Reliability of Hospital Discharge Abstracts (1977)

Pages
128

Size
8.5 x 10

ISBN
0309026148

Institute of Medicine

INSTITUTE OF MEDICINE

REPORT OF A STUDY

Reliability of Hospital Discharge Abstracts

February 1977

NATIONAL ACADEMY of SCIENCES

Washington, D.C.

NOTICE

The project that is the subject of this report was approved by the Governing Board of the National Research Council, whose members are drawn from the Councils of the National Academy of Sciences, the National Academy of Engineering, and the Institute of Medicine. The members of the Committee responsible for the report were chosen for their special competences and with regard for appropriate balance.

This report has been reviewed by a group other than the authors according to procedures approved by a Report Review Committee consisting of members of the National Academy of Sciences, the National Academy of Engineering, and the Institute of Medicine.

Supported by U.S. Department of Health, Education, and Welfare Contract No. 282-75-0437 PM.

The Institute of Medicine was chartered in 1970 by the National Academy of Sciences to enlist distinguished members of appropriate professions in the examination of policy matters pertaining to the health of the public. In this, the Institute acts under both the Academy's 1863 Congressional charter responsibility to be an advisor to the Federal Government, and its own initiative in identifying issues of medical care, research, and education.

Publication IOM-77-01

Library of Congress Catalog Number 77-76685
ISBN 0-309-02614-8

Available from:

Printing and Publishing Office
National Academy of Sciences
2101 Constitution Avenue, N.W.
Washington, D.C. 20418

Printed in the United States of America

NATIONAL ACADEMY OF SCIENCES
2101 CONSTITUTION AVENUE
WASHINGTON, D. C. 20418

INSTITUTE OF MEDICINE
OFFICE OF THE PRESIDENT

February 28, 1977

The Honorable Harrison A. Williams, Jr.
Chairman
Committee on Human Resources
United States Senate
Washington, D. C. 20501

Dear Mr. Chairman:

I am pleased to present to the Committee on Human Resources the final report of one component of a study of health care quality assurance programs conducted by the Institute of Medicine, National Academy of Sciences, pursuant to Section 4 of the Health Maintenance Organization Act of 1973 (P.L. 93-222).

The enclosed volume presents the findings of an assessment of the reliability of hospital utilization information generated by private abstracting services and based on abstracts of the hospital medical record. The purpose was to determine the usefulness of such information for evaluating the impact of Professional Standards Review Organizations.

We will be happy to discuss the report in greater detail with the members and staff of your committee.

Sincerely yours,

David A. Hamburg, M.D.
President

Enclosure

NATIONAL ACADEMY OF SCIENCES
2101 CONSTITUTION AVENUE
WASHINGTON, D. C. 20418

INSTITUTE OF MEDICINE
OFFICE OF THE PRESIDENT

February 28, 1977

The Honorable Harley O. Staggers
Chairman
Committee on Interstate and Foreign Commerce
U. S. House of Representatives
Washington, D. C. 20515

Dear Mr. Chairman:

I am pleased to present to the Committee on Interstate and Foreign Commerce the final report of one component of a study of health care quality assurance programs conducted by the Institute of Medicine, National Academy of Sciences, pursuant to Section 4 of the Health Maintenance Organization Act of 1973 (P.L. 93-222).

The enclosed volume presents the findings of an assessment of the reliability of hospital utilization information generated by private abstracting services and based on abstracts of the hospital medical record. The purpose was to determine the usefulness of such information for evaluating the impact of Professional Standards Review Organizations.

We will be happy to discuss the report in greater detail with the members and staff of your committee.

Sincerely yours,

David A. Hamburg, M.D.
President

Enclosure

NATIONAL ACADEMY OF SCIENCES
2101 CONSTITUTION AVENUE
WASHINGTON, D. C. 20418

INSTITUTE OF MEDICINE
OFFICE OF THE PRESIDENT

February 28, 1977

James F. Dickson, III, M.D.
Acting Assistant Secretary for Health
Department of Health, Education, and
Welfare
200 Independence Avenue, S.W.
Washington, D. C. 20201

Dear Dr. Dickson:

I am pleased to present to the Department of Health, Education and Welfare the final report of one component of a study of health care quality assurance programs conducted by the Institute of Medicine, National Academy of Sciences, pursuant to Section 4 of the Health Maintenance Organization Act of 1973 (P.L. 93-222).

The enclosed volume presents the findings of an assessment of the reliability of hospital utilization information generated by private abstracting services and based on abstracts of the hospital medical record. The purpose was to determine the usefulness of such information for evaluating the impact of Professional Standards Review Organizations.

We will be happy to discuss the report in greater detail with you or members of your staff.

Sincerely yours,

David Hamburg

David A. Hamburg, M.D.
President

Enclosure

RELIABILITY
OF HOSPITAL
DISCHARGE ABSTRACTS

TECHNICAL ADVISORS

Faye BROWN, R.R.A., Director, Medical Information Service, St. Helena Hospital and Health Center, Deer Park, California

Jacob J. FELDMAN, Ph.D., Associate Director for Analysis, National Center for Health Statistics, Rockville, Maryland

Donald C. RIEDEL, Ph.D., Chairman of the Center for the Study of Health Services, Yale University, New Haven, Connecticut

Joseph STEINBERG, President, Survey Design, Inc., Silver Spring, Maryland

George WERKING, Mathematical Statistician, Bureau of Labor Statistics, Washington, D.C.

STEERING COMMITTEE

Robert J. HAGGERTY, M.D., (Chairman) Roger I. Lee Professor of Public Health, Harvard School of Public Health, Boston, Massachusetts

*Carl EISDORFER, Ph.D., M.D., Professor and Chairman, Department of Psychiatry and Behavioral Sciences, School of Medicine, University of Washington, Seattle, Washington

Jacob J. FELDMAN, Ph.D., Associate Director for Analysis, National Center for Health Statistics, Rockville, Maryland

John R. GAMBLE, M.D., Chairman, Department of Medicine, Pacific Medical Center, San Francisco, California

*Melvin A. GLASSER, LL.D., Director, Social Security Department, United Auto Workers, International Union, Detroit, Michigan

Sidney KATZ, M.D., Professor and Director, Office of Health Services Education and Research, College of Human Medicine, Michigan State University, E. Lansing, Michigan

*B. Harvey MINCHEW, M.D., Internist, Ellicott City, Maryland

Alan R. NELSON, M.D., Internist, Salt Lake City, Utah

Howard N. NEWMAN, J.D., President, Dartmouth-Hitchcock Medical Center, Hanover, New Hampshire

Donald C. RIEDEL, Ph.D., Chairman of the Center for the Study of Health Services, Yale University, New Haven, Connecticut

*Marc J. ROBERTS, Ph.D., Professor of Political Economy and Health Policy, Academic Director of the Executive Programs in Health Policies and Management, Harvard University, Boston, Massachusetts

Sam SHAPIRO, B.S., Director, Health Services Research and Development Center, Professor of Health Care Organization, Department of Health Services Administration, The Johns Hopkins University, Baltimore, Maryland

Barbara STARFIELD, M.D., Professor and Division Head, Health Care Organization and Associate Professor of Pediatrics, The Johns Hopkins University, School of Hygiene and Public Health, Baltimore, Maryland

*Participation was limited to the planning phases of the study.

STAFF

Linda K. DEMLO, Ph.D., Study Director

Sarah Spaght BROWN, M.P.H.

Paul M. CAMPBELL, M.S.

FIELD TEAM

Judy HILL, R.R.A., Lynchburg, Virginia

Sally PERRINE, R.R.A., West Coast Cancer Foundation, Emeryville, California

Barbara PUDER, R.R.A., St. Louis University, Department of Medicine, St. Louis, Missouri

CONTENTS

FOREWORD

As part of its previously published study on "Assessing Quality in Health Care: An Evaluation," the Institute of Medicine was asked by the Department of Health, Education, and Welfare to assess the reliability of hospital utilization data compiled by private abstracting services and based on abstracts of medical records. The request stemmed from the need to identify an existing data base with sufficient reliability to serve as a baseline for evaluating the effectiveness of Professional Standards Review Organizations, since baseline data had not been gathered prior to instituting the PSRO program. However, hospital discharge abstracts are potentially useful for broader health services research, administration, and policy formulation activities, as well. It was in this broader context that the Institute agreed to conduct the study.

The findings indicate that for the study year (1974), the reliability of nationally aggregated, abstracted information varies. While some data items are quite reliable, others are not. Uniform, comprehensive, and reliable data bases are essential in order to plan for programs to meet the health needs of the nation and evaluate their effectiveness. If one assumes that the levels of accuracy of the 1974 data persist, a concerted effort is required by the health community to improve medical recording, abstracting, and information processing practices, so that abstracted data can be used to their full potential.

The study also highlights the need to design prospective evaluations of major innovations in the delivery of health and other social services and to gather baseline data prior to program implementation. This is the only way to assure the availability of adequate information to assess the effectiveness of such programs and to determine whether they warrant continuing public expenditures.

Robert J. Haggerty, M.D.
Chairman, Steering Committee

ACKNOWLEDGMENTS

This study would not have been possible without the cooperation of the fifty participating hospitals and the abstracting services to which they subscribe: the California Health Data Corporation, the Commission on Hospital and Professional Activities, the Hospital Utilization Project, and the QUEST system of Blue Cross of Northeast Ohio. Numerous individuals within each service were extremely helpful in explaining their respective abstracting and data processing procedures and selecting the samples of abstracts. Their interest and support is gratefully acknowledged.

Special thanks go to Jacob J. Feldman and Donald C. Riedel, who served in a dual advisory capacity. In addition to their broader steering committee responsibilities, they devoted many hours to reviewing the methods and analysis of the re-abstracting study and shaping the final recommendations.

Faye Brown's assistance in recruiting the field team and developing the re-abstracting procedures used in the field work is also deeply appreciated. The three field team members followed arduous travel schedules and spent long hours in medical record departments in order to complete the field work within the allotted time.

