ethical Review Committee
8. Approval by individual hospital research committees
9. Principal Investigator training (HMRI - 2 days)
10. Updating diagnostic and procedure coding books
11. Review code books
12. Determine rater testing with representative group
13. Determine timetable with each hospital
14. Computer analyst draws sample from MOH database for Selection of Cases
15. Listing of selected cases to each hospitals at least one week prior to arrival of research team
16. Health record clerk - Pulls charts
17. Assistant checks to ensure correct sample
18. Administer Hospital Questionnaire
19. Assistant replaces face sheet and deletes all codes
20. Records re-examined, re-coded, and re-abstracted
21. Assistant enters original data from MOH on re-abstracting form
22. Reconciliation stage - Data is compared and discrepancies noted
23. Type of discrepancy determined
24. Review with Supervisor Coders
25. Face sheets replaced and charts re-filed
26. Intra-rater reliability testing - 10% Subsample re-coded and re-abstracted by Investigator - One week after initial stage completed
27. Data entry
28. Analysis of data
29. Results obtained
30. Writing of report
31. Approval from Thesis Committee
32. Final approval and mark from HCEP Thesis Screening Panel
33. Thesis filed in Library
MINUTES

Thesis Screening Panel Meeting

June 10, 1991
9:00 am - 12:30 pm

Attending: Morris Barer (chair), Geoff Anderson, Godwin Eni (until 9:45), Brenda Morrison, Konnie Peet, Chris van Netten.

1 Linda Brown Thesis Proposal

Title: "Reliability of HMRI Database: A Re-abstracting Study"

The Panel determined that this proposal be conditionally approved, subject to the student being able to work through a number of fundamental design issues to the satisfaction of her Committee. It was noted that the proposal lacked two signatures (G. Anderson; Barer).

Specifically, the student must satisfy the following conditions in order that the proposal be formally approved by the Panel:

(a) ascertain whether there is an abstracting test that would establish the student’s placement in a distribution of skills of the abstracters with which she will be compared; undergo the test; satisfy her Committee that she can be considered to possess at least equivalent skills at the abstracting in which she will be involved;

(b) if the student cannot establish herself as an acceptable coder/abstracter for the purposes of the thesis, she will need to identify an alternative source of coding/abstracting, that is able to meet the 'quality' test noted above;

(c) design and sample size details to be sorted out with Committee;

(d) signatures of all Committee members needed on proposal before it is filed with Department.

Either (a) or (b), and (c) must be worked out to the satisfaction of the student’s Committee, at which point the Committee will report to the Panel that the student has met the conditions set out by the Panel.
BACK TO THE BASICS WITH ICD-9
A CODER'S REVIEW WORKSHOP

AGENDA

1. INTRODUCTION
2. CHAPTER BY CHAPTER REVIEW
3. CODING EXERCISES
4. PROCEDURE CODING GUIDELINES
5. RULES FOR SELECTION OF MOST RESPONSIBLE DIAGNOSIS AND PRINCIPLE PROCEDURE
6. CONCLUSION
A DIAGNOSIS TYPING WORKSHOP

AGENDA

1. WELCOME AND INTRODUCTION
2. SELECTION CRITERIA FOR MOST RESPONSIBLE DIAGNOSIS
3. EXERCISE - SELECTION OF MOST RESPONSIBLE DIAGNOSIS
4. DIAGNOSIS TYPING DEFINITIONS REVIEWED
5. THE IMPACT OF DIAGNOSIS TYPING DECISIONS
   - HOW ARE CMG'S AND RIW'S USED BY HOSPITAL ADMINISTRATION
   - DIAGNOSIS TYPES WHICH DO (DON'T) AFFECT CMG AND RIW ASSIGNMENT
   - CMG RELATED ERROR MESSAGES
6. DIAGNOSIS TYPING WORKSHOPS
7. NEWBORN CODING AND DIAGNOSIS TYPING GUIDELINES
8. ADJOURNMENT