Within the Department of Health, Education, and Welfare we wish to thank Martin A. Baum of the Office of Quality Standards, who was instrumental in initiating the study request and monitoring its progress. At the National Center for Health Statistics, Delray Green helped delineate differences between the ICDA-8 and H-ICDA diagnostic classifications; W. Edward Bacon reviewed the study design and provided technical information used by Joseph Steinberg in developing the sampling methodology.

For data processing and support services, we are indebted to the staff of the Commission on Human Resources of the National Academy of Sciences. Susan Seefried's programming assistance was particularly valuable.

Sandra Matthews deserves special recognition for her skillful administrative assistance throughout the course of the study and her important role in coordinating the field work and preparing the final report.

Linda K. Demlo, Ph.D.
Study Director

Chapter 1

INTRODUCTION

The study reported here was conducted by the Institute of Medicine at the request of the Office of Quality Standards and the Office of Planning, Evaluation and Legislation, both in the Department of Health, Education, and Welfare. It was administered as part of a larger study focused on broader issues of quality assurance in medical care as requested by Congress in Public Law 93-222. 1/ Because the purpose and methods of this project were quite different from the larger study, the findings are reported separately.

ORIGINS AND OBJECTIVES

The purpose of this study was to assess the reliability of hospital utilization data compiled by private abstracting services and based on abstracts of medical records. This information was needed by the Office of Quality Standards to identify an existing and accurate source of data to serve as a baseline for measuring the impact of Professional Standards Review Organizations (PSROs), since baseline data were not gathered prior to implementation of the PSRO program. 2/ Several potential sources of data were identified which would help to determine whether PSROs have a demonstrable influence on costs and utilization of health services, and further, whether the benefits are worth the investment of public funds.

1
Institute of Medicine, Assessing Quality in Health Care: An Evaluation (Washington, D. C.: National Academy of Sciences, November 1976).

2
The establishment of PSROs was authorized by Congress in 1972 through Public Law 92-603. PSROs are intended to assure that medical services financed by Medicare, Medicaid, and Maternal and Child Health Programs conform to appropriate professional standards, are medically necessary, and in the case of inpatient services, could not have been performed equally effectively on an outpatient basis or in an inpatient facility of a different type.

The baseline data option which seemed most practical in terms of cost and accessibility was utilization data routinely gathered and processed by private abstracting services, provided the data were sufficiently reliable.

Because the need to evaluate the PSRO program provided the initial impetus for the re-abstracting study, many of the study's methodological details were influenced by the PSRO Evaluation Plan. 3/ Implications of the findings, however, are much broader than the evaluation of PSROs.

There are about 30 abstracting services in the United States which process hospital utilization information for subscriber hospitals. Typical reports specify the types of services used, the length of stay, and patient status upon discharge for each hospital admission. The data are abstracted from medical records by hospital personnel when patients are discharged, according to a prescribed format. The information is forwarded to the abstracting service for editing and processing and eventually returned to the hospital in the form of statistical profiles. The profiles enable hospital administrators to determine the manner in which hospital resources are used to provide patient care and, frequently, how the patient care patterns of a particular hospital compare with those of similar hospitals. The comparative information may identify areas in which a given hospital's performance differs from other hospitals. This knowledge, in turn, may lead to a more careful analysis of the reasons for differences and eventually result in internal hospital improvements.

Many of the larger abstracting services have been in existence for several years and have the capability of providing longitudinal information about utilization patterns in their subscriber hospitals. The data can be aggregated at both regional and national levels and are potentially useful for addressing health services research and policy issues, in addition to internal hospital management problems. Prior to this study, the question of accuracy or reliability particularly for national estimates, had not been systematically examined.

Specific Issues Addressed

The re-abstracting study had two primary objectives:

- To determine the reliability of specified information items included in the original abstract for use in national statistical estimates; and

---

3/ U.S. Department of Health, Education, and Welfare, Office of the Assistant Secretary for Health, Office of Professional Standards Review, Program Evaluation Plan: Professional Standards Review Organizations, by Martin A. Baum et al. (22 September 1975).

- To assess selected characteristics of the abstracting process within the hospital and interactions between the hospital and abstract service, which may be associated with varying levels of data reliability.

The first objective was intended to assist in determining whether the abstracted data were sufficiently accurate to constitute the baseline for assessing the impact of PSROs over time. Accuracy was measured by comparing the results of an independent abstracting of medical records by a trained field team with the original abstract provided by the abstract service and noting the frequency and types of discrepancies between them.

The second objective was included to assist in identifying areas for improvement by either the hospital or abstract service, which could increase the reliability of future data. It was further viewed as a protection against erroneously associating poor reliability of data with an abstract service when factors within the hospital may have been responsible. Finally, if specific hospital characteristics (for example, the level of training of coders completing the abstracts) were identified with poor quality data, it might then be possible to find alternative sources of data for such hospitals. Within each study hospital, the medical record department supervisor completed a questionnaire, which provided the information to address this objective.

An important issue not addressed by this study is the antecedent question of the quality of the medical record itself. Abstracted information clearly can be only as accurate as the record from which it is drawn. If important information is ambiguously noted or absent, the abstract will be equally inadequate. There is also some evidence to suggest that the quality of medical record keeping is related to the quality of care delivered. 4/ These issues were outside the scope of the current study, but are important and deserve further consideration.

4
See for example, Alan E. Zuckerman et al., "Validating the Content of Pediatric Out-patient Medical Records by Means of Tape-Recording Doctor-Patient Encounters," Pediatrics 56 (September 1975).

Chapter 2

RELATED RESEARCH

The reliability of hospital discharge information abstracted from medical records has not been analyzed extensively. However, several studies were helpful in defining the scope of this study and outlining appropriate research methods. (See Appendix A for Bibliography.)

The Uniform Hospital Discharge Data Demonstration (UHDDD) sought to demonstrate the feasibility of collecting a minimum basic data set with uniform definitions using existing hospital discharge abstract systems. 1/ In addition, the study was intended to determine the degree to which these data could satisfy a multiplicity of user requirements. The findings confirmed the feasibility of using the specified data set. Subsequently, the data items formed the basis for the PSRO information reporting requirements and, in turn, were incorporated in the Institute of Medicine's study.

One phase of the demonstration required a re-abstracting of a sample of medical records at several test sites to determine the success of the hospital coders in applying the new UHDDD definitions. 2/ The UHDDD concepts and procedures were reviewed carefully and some were adapted for this study. The estimates of error rates by abstract item were considered in determining the sample size. A revision of the UHDDD method for analyzing discrepancies between the initial abstracts and re-abstracts was also utilized. There are two principal limitations to the UHDDD reliability assessment for answering the current questions, however: the hospitals which participated in the demonstration were volunteers and could be expected to have a higher performance level

---

1 David A. Hodgson, Lawrence E. Kucken, and James M. Ensign, The Uniform Hospital Discharge Data Demonstration Summary Report, U. S. Department of Health, Education, and Welfare, Public Health Service, Pubn. No. HRA-74-3102, July 1973.

2 Benjamin C. Duggar, Evaluation of the Uniform Hospital Discharge Abstract Site Test, (Springfield, Va.: National Technical Information Service #PB-223 405, March 1973).

than would be found in a national probability sample of hospitals; and the UHDDD studied how successfully a small group of hospitals was able to apply new definitions for abstract items, and not the reliability of nationally aggregated data processed by abstracting services.

In a related study, Hendrickson and Myers reviewed utilization statistics provided by an abstract service for over 16,000 admissions to an urban teaching hospital during one year. 3/ After noting apparent errors, a random sample of medical records was examined to ascertain the reliability of the abstracted information. The investigation was restricted to the coding of routine laboratory tests, particularly for patients from diagnostic groupings suspected of having errors, based on the initial review of aggregate statistics. The authors uncovered a "variable but significant error rate" in the recording of each entry studied, and concluded that such errors "caused the data to be unsuitable for a quality of care study since the error rate was either very high or there was evidence of a recording bias." 4/

Hendrickson and Myers noted the uneven quality, background, and training of coders as an important cause of inadequate data. They concluded that "possible errors in data recording could be recognized most efficiently through check programs at the abstracting service and that by using such programs, the abstracting service could and should initiate improvements in hospital practices and coder training to minimize errors." 5/ Since the research was conducted in one hospital only and for one year (1969), its generalizability is limited. Most major abstracting services have refined their data collection and analysis capabilities and introduced more careful edit checks since that time. In addition, the coding of diagnoses and procedures, which are the abstracted items of primary importance for the current PSRO evaluation plan, was not examined in detail.

More recently the Cooperative Health Information Center of Vermont (CHIC) examined the reliability and internal consistency of abstracted hospital discharge data. 6/ Computer edit checks were used to examine incompatibilities

---

3
Leslie Hendrickson and Jeffrey Myers, "Some Sources and Potential Consequences of Errors in Medical Data Recording," Methods of Information in Medicine 12 (January 1973): 38-45.

4
Hendrickson and Myers, p. 38.

5
Ibid.

6
Cooperative Health Information Center of Vermont, Inc., Notes on Data Quality, South Burlington, Vt., September 1974.

in the coding of age, sex, diagnoses, and procedures for some 95,000 abstracts, as an indirect measure of coding accuracy in situations where a more direct assessment of reliability would not be possible. The compatibility between final diagnoses and surgical procedures was analyzed. Finally, diagnostic information from abstracts for patients who died in the hospital was compared with cause of death statements on death records. Discrepancies were greatest for this final analysis.

The CHIC data suggest that reliability varies by hospital, strengthening the argument that the quality of individual coders and the care with which abstracting is conducted within hospitals is a key variable in determining reliability. The authors also note that the acceptability of varying degrees of data accuracy is closely related to the purposes for which the data are used.

Two of the abstract services included in this study have recently conducted their own internal analyses of data reliability through limited re-abstracting studies.

The Commission on Professional and Hospital Activities (CPHA) in Ann Arbor has analyzed the consistency and accuracy of coding and abstracting and the relationship of such information to selected hospital and medical record department characteristics, such as the training and experience of coders and abstractors. 7/ The study was based on a nationally representative sample of 171 hospitals and a diagnostic-specific sample of 1,444 abstracts. Participating hospitals were asked to re-abstract a sample of records. When the initial CPHA abstract was compared to the re-abstract (both completed by hospital personnel), 91 percent of the codes for final diagnosis agreed through four digits. Because of the high level of agreement, CPHA questioned whether participating hospitals had in fact re-abstracted the specified records, which demands a detailed review of the chart, or whether codes placed on the face sheet at the time of initial abstracting had merely been copied on the re-abstract. Accordingly, all hospitals were contacted again and asked about their method of completing the re-abstract. Ninety-three hospitals indicated that they had actually re-abstracted the information. For those hospitals the consistency rate for final diagnosis coded to four digits was 88 percent. This figure may over-estimate coding accuracy, since it is based on two sessions of abstracting, presumably conducted by the same hospital employees, following their customary abstracting procedures. Only an independent re-abstracting of records could address this issue, however.

---

7 Roland J. Loup, "Consistency of Final Diagnosis Coding in a Hospital Patient Record Data System," paper presented at the Seventh Annual Interdisciplinary Conference on Health Records, Chicago, Ill., June 1976.

The Hospital Utilization Project (HUP) in Pittsburgh, Pennsylvania has also examined the reliability of its own abstracted information. 8/ Preliminary results of this study should be available in the near future.

Regardless of the differences among these studies, each noted the importance of careful training of hospital abstractors and meticulous editing, preferably by computer, once abstracts have been completed. The reliability estimates for selected items abstracted from the medical record and their accompanying methods provide useful contributions to the body of knowledge about data systems. Nevertheless, the reliability of hospital discharge abstracts for estimating national utilization patterns was not determined. This, then, was the charge of the Institute of Medicine's re-abstracting study.

8
Jean Bartholo, R.R.A., Hospital Utilization Project (HUP), Pittsburgh, telephone conversation, Fall 1975.

Chapter 3

STUDY METHODS

In the Institute of Medicine's re-abstracting study a field team independently re-abstracted selected medical records within participating study hospitals. The results of the independent re-abstracting were compared with information on original abstracts obtained from private abstracting services. Discrepancies between the two abstracts were noted and the medical records were re-examined in an attempt to understand the reasons for discrepancies. These data, supplemented by information on abstracting procedures within study hospitals, constituted the basis for the analysis.

More detailed information on the sampling plan, survey instruments, field work, and data processing is provided below.

SAMPLING PLAN

A three-stage sampling plan was used, which included an initial national sample of hospitals suitable for use in the larger PSRO evaluation; a smaller subsample of hospitals which were included in this study; and within each study hospital, a sample of discharge abstracts, which formed the basis for the re-abstracting. (See Appendix B for a more extensive discussion of the sample design.)

Initial National Sample of Hospitals

The initial sample of hospitals for the larger PSRO evaluation plan was chosen from the 5,944 non-federal, short-term general hospitals included in the 1973 Master Facilities Inventory (MFI), compiled by the National Center for Health Statistics. The MFI contains 27 hospitals with 1,000 or more beds. Because of the importance of these hospitals in treating Medicaid patients and the likelihood that they might use internal data systems, rather than subscribe to abstracting services, all 27 hospitals were included in the sample with certainty. They are referred to as "certainty" hospitals. From the remaining MFI hospitals, a national probability sample was drawn, using a two-way controlled

selection process, in order to increase the precision of study estimates beyond that resulting from simple stratification. 1/

The following stratification variables were used:

- nine census divisions;
- hospital bed size: under 50 beds, 50-99, 100-199, 200-299, 300-499, and 500-999 beds;
- population density: SMSA and non-SMSA; and
- ownership: government, voluntary, and proprietary.

The resulting sample consisted of ten panels of hospitals, each of which constitutes a national sample in its own right. Each panel can be combined with other panels, plus the certainty hospitals, to create a range of representations of the national hospital universe, depending on the sample size and level of precision desired. Two of the ten panels were randomly selected and combined with the certainty hospitals to serve as the sampling frame for the re-abstracting study.

Subsample for Re-abstracting Study

Selection of the subsample of hospitals for inclusion in the re-abstracting study was governed by the desire to include only hospitals which subscribe to an abstracting service and to limit the number of abstract services included, while maintaining reasonable national coverage and staying within resource constraints.

The abstract services selected for inclusion in the study were chosen because together they cover about 65 percent of all discharges from short-stay general hospitals and represent a reasonable national geographic spread. They include the Commission on Professional and Hospital Activities in Ann Arbor, whose abstracting system, the Professional Activity Study (PAS), is used throughout the United States, but slightly less in the southern states; the Hospital Utilization Project (HUP) in Pittsburgh, which includes hospitals in Pennsylvania as well as several southern and eastern states; the California Health Data Corporation (CHDC) in Sacramento, which covers California and selected western states; and the QUEST system of Blue Cross of Northeast Ohio, which covers selected areas in the midwest.

---

1 R. Goodman and L. Kish, "Controlled Selection - A Technique in Probability Sampling," Journal of the American Statistical Association, 45 (September 1950): 350-72; see also Irene Hess, Donald C. Riedel, and Thomas B. Fitzpatrick, Probability Sampling of Hospitals and Patients, 2nd ed. (Ann Arbor: Health Administration Press, 1975).

Each of the 708 hospitals included in the two panels which constituted the sampling frame was compared against lists of subscribers to the abstracting services to identify hospitals which had subscribed to one of the four services for all of calendar year 1974. 2/ From these 196 eligible hospitals, a final sample of 65 hospitals was drawn for inclusion in the re-abstracting study, using an approximation to controlled selection. These 65 hospitals, plus five hospitals with internal data systems, were asked in writing and by telephone to participate in this study. Fifty agreed. A copy of the letter to the hospital administrator is in Appendix C.

Table 1 shows the initial sample and the number of hospitals which eventually agreed to participate by abstract service. Despite the care with which the abstract services defined the frames for selecting hospitals, after the initial contact it became apparent that six were not eligible for inclusion in the study and were removed from the sample. One hospital was erroneously classified as a QUEST participant. Two CHDC hospitals had not subscribed to the service for all of 1974. Three hospitals with internal data systems were eliminated because of their coding terminology. 3/

Table 1. Sample of Hospitals

Abstract service	Number of hospitals contacted	Number ineligible for inclusion	Number declining to participate	Number participating in the study
CHDC	10	2	2	6
HUP	8	-	1	7
QUEST	8	1	0	7
PAS	39	-	11	28
Internal	5	3	0	2
Total	70	6	14	50

2
Utilization data from 1974 will be used as a baseline for assessing the effects of PSROs over time. If abstracted data from this study are judged to be reliable, they could be supplemented by specially gathered information from hospitals not subscribing to these abstract services, which were included in the sampling frame but excluded from the study, in order to provide true national estimates of utilization patterns.

3
PSROs will develop data analysis competence in two disease classifications only: ICDA-8 (Eighth Revision, International Classification of Diseases, National Center for Health Statistics, Public Health Service, Publication No. 1693, Vol. 2, Deceember, 1968) and H-ICDA (Hospital Adaptation of the International Classification of Diseases, 2nd ed., Commission on Professional and Hospital Activities, Ann Arbor, Michigan, 1973). The three hospitals not included had been using SNDO (the Standard Nomenclature for Diseases and Operations) in 1974, even though all three had switched to either H-ICDA or ICDA-8 by mid-1975.

The hospitals which declined to participate in the study usually gave one of three reasons. Several had just completed either another major research project or an audit by the Joint Commission for the Accreditation of Hospitals (JCAH) and were reluctant to undertake another outside project. Others, particularly large urban hospitals, felt that they did not have the personnel needed to locate specified medical records, even though the study had been designed to minimize demands on medical record department staff. A third group expressed a vague distrust of research projects, the PSRO program, and reporting requirements in general and was also concerned about confidentiality of medical information. The high participation rate for QUEST hospitals was undoubtedly influenced by personal letters from a vice president of Blue Cross of Northeast Ohio to the administrators of each sampled hospital, encouraging their cooperation. It is possible that the large size of the PAS system may have instilled less sense of identification among subscriber hospitals, which then felt less compelled to participate. Furthermore, a few of the PAS hospitals were participants in the CPHA reliability study (see page 6) and may have viewed the IOM study as repetitive. In general, the non-participants did not appear to differ from participants in any systematic manner, with the possible exception of a lower rate of cooperation among proprietary hospitals. The hospital weights were adjusted to reflect the influence of the reduced sample size.

Sample of Discharge Abstracts

In selecting abstracts within each study hospital, special attention was given to the 14 diagnoses to be used in evaluating PSRO impact on utilization of hospital services, which were selected because they met most of the following criteria:

- high frequency of occurrence,
- high cost of care,
- diagnosed with a high degree of reliability,
- amenable to data gathering through the Uniform Hospital Discharge Data Set, and
- stable utilization pattern during recent time periods. 4/

4
U.S. Department of Health Education and Welfare, Office of the Assistant Secretary for Health, Office of Professional Standards Review, Program Evaluation Plan: Professional Standards Review Organizations, by Martin A. Baum et al. (22 September 1975), pp. 119-21.

If the sample had been limited to only those abstracts on which any of the 14 diagnoses were listed as principal 5/, it would be possible to estimate the frequency with which a given diagnosis was listed as principal and should not have been (a false positive). It would not be possible to estimate the frequency with which a given diagnosis should have been listed as principal and was not (a false negative). To satisfy completely the latter estimating requirements and calculate precise net and gross difference rates, it would be necessary to sample all diagnoses for a particular population--in this case, Medicare and Medicaid discharges during 1974. However, this approach would have exceeded available resources, if it were also to meet the requirement of concentrating adequately on the 14 evaluation diagnoses. Thus, it became necessary to balance the objectives of focusing on the 14 evaluation diagnoses, while at the same time achieving some representation from other conditions.

A compromise sampling plan evolved that included a higher sampling ratio for the 14 diagnoses included in the PSRO evaluation (referred to as "target" diagnoses) when they were coded as primary or principal. For 12 of the target diagnoses, some abstracts on which the 12 diagnoses were listed as "other" diagnoses were also sampled, but at a lower rate. 6/ This permitted an estimate of the extent to which confusion over designation of principal as opposed to other diagnosis led to errors. Finally, an added sampling element was developed to account for seven of the target diagnoses which are frequently and erroneously interchanged with specific different diagnoses (referred to as "satellite" diagnoses). 7/ As an example, a principal diagnosis of acute tonsillitis (a "satellite" diagnosis) is often coded erroneously, in place of hypertrophy of the tonsils/adenoids (a "target" diagnosis). The target and corresponding satellite diagnostic categories are listed below. Diagnostic code numbers from both the H-ICDA-2 and ICDA-8 classification schemes are found in Appendix D. For some diagnoses the codes are not directly comparable and this is so noted.

5
According to UHDDS, the principal diagnosis is "the condition established after study to be chiefly responsible for occasioning the admission of the patient to the hospital for care." Other diagnoses include "all conditions that coexist at the time of admission, or develop subsequently, which affect the treatment received and/or the length of stay. Diagnoses that relate to an earlier episode which have no bearing on this hospital stay are to be excluded." See: U.S. DHEW, Uniform Hospital Abstract: Minimum Basic Data Set, Vital and Health Statistics, Series 4: Documents and Committee Reports, No. 14 (August 1974).

6
Sampling frames were not established for hypertrophy of tonsils/adenoids or deliveries when they were coded as an "other" diagnosis, since this rarely occurs.

7
For the remaining target diagnoses, coding errors were not expected to stem primarily from confusion between a target diagnosis and a pre-determined different diagnosis. Therefore, "satellite" diagnoses were not defined.

Target Diagnoses	Satellite Diagnoses
Chronic ischemic heart disease	Subacute ischemic heart disease
Acute myocardial infarction	Subacute ischemic heart disease
Inguinal hernia without obstruction	Inguinal hernia with obstruction
Hypertrophy of tonsils and adenoids	Acute tonsillitis
Fracture of neck of femur	Fracture of other and unspecified parts of femur
Displacement of intervertebral disc	Low back pain
Neuroses	Personality disorders
Cerebrovascular disease	-
Diabetes mellitus with and without acidosis	-
Malignant neoplasm of the breast	-
End stage renal disease	-
Cholecystitis/cholelithiasis	-
Delivery with and without complications	-
Cataract	-

After completing agreements of participation with each of the study hospitals, the abstracting services and the two hospitals with internal data systems forwarded to the Institute of Medicine copies of discharge abstracts for patients discharged during 1974, for whom the anticipated principal source of payment on admission was either Medicare or Medicaid and who had any of the diagnoses of interest. Either a manual or computerized sampling process was used (depending on volume) to select a pre-specified number of abstracts per diagnostic category per hospital from each of the following categories of independent sampling frames:

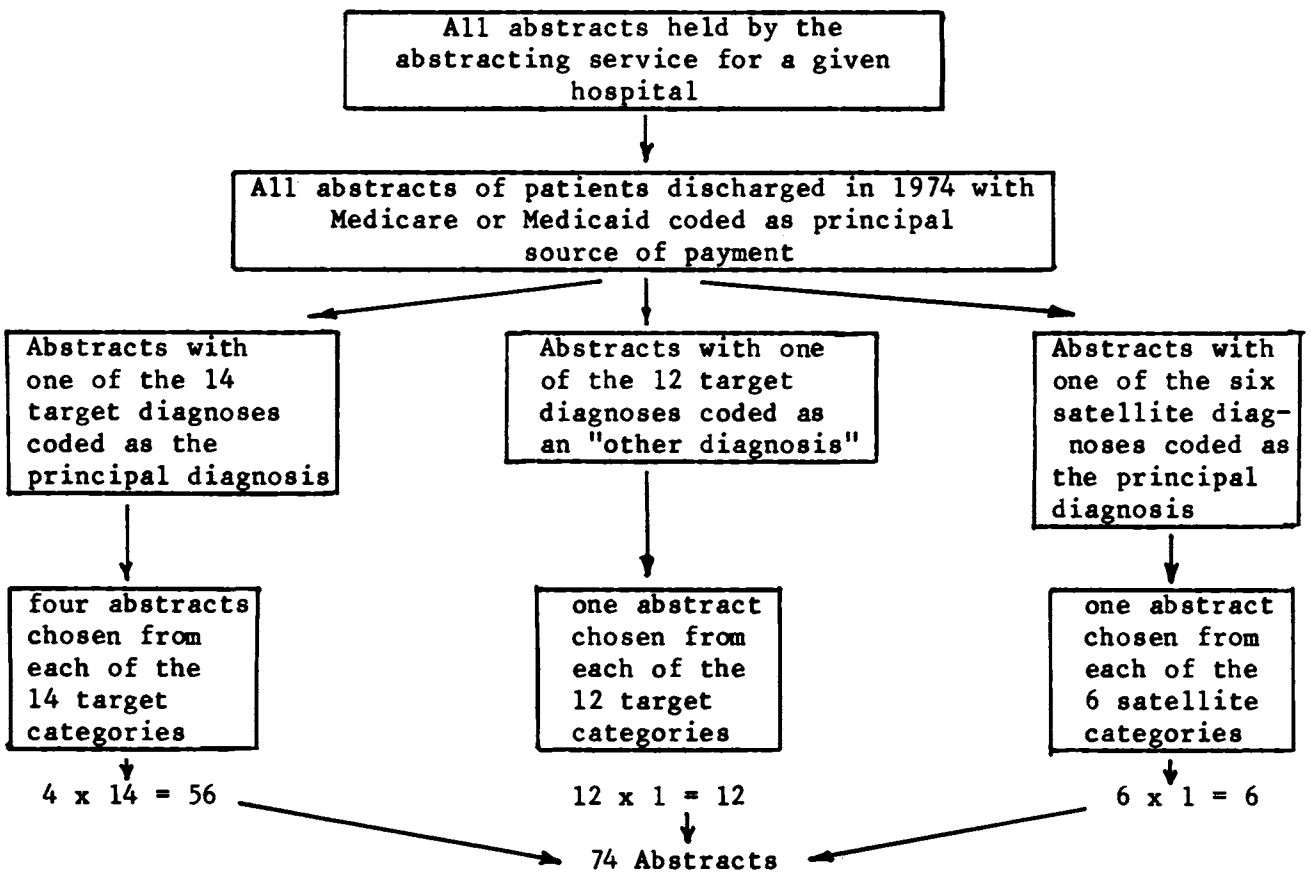
- "Target" diagnoses when they were coded as "Principal"--for each of the 14 target diagnoses, four abstracts were to be selected per hospital, using a systematic probability sampling technique, which would yield 56 abstracts per hospital in this category.

- "Target" diagnoses when they were coded as "Other" than principal--for each of 12 target diagnoses, one abstract was to be randomly selected per hospital to yield 12 abstracts per hospital in this category.

- "Satellite" diagnoses when they were coded as "Principal", regardless of whether or not a target diagnosis was coded as "Other" than principal--for each of the six satellite diagnoses, one abstract was to be randomly selected per hospital to yield six abstracts per hospital in this category.

This sampling plan was designed to yield a maximum of 74 abstracts per hospital or a total of 3,700 abstracts to be re-abstracted. However, some hospitals did not have enough discharges per diagnostic category to meet the sampling requirements. Furthermore, for some of the abstracts included in the sample, the hospital medical record department personnel were unable to retrieve the corresponding medical record. Thus, the number of records actually re-abstracted was reduced accordingly, as shown in Figure 1.

Figure 1. Sample of Abstracts



74 Abstracts x 50 hospitals = 3,700 Abstracts Expected

265 Abstracts not available in sampling frames
134 Records not available in record departments
399 Abstracts Excluded

3,700 Expected - 399 Excluded = 3,301 Abstracts Available for Re-abstracting

With this sampling plan and a weighted analysis, it is possible to generalize the results of the re-abstracting study nationally to all 1974 discharges for Medicare and Medicaid patients who had one or more of the diagnoses studied and were treated in hospitals subscribing to one of the four participating abstracting services or larger hospitals with more than 1,000 beds and internal data systems.

It is important to note that the purpose of the study is only to generate national estimates at one point in time and not to compare the relative reliability of data processed through the four participating abstracting services. Because the PSRO evaluation plan requires national estimates of patterns of care and no single system can generate national estimates, it was necessary to include several abstracting services. However, at no point did the study objectives focus on an assessment of the competence or quality of individual abstracting services. The data do not permit drawing such conclusions. Furthermore, all participating services have revised their procedures and edit systems since 1974, and the study findings can not be generalized to current data.

SURVEY INSTRUMENTS

Two survey instruments were developed: the re-abstracting form was intended to record information obtained in the process of re-abstracting selected medical records (see Appendix E); the medical record department questionnaire was intended to capture information describing the abstracting process within each study hospital (see Appendix F).

Re-abstracting Form

Information items to be re-abstracted included date of hospital admission, date of discharge, sex, date of birth or age (depending on the abstract service used by a particular hospital), principal expected source of payment, principal and other diagnoses, and principal and other procedures. Only principal diagnosis and principal procedure were coded, however. These seven items were considered to be the minimum information for the PSRO evaluation. Many other items are coded on a discharge abstract, such as information on laboratory procedures, attending physician, and patient disposition. But the complexities and costs of adding more items to the re-abstracting process restricted the analysis to those mentioned above.

All coding was in accord with item definitions developed for the Uniform Hospital Discharge Data Set (included in Appendix E), since PSROs will use UHDDS definitions.

Several steps were involved in the re-abstracting process, each intended to increase understanding of the reasons for inaccuracy in utilization statistics and the relative degree of error that would be involved if such data were used for evaluative purposes. For example, simply determining that the patient's age on the original abstract was incorrect does not enable one

to learn the seriousness of that error for evaluation. An error of one year may be relatively minor; consistent errors ranging from 10 to 15 years are probably unacceptable. For that reason, when the original abstract and the re-abstract disagreed on the coding of a given item, the information from the original abstract was transferred to the form (column 3) in order to facilitate further analysis. In addition, the form allows for recording the correct information if both the original abstract and re-abstract are judged to be inaccurate upon closer review of the medical record.

When discrepancies were found between the original abstract and the re-abstract, an attempt was made to determine the reasons for the differences, since such information might be used to improve the abstracting process in either the hospital or abstracting service. For information on admission date, discharge date, sex, date of birth or age, and anticipated principal source of payment, the following general reasons for discrepancy were provided:

- Clerical--discrepancy attributable to mistakes of a coding clerk, such as obvious transposition of numbers for patient age.

- Completeness--inaccurate information stemming from an inadequate review of the medical record. For example, an item may be missing from the admitting sheet, but clearly stated in the discharge summary.

- Procedural--discrepancy resulting from an identifiable practice within a hospital, which results in systematic differences between the original and the re-abstract. For example, a routinely and consistently different interpretation of data set items, such as inverting day and month in admission and discharge dates.

The reasons for discrepancies on principal diagnosis and procedure were divided into two broad categories: ordering and coding. However, the possibility of an ordering discrepancy had to be eliminated before considering the possibility of a coding discrepancy.

In general, ordering discrepancies stem from uncertainty about whether a diagnosis or procedure should be regarded as "principal" or "other." The following specific types of ordering discrepancies were considered:

- Ordering: Abstract Service Policy--these discrepancies reflect differences between abstract service and UHDDS definitions, which should occur consistently if hospital coders adhere to the abstract service definitions. For example, a patient is admitted for an open fracture reduction and, while on the operating table, suffers an acute myocardial infarction which keeps him in the hospital three months. The UHDDS definition requires that fracture be coded as the principal diagnosis, because the cause of admission was a fracture. However, if the hospital participates in QUEST, acute myocardial infarction probably would be coded as principal diagnosis, since QUEST's non-specific

interpretation of principal discharge diagnosis is assumed to reflect the most important diagnosis that best accounts for the days of care. 8/

- Ordering: Hospital Policy--these discrepancies reflect systematic differences between the definitions of a particular hospital for selecting principal diagnosis and/or procedure and the definitions of both the abstracting service and UHDDS definitions. As an example, in some hospitals the diagnosis listed first by the physician is routinely coded as principal.

- Ordering: Judgment--these discrepancies reflect a difference of opinion or professional judgment in interpreting the medical record and determining which diagnosis/procedure should be regarded as principal. This option is particularly important in instances in which there is no "correct" answer, but rather, the code on either the original abstract or re-abstract is equally acceptable and there are no clear-cut guidelines to determine which is preferable.

- Ordering: Other--this option was provided in case there were situations which did not fit any of the above. The field team was instructed to write a note explaining the necessity to use this reason.

After the ordering options were eliminated as possible reasons for explaining discrepancies between the original abstract and re-abstract, the following coding options could be considered:

- Coding: Procedural--these discrepancies stem from routine and systematic misuse or misunderstanding of the coding system. Examples include reliance on the index without reference to tabular listings, failure to heed inclusion or exclusion advice from the tabular listing, double coding when single coding is indicated, and single coding when double coding is required.

- Coding: Judgment--these discrepancies result from the absence of complete word-for-word agreement between the recorded diagnosis or procedure and the wording in the coding manual, which means that judgment must be exercised in determining what the code should be. For example, a diagnosis may be listed on the record as recurrent. It is unclear whether "acute" or "chronic" is actually the more appropriate qualifier for coding purposes, and these are the only two options available in the coding manual.

8
QUEST Manual (Cleveland: Blue Cross of Northeast Ohio, April, 1967), p. 16; Jerome E. Goodman, Manager, Health Data Systems, and John P. Sturges, Research Analyst, personal communication, 1 November 1975 and confirmed 12 January 1977.

- Coding: Other--this option was provided for coding discrepancies which could not be explained by the reasons listed above. Again, the field team was instructed to write a note to explain the necessity for its use. It was expected to be used most frequently to explain differences in the degree of detail and significance reflected in the coding of procedures.

- Coding: Dependent--this option applied only to discrepancies on principal procedure and was used if an earlier discrepancy in the selection of principal diagnosis resulted in a corresponding or dependent discrepancy in the selection of principal procedure.

For each discrepancy, the field team was instructed to code only one explanatory reason. In certain instances this required interpretation and subjective assessment in order to determine which reason might be influential. The reliability of these responses may be less than for the remainder of the data. Nevertheless, the potential value of the information was judged to outweigh its partially subjective nature.

Medical Record Department Questionnaire

Prior to the initiation of field work the supervisor of the medical record department in each study hospital was asked to complete a questionnaire on variables expected to be associated with varying levels of reliability of abstracted data. All responses related to calendar year 1974 in order to conform with the abstracted data. The questionnaire is included as Appendix F.

An initial set of questions asked about training levels of persons abstracting the medical records and, in particular, training of the supervisor. Additional questions reflected participation in training programs and contacts with the abstract service intended to improve the abstracting process. The extent of such activities was thought to vary by hospital and possibly be associated with the quality of abstracted data.

Another section of the questionnaire obtained information on the process by which abstracting occurred within the hospital. In particular, these items attempted to determine when information was abstracted, average work load, and the portions of the record used for abstracting, all of which may affect the completeness of information upon which judgments are made. A related question determined whether the record department had any method of checking the accuracy of completed abstracts. A section on uses of reports prepared by the abstract service was included because it has been hypothesized that only in hospitals where the data are used routinely will the data be reliable.

Finally, the section about the definition of information items (i.e., diagnoses, procedures, and lengths of stay) was included to determine whether the hospital followed a unique pattern which did not conform with the definitions recommended by the abstract service. If so, the hospital procedure could be internally correct, but still result in discrepancies if either UHDDS or abstract service definitions were used as the basis for comparison.

FIELD WORK

The field work was conducted by three Registered Record Administrators (RRAs), recruited because of their extensive experience in hospital medical record departments and related research activities. They were specially trained for this study by a medical record consultant under the general guidance of the Institute staff. Members of the field team were asked first to determine the best possible code for each item by fully exploring the medcal record and then to attempt to understand the reasons for discrepancies. They were explicitly instructed not to simply determine whether the original code was "correct" or "incorrect."

Before the field work began, the supervisor of the medical record department in each study hospital was asked to locate and have available the selected medical records, which were identified by medical record number and admission date. Therefore, when the member of the field team arrived, she first reviewed the completed questionnaire to acquaint herself with any unusual hospital procedures that would influence the abstracts and immediately began re-abstracting. No substitutions were made for records which were not available. To avoid conditioning the re-abstracting, the original abstracts were contained in a sealed envelope, to be opened after all records had been re-abstracted. The completed re-abstracts were then compared with the original abstracts provided by the abstract service and any differences between the two were noted. If a discrepancy between the abstract and re-abstract was observed, the original medical record was checked again to determine what information appeared to be correct, and what factors might account for the discrepancy. Details on the re-abstracting process are found in Appendix E.

To check the reliability of the field work, a subsample of abstracts was independently "re-re-abstracted" by a consultant who did not participate in the initial field work. Comparisons were made between these results and those compiled by the field team. In conducting the assessment, the consultant did not know which member of the field team had done the initial abstracting, nor did she know whether discrepancies were initially detected. This assessment process and its results are described in detail in Appendix G.

CONFIDENTIALITY

Because the study required working with information taken directly from individual patient records, specific measures were taken to assure confidentiality. The medical record number was used only to access the relevant record and to assist in compiling the abstracted information. In no instance was the name of the patient or attending physician recorded. The report contains statistical summaries only, which do not permit the identification of patient, physician, hospital, or abstracting service. Statements to this effect were included in the initial introductory letter to hospital administrators, as a legend on the re-abstracting form, and in a three-way contract with the hospital, the abstract service, and the Institute of Medicine that permitted the Institute to obtain the necessary information. The computer tapes on which

the data were compiled for analytic purposes were stored in the National Academy of Sciences Computer Center, which is accessible only to authorized individuals.

DATA PROCESSING

After completion of the field work, the re-abstracts and original abstracts were returned to the Institute for data processing. Each abstract and medical record department questionnaire was first scanned visually, keypunched and verified, and subjected to computer edits. After the accuracy of the raw data was assured, weights were added for use in the analysis.

A single composite weight was assigned to each abstract which reflected its probability of inclusion at each step in the sampling process. (See Appendix B for additional information.) The weights were applied throughout the analysis to permit generalizing to the broader universe of diagnoses, patients, and hospitals to which the findings relate; and the weighted percentages are reported. However, only the unweighted totals are included in the tables presented in the following chapter.

Chapter 4

ANALYSIS

In general, the analysis sought to assess the reliability of the seven abstracted information items chosen for study by determining how often there was a discrepancy between the original abstract and the re-abstract. A series of analyses was conducted including simple calculations of the frequencies of discrepancies on each information item and more extensive assessments of the 14 diagnostic categories. Factors which might influence the accuracy of diagnostic and procedural coding were also explored.

TOTAL FREQUENCIES OF DISCREPANCIES

Table 2 shows the weighted frequencies of discrepancies between the Institute of Medicine re-abstract and original abstract for each item in all study hospitals. Where discrepancies were detected, the correct data source is also noted. Except for diagnoses and procedures, the data are highly reliable. Where there were discrepancies, the Institute of Medicine re-abstract was more likely to be found correct. However, for 10.7 percent of the diagnoses and 16.3 percent of the procedures, even though there were discrepancies, the correct data source could not be determined. The accuracy of data on anticipated principal source of payment should not imply that this information is generally correct. The study was intended to include only Medicare and Medicaid discharges, and abstracts were sampled accordingly. However, there may be additional actual Medicare or Medicaid discharges which were erroneously abstracted as having other sources of payment, and therefore, were not eligible for inclusion in the sample.

The percentages of abstracts in Table 2 with no discrepancy and those for which the original abstract was determined to be correct are combined in Table 3, which indicates the percentage of original abstracts with correct data items, by abstract services individually and combined. The lower level of accuracy in Abstract Service A for date of birth or age is thought to stem from an error in a computer program used by the abstract service to relate date of birth and age, since the error was consistent and no other reason could be identified. The differences in accuracy of diagnostic and

Table 2. Discrepancy Between IOM Re-abstract and Original Abstract and the Correct Data Source for Selected Items (weighted percent)

Selected items	Percent with no discrepancy	Correct data source where a discrepancy exists				
		Abstract	Re-abstract	Neither	Indeterminate	Total
Admission date	99.7	0.2	0.1	-	-	100.0%
Discharge date	99.2	0.7	0.1	-	-	100.0
Date of birth/age	97.7	0.3	1.4	0.1	0.4	99.9
Sex	99.1	0.2	0.7	-	-	100.0
Payment source	98.1	0.7	0.7	-	0.5	100.0
Principal Diagnoses	65.2	1.6	22.2	0.2	10.7	99.9
Principal Procedures	73.2	2.5	7.8	0.2	16.3	100.0

Note: Unweighted N = 3,301 abstracts.

Table 3. Original Abstracts with Correct Data for Selected Items by Abstract Service (weighted percent)

Selected items	All services	Abstract service				
		A	B	C	D	E
Admission date	99.9	100.0	99.9	99.8	100.0	100.0%
Discharge date	99.9	100.0	100.0	99.9	100.0	98.8
Date of birth/age	98.0	68.0	98.9	99.7	99.8	91.4
Sex	99.3	100.0	99.7	99.6	93.2	99.4
Payment source	98.8	96.9	96.9	99.7	99.4	100.0
Principal Diagnoses	66.8	71.8	76.6	64.8	68.6	67.6
Principal Procedures	75.7	75.2	86.0	73.5	82.6	89.5

Note: Unweighted N = 3,301 abstracts.

procedural data among abstract services are statistically significant. The generally lower reliability of this information required further exploration since a variety of factors could have been influential, including:

- The potential inadequacy of current nomenclature, coding guidelines, or medical recording for definitively determining primary diagnosis and the necessity of relying on professional judgment, particularly if a patient has multiple diagnoses, which may lessen reliability;
- The degree of coding refinement (four-digit, three-digit, or broader diagnostic groupings);
- The contribution of individual diagnoses to the overall discrepancy rates; and
- The contribution of structural and functional factors within the hospital which may affect the reliability of abstracted information, such as training of abstractors, procedures and supervision within the medical record department, and the thoroughness of the record review before determining principal diagnosis and procedure.

REASONS FOR DISCREPANCIES

In an initial attempt to understand the lower reliability of information on principal diagnosis and procedure, the reasons assigned by the field team to explain the discrepancies were analyzed. As noted in Chapter 3, the possibility of an ordering discrepancy (a discrepancy caused by uncertainty over whether a diagnosis or procedure should be regarded as "principal" or "other") was to be considered before the possibility of errors in coding. Tables 4, 5, and 6 show the reasons for discrepancy, according to the correct data source, for diagnostic codes compared at the fourth and third digit and for procedure codes.

For diagnostic information, coding discrepancies occurred more frequently than ordering discrepancies when either the original abstract or Institute of Medicine re-abstract was correct and codes were compared to the fourth digit. For three-digit comparisons, however, ordering discrepancies were more likely to occur. Ordering decisions also were more important for both three and four-digit comparisons if both abstracts were incorrect or the correct data source could not be determined. For information on procedures, coding was more troublesome than ordering, regardless of the correct data source. Ordering decisions were not linked with policies of specific abstract services, with the possible exception of discrepancies related to designation of principal procedure. Policies of particular hospitals did lead to ordering errors. This occurred more frequently for diagnoses than for procedures. Anecdotal information provided by the field team indicates that in some hospitals the diagnosis listed first by the physician is coded as primary; in at least one hospital, coders assign a diagnosis on the basis of their knowledge of cases usually treated by the attending physician or the ward on which the patient stayed.

Table 4. Reason for Discrepancy in Principal Diagnostic Codes Compared to the Fourth Digit by Correct Data Source (weighted percent)

Reason for discrepancy	Correct data source			
	Abstract	Re-abstract	Neither	Indeterminate
Ordering-abstract service policy	-	-	-	-
Ordering-hospital policy	-	26.1	-	0.7
Ordering-judgment	36.0	15.3	58.8	72.4
Ordering-other	-	5.5	-	-
Coding-procedural	22.5	16.1	29.4	0.1
Coding-judgment	11.5	32.3	11.8	24.9
Coding-other	30.0	4.6	-	1.7
Total	100.0%	99.9%	100.0%	99.8%
(Percent of total number of abstracts*)	(1.5)	(21.8)	(0.2)	(10.5)

Note: Unweighted N = 953 abstracts with discrepancies.
*These percents are slightly less than comparable percents in Tables 2, 3, and 7. For a few abstracts on which there were discrepancies, the field team neglected to determine the correct source of data. Those abstracts were proportionately distributed among columns for correct data source in Tables 2, 3, and 7. This was not done here, and therefore, the percents are lower.

Perhaps the most important discrepancies from the standpoint of reliability are those for which the correct data source could not be determined (the "Indeterminate" columns in Tables 4-6). Four-digit diagnostic discrepancies in this category constitute 10.5 percent of the total number of abstracts in the study when all diagnoses are combined. The most frequent reason for this decision was "Ordering-Judgment," which was to be chosen when the discrepancy reflected a difference of professional opinion in interpreting the medical record. The next most frequent reason was "Coding-Judgment" with similar implications. When diagnoses are compared at three digits the percent of abstracts assigned to the "Indeterminate" category remains about the same (9.8 percent), and the reasons for discrepancy are similar to those for four-digit coding. This suggests that for any aggregate sample of abstracts a sizable portion will not be reliable in the sense of being coded similarly

on repeated occasions. Furthermore, there is no identifiable or correctable "error." Instead, with current nomenclature, coding guidelines, and recording practices, this must be accepted as background "noise," which probably is difficult to eliminate.

Examples of incompatibilities between the principal diagnosis on the original abstract and that assigned by the field team, when the correct data source could not be determined, are found in Appendix H for three study diagnoses. In some cases the diagnoses appear to be unrelated. The discrepancy apparently stems from co-morbidity and the difficulty of determining which diagnosis is principal. In other cases, however, the two diagnoses appear to relate to the same general constellation of diseases or body system. For such cases, the utility of the original data probably depends on the purpose to which they are applied.

Table 5. Reason for Discrepancy in Principal Diagnostic Codes Compared to the Third Digit by Correct Data Source (weighted percent)

Reason for discrepancy	Correct data source			
	Abstract	Re-abstract	Neither	Indeterminate
Ordering-abstract service policy	-	-	-	-
Ordering-hospital policy	-	39.0	-	0.8
Ordering-judgment	65.1	22.7	66.7	76.5
Ordering-other	-	8.3	-	-
Coding-procedural	22.2	14.1	33.3	-
Coding-judgment	4.8	12.4	-	21.3
Coding-other	7.9	3.5	-	1.4
Total	100.0%	100.0%	100.0%	100.0%
(Percent of total number of abstracts*)	(0.8)	(14.6)	(0.2)	(9.8)

Note: Unweighted N = 714 abstracts with discrepancies.
*These percents are slightly less than comparable percents in Tables 2, 3, and 7. For a few abstracts on which there were discrepancies, the field team neglected to determine the correct source of data. Those abstracts were proportionately distributed among columns for correct data source in Tables 2, 3, and 7. This was not done here, and therefore, the percents are lower.

Selection of the "Indeterminate" option for discrepancies on principal procedure (16.3 percent of the total number of abstracts in the study) was most frequently explained by the "Coding-Other" option. In almost every case, the field team decided that no procedure was important enough to warrant coding, even though a code appeared on the original abstract. The field team had been specifically instructed to use this option in such instances. The next most frequent reason is "Ordering-Judgment" which again does not imply an identifiable "error," but rather, a difference in judgment in selecting the principal procedure.

Table 6. Reason for Discrepancy for Principal Procedures by Correct Data Source (weighted percent)

Reason for discrepancy	Correct data source			
	Abstract	Re-abstract	Neither	Indeterminate
Ordering-abstract service policy	-	11.7	-	-
Ordering-hospital policy	-	5.0	-	-
Ordering-judgment	1.1	12.2	-	34.7
Ordering-other	13.5	1.9	-	-
Coding-procedural	24.5	20.1	39.1	0.4
Coding-judgment	6.7	7.0	-	7.7
Coding-other	53.1	41.6	60.9	56.9
Dependent	1.1	0.5	-	0.2
Total	100.0%	100.0%	100.0%	99.9%
(Percent of total number of abstracts)	(2.5)	(7.8)	(0.2)	(16.3)

Note: Unweighted N = 901 abstracts with discrepancies.

For four-digit diagnostic discrepancies where the original abstract was correct (1.5 percent of the total) the most likely reason for discrepancy was "Ordering-Judgment," followed by "Coding-Other." For three-digit diagnostic discrepancies, the "Ordering-Judgment" option remained important, but coding problems were "procedural," rather than "other." Non-specific reasons related to coding were the most important justification for discrepancies on procedures where the original abstract was correct.

For those abstracts where the re-abstract was the correct data source for four-digit diagnostic coding (21.8 percent of the total number of abstracts), the most frequent reason for discrepancy reflected judgmental decisions regarding coding, followed by ordering policies of particular hospitals. As noted earlier, such policies usually required that the diagnosis listed first on the face sheet of the medical record be coded as principal. When diagnoses were compared to the third digit, the "Coding-Judgment" option was used less often and ordering decisions linked with hospital policies were more important. Discrepancies on principal procedure where the re-abstract was correct (7.8 percent of the total) were primarily related to coding problems, rather than difficulty in determining which procedure should be primary. The most frequent coding problems were non-specific "Other" reasons or procedural reasons, caused by routine and systematic misuse or misunderstanding of the coding system.

Discrepancies where both the original and Institute of Medicine abstracts were incorrect occurred so infrequently that the associated reasons will not be discussed.

INFLUENCE OF DIAGNOSTIC GROUPINGS

The initial tables in this chapter showed the percent of abstracts with no discrepancies for all principal diagnoses combined and compared to the fourth digit. Depending on the reasons for analysis, less specific coding may be equally appropriate and result in increased reliability. In Tables 4 and 5, the frequency of discrepancies stemming from differences in coding judgment was reduced when codes were compared to three digits rather than four. In this section, the influence of differing levels of diagnostic groupings is explored in more detail.

For most of the 14 target diagnoses, three-digit analyses may be acceptable for determining basic utilization patterns such as admission rates, with a few exceptions. Three-digit coding is probably inappropriate for end stage renal disease and low back pain, because the conditions encompassed in the first three digits would include some diagnoses specifically excluded from the sampling frame. For other diagnoses, the fourth digit may be important to indicate the presence of complications that may influence length of stay. Specially derived classifications that group homogeneous patients on clinically meaningful diagnostic criteria also may be useful in displaying patterns of care. The CPHA List A is one such classification scheme which fluctuates between three- and four-digit coding in specificity.

Table 7 shows the influence of diagnostic groupings on reliability of diagnostic coding. Four-digit and three-digit categories are further subdivided to include all principal diagnoses combined, abstracts sampled because a target diagnosis was listed as principal, abstracts sampled because a satellite diagnosis was listed as principal, 1/ and abstracts with other

---

1
Satellite diagnoses are thought to be frequently and erroneously coded as primary in place of the correct target diagnoses. They were included in the sample to determine the extent to which this type of error may occur. See pp. 12-14.

diagnoses listed as principal which were sampled because a target diagnosis was coded as "other than principal."

Table 7. Discrepancy between the Original Abstract and IOM Re-abstract at Differing Levels of Coding Refinement and the Correct Data Source where a Discrepancy Exists for Principal Diagnosis (weighted percent)

Level of coding refinement	No discrepancy	Correct data source where a discrepancy exists				
		Abstract	Re-abstract	Neither	Indeterminate	Total
	(A)	(B)	(C)	(D)	(E)	
Four digit						
All diagnoses	65.2	1.6	22.2	0.2	10.7	99.9%
Targets	72.8	1.6	19.7	0.3	5.5	99.9
Satellites	67.8	3.4	21.6	-	7.2	100.0
Others	55.9	1.4	25.4	0.1	17.2	100.0
Three digit						
*All diagnoses	74.0	0.9	14.8	0.2	10.0	99.9
*Targets	80.9	0.9	13.7	0.3	4.2	100.0
*Satellites	70.8	3.9	17.2	-	8.1	100.0
Others	66.1	0.7	16.0	0.1	17.1	100.0
CPHA List A	70.2	1.1	18.7	0.2	9.8	100.0

Note: Unweighted N = 3,301 abstracts.
*Excludes end stage renal disease and low back pain.

According to Table 7, codes compared to three digits are consistently more reliable than four-digit comparisons. The reliability of coding with List A groupings lies midway between the three and four-digit levels, as would be expected on the basis of List A specificity. Within both three and four-digit categories, the target diagnoses combined have a higher reliability than all diagnoses combined and a lower percent of abstracts for which the correct data source could not be determined. This may result in part from the selection criteria, which required that the conditions chosen for the PSRO evaluation plan be diagnosed with a high degree of accuracy, as well as the higher sampling ratios for these conditions.

Abstracts included in the sample because a satellite diagnosis was listed as principal were analyzed further, although the data are not presented here. Where there was a discrepancy between the original and IOM abstracts, in 45.7 percent of the cases the field team determined that the target diagnosis corresponding to the satellite diagnosis should have been coded as principal instead. A similar, but less consistent, pattern existed for the abstracts sampled because a target diagnosis was coded as "other." The satellite and other diagnoses were selected with a lower sampling ratio and the reliability estimates may be less stable because of the smaller sample size and weighting

factors. In any case, the data base becomes more reliable when all diagnoses other than target diagnoses are excluded.

If the cases in Table 7 for which the correct source of data for principal diagnosis could not be determined (the "Indeterminate" column) are redistributed to columns where either the original abstract or the re-abstract was correct, and these percents in turn are added to the percent of abstracts with no discrepancy, it becomes possible to calculate ranges of "correct" data for either the original or Institute of Medicine abstracts. This is shown in Table 8. The lower figures in both columns show the percent of abstracts for which there were no discrepancies plus those for which either the abstract or re-abstract was determined to be correct. The upper ranges do not reflect "correct" data in the same sense. Instead, the upper figures include the lower percentage plus all abstracts in the "Indeterminate" column, on the assumption that since the correct data source could not be determined, either the original or re-abstract is equally likely to be correct. In reality, this is highly unlikely. The higher percents do not indicate reliable data, if reliability is defined as consistency of repeated assessments. Therefore, the "best estimates" of reliable data were calculated by determining for what percent of all abstracts in columns B, C, and D in Table 7 either the original abstract or re-abstract was determined to be correct, applying those percents to column E, and adding those figures to the respective lower ranges. They are probably better estimates of maximum levels of accuracy for principal diagnosis on either the original or re-abstract than the upper figures in the ranges would be.

Table 8. Re-Distribution of Indeterminate Discrepancies to Either Original Abstract or IOM Re-abstract and Resulting Ranges of "Correct" Data for Principal Diagnosis

Level of coding refinement	Original abstract*	Best estimate	IOM re-abstract**	Best estimate
Four digit				
All diagnoses	66.8-77.5%	67.5%	87.4-98.1%	97.2%
Targets	74.4-79.9	74.8	92.5-98.0	97.5
Satellites	71.2-78.4	72.2	89.4-96.6	95.6
Others	57.3-74.5	58.2	81.3-98.5	97.5
Three digit				
All diagnoses	74.9-84.9	75.5	88.8-98.8	98.1
Targets	81.8-86.0	82.1	94.6-98.8	98.5
Satellites	74.7-82.8	76.2	88.0-96.1	94.6
Others	66.8-83.9	67.5	82.1-99.2	98.4
CPHA List A	71.3-81.1	71.8	88.9-98.7	98.1

*The lower range equals the sum of the percentages of abstracts in columns A and B in Table 7; the upper range equals the sum of the percentages of abstracts in columns A, B, and E, in Table 7.

**The lower range equals the sum of the percentages of abstracts in columns A and C in Table 7; the upper range equals the sum of the percentages of abstracts in columns A, C, and E, in Table 7.

DIAGNOSTIC SPECIFIC DISCREPANCIES

In this section, individual principal diagnoses are examined to consider further the extent to which inaccuracies in the coding of specific diagnoses may affect the reliability of aggregated data. Table 9 shows the weighted frequency of discrepancy between the original abstract and

Table 9. Weighted Frequency of Discrepancy Between the Original and IOM Abstracts and the Correct Data Source where a Discrepancy Exists

Principal diagnosis	Percent of all abstracts that each diagnosis represents	Percent with no discrepancy	Correct data source where a discrepancy exists: Abstract	Re-abstract	Neither	Indeterminate	Total
Cerebrovascular disease	10.6	78.4	1.2	14.9	-	5.5	100.0%
Chronic ischemic heart disease	9.1	30.2	3.9	52.7	1.7	11.5	100.0
Acute myocardial infarction	5.5	68.2	2.9	24.1	0.5	4.3	100.0
*Subacute ischemic heart disease	1.3	63.7	5.3	16.3	-	14.7	100.0
Diabetes	3.4	60.8	0.6	29.1	-	9.4	99.9
Carcinoma of the breast	1.5	75.7	-	24.3	-	-	100.0
End stage renal disease	1.0	64.9	2.1	25.4	-	7.6	100.0
Cholelithiasis/ cholecystitis	3.9	73.2	0.4	16.3	0.2	9.9	100.0
Hernia without obstruction	2.5	89.1	-	9.3	-	1.6	100.0
*Hernia with obstruction	0.3	81.0	2.4	16.6	-	-	100.0
Delivery	4.2	86.5	1.4	4.3	-	7.8	100.0
Hypertrophy of tonsils and adenoids	1.3	97.0	-	2.5	-	0.5	100.0
*Acute tonsillitis	0.1	71.4	4.8	19.0	-	4.8	100.0
Fracture of neck of femur	3.7	81.1	3.7	14.9	-	0.3	100.0
*Unspecified fracture	0.5	68.7	-	31.3	-	-	100.0
Displacement of intervertebral disc	1.1	77.8	3.2	14.0	-	5.0	100.0
*Low back pain	0.3	56.3	-	25.0	-	18.7	100.0
Cataract	7.5	94.3	-	4.4	-	1.3	100.0
Neuroses	1.5	80.1	-	14.6	-	5.3	100.0
*Personality disorders	0.1	100.0	-	-	-	-	100.0
**Other	40.5	56.9	1.4	24.4	0.1	17.2	100.0
Total	99.9%	65.2	1.6	22.2	0.2	10.7	99.9

*Satellite diagnoses are listed beneath the target diagnoses (see pages 12-14 and Appendix D). **This category includes all primary diagnoses on abstracts which were included in the sample because a target diagnosis was coded as "Other than primary."

Institute of Medicine re-abstract for target and satellite diagnoses when they are coded as principal on the original abstract. The principal diagnoses on abstracts which were selected because a target diagnosis was coded as other than principal are included in the "other" category. Where discrepancies were found, the correct data source is also shown. The percent of cases with no discrepancy ranges from a low of 30.2 percent for chronic ischemic heart disease to 100 percent for the satellite category of personality disorders. Similar variation is shown in the percent of abstracts for which the correct data source could not be determined. The highest percentages are for chronic and subacute ischemic heart disease, low back pain, and "other" diagnoses. 2/

Abstracts for Medicare patients tended to be less reliable than for Medicaid patients, particularly for chronic and subacute ischemic heart disease, end stage renal disease, fracture of other and unspecified parts of the femur, and low back pain. However, abstracts for Medicare patients with breast cancer or inguinal hernia without mention of obstruction were more reliable than for Medicaid patients. Medicare patients are usually older and tend to have

---

2
It might be noted that the distribution of diagnoses within the total population of abstracts in Table 9 differs slightly from other readily available data sources. Comparisons must be viewed with caution because of differences in populations, diagnostic groupings, and methods for determining principal diagnosis. In particular, this study combined data for both Medicare and Medicaid patients and exactly comparable data are not available. Nevertheless, National Center for Health Statistics data for persons 65 years of age and older, which approximates the Medicare population, indicate that the percents of all discharges for patients with cerebrovascular disease, chronic ischemic heart disease, acute myocardial infarction, and diabetes mellitus (based on the first diagnosis listed on the face sheet of the medical record) are 6.2, 9.6, 2.7, and 2.7, respectively. [DHEW, Health Resources Administration, National Center for Health Statistics, Utilization of Short-Stay Hospitals: Annual Summary for U.S., 1974, Vital and Health Statistics, Series 13, No. 26, DHEW Pub. No. (HRA) 76-1777, (Washington, D.C.: U.S. Government Printing Office, September 1976), pp. 42-43.] This contrasts with 11.5, 9.8, 6.1, and 3.3 percent for comparable principal diagnoses in the Medicare portion of the Institute's sample. Data on Medicaid discharges were obtained from the states of Texas and Michigan. [Morris Williamson, State Department of Public Welfare, Austin, Texas, personal letter, 21 October 1976; Hermann A. Ziel, Chief, Bureau of Health Care Administration, Michigan Department of Public Health, personal letter, 20 October 1976; Paul N. Shaheen, Chief, Program Development and Evaluation, Michigan Department of Public Health, personal letter, 12 November 1976.] Again there is some variation, and the Medicaid data may not be complete. But the rank ordering of study diagnoses in the state data and the Institute's sample is comparable. Some distortion in frequencies may stem from sampling error and the weighting factors used in this study, but the extent to which this may have occurred cannot be determined precisely because of the absence of directly comparable data. In particular, the information was gathered in different ways, using different definitions and different methods for determining principal diagnosis. However, the distribution of diagnoses should not significantly affect overall considerations of reliability.

multiple diagnoses; this may explain the generally lower reliability of Medicare data.

The reasons for discrepancy were re-examined in order to determine whether special problems were associated with particular diagnoses (see Table 10). Because of the small number of abstracts, the reasons were grouped into those associated with determining principal diagnosis (ordering) and those associated with assigning a code number after the principal diagnosis had been selected (coding). If the correct data source could not be determined, the reasons for discrepancy tended to reflect problems in ordering. Where the Institute of Medicine re-abstract was determined to be correct, most diagnoses with high discrepancy rates and a likelihood of co-morbidity (especially chronic ischemic heart disease and diabetes) were associated with ordering problems. Coding discrepancies may reflect imprecision in coding guidelines. For example, with breast cancer there appears to be some uncertainty about whether attention should be directed to the site of the malignancy or the characteristics of the malignancy. Additional coding discrepancies may be limited to the fourth digit, as discussed below.

Table 11 shows the percent of abstracts with no discrepancy by principal diagnosis at differing levels of coding refinement. For all diagnoses, the percent increases as one moves from four-digit to three-digit comparisons. The larger increases tend to occur for diagnoses which had a high percentage of coding discrepancies in Table 10. As would be expected, the percent with no discrepancy on List A fluctuates, but is usually between the other two levels. The influence of the fourth digit is perhaps best illustrated by acute myocardial infarction. The H-ICDA codes for this diagnosis include ten options, ranging from 410.0 to 410.9. The increase in reliability which occurs when codes are compared only to the third digit reflects the likelihood of mistakes at the fourth digit. However, the loss in specificity of information obtained from three-digit analyses should also be noted. The code for hypertrophy of tonsils and adenoids is the same in all three comparisons, and therefore, the percent with no discrepancy is the same.

One might hypothesize that differing levels of reliability may be linked with use of particular coding schemes. H-ICDA includes more specific code numbers than ICDA-8, which might be associated with less reliability, since additional opportunities for error would arise. Alternatively, H-ICDA might be associated with greater reliability, since the increased precision from more specific codes might lessen the need to rely on judgment. Both three- and four-digit comparisons between H-ICDA and ICDA-8 were made. Some statistically significant differences were noted, but they were not sufficiently consistent or large to distort the conclusions drawn from the combined data. The differences may suggest that sets of information for individual diagnoses using different coding systems may not be directly comparable because of differing levels of accuracy. On the other hand, this may be a function of the hospital or the training of persons using a particular system and, in effect, be an artifact of the structural characteristics, discussed later.

Table 10. Reasons for Discrepancy by Principal Diagnosis and Correct Source of Data, Based on Four-Digit Comparisons (weighted percent)

Principal diagnosis	Correct data source: IOM re-abstract			Correct data source: Indeterminate		
	Ordering	Coding	Unweighted number of abstracts	Ordering	Coding	Unweighted number of abstracts
Cerebro-vascular disease	39.7%	60.3%	44	68.6%	31.4%	12
Chronic ischemic heart disease	82.3	17.8	126	93.0	7.0	30
*Acute myocardial infarction	12.1	87.4	49	42.9	57.1	8
Subacute ischemic heart disease	19.4	80.7	8	53.6	46.4	5
Diabetes	71.0	29.1	65	70.8	29.2	15
Carcinoma of the breast	24.5	75.4	48	-	-	-
End stage renal disease	29.7	70.2	44	45.5	54.6	14
*Cholelithiasis/ Cholecystitis	51.7	44.9	42	79.6	20.4	21
Hernia without obstruction	17.7	82.3	21	100.0	-	2
Hernia with obstruction	-	100.0	7	-	-	-
Delivery	-	99.9	13	12.2	87.8	9
Hypertrophy of tonsils and adenoids	60.0	40.0	6	100.0	-	2
Acute tonsillitis	25.0	75.0	5	100.0	-	1
Fracture of neck of femur	6.6	93.4	38	100.0	-	1
Unspecified fracture	-	100.0	11	-	-	-
Displacement of intervertebral disc	9.1	90.0	24	75.0	25.0	9
Low back pain	33.3	66.6	12	100.0	-	2
Cataract	-	100.0	14	33.3	66.7	2
Neuroses	51.5	48.4	18	66.7	33.3	8
Personality disorders	-	-	3	-	-	-

*The total number of cases where the re-abstract is correct is slightly less than 100 percent because for a few abstracts the field team neglected to determine a reason for discrepancy.

Table 11. Abstracts With No Discrepancy Between IOM Re-abstract and Original Abstract by Differing Levels of Diagnostic Coding Refinement (weighted percent)

Principal diagnosis	Four-digit comparison	Three-digit comparison	CPHA list A
Cerebrovascular disease	78.4%	84.4%	84.0%
Chronic ischemic heart disease	30.2	33.6	33.6
Acute myocardial infarction	68.2	93.9	93.9
Diabetes	60.8	68.2	61.9
Carcinoma of the breast	75.7	79.0	79.0
*End stage renal disease	64.9	-	66.4
Cholelithiasis/ cholecystitis	73.2	82.6	93.8
Hernia without obstruction	89.1	93.1	92.7
Delivery	86.5	89.8	89.4
Hypertrophy of tonsils and adenoids	97.0	97.0	97.0
Fracture of neck of femur	81.1	97.6	97.6
Displacement of inter-vertebral disc	77.8	85.5	70.9
Cataract	94.3	99.4	94.2
Neuroses	80.1	81.8	80.5

*Three-digit analysis is inappropriate for this diagnosis.

The tables presented to this point are based on cases for which a target diagnosis was listed on the original abstract as primary and the field team either agreed or disagreed with that determination. If there was disagreement, the original diagnosis may be regarded as a false positive. However, there may also be cases for which the same target diagnosis should have been listed as primary, but was not. These cases may be regarded as false negatives. The sampling plan permits a rough approximation of the extent to which both types of errors may occur. Perhaps more importantly, their influence on admission rates and lengths of stay can be explored. Table 12 helps to explain the methods for calculating these estimates.

In Table 12, the cases included in cell "a" are those for which the same target diagnosis was coded as principal on both the original abstract and re-abstract. The total number of differences affecting that figure for any specific diagnosis is equal to the number of cases included in that class on the original abstract, but not on the re-abstract (cell "c"), plus the number

included in that class on the re-abstract, but not on the original abstract (cell "b"). Cell "d" includes all cases from the study population which do not have the target diagnosis coded as primary on either the original or re-abstract.

The sum of the number of cases in cells "b" and "c", divided by the total number of cases in the population irrespective of diagnosis (N), may be termed the gross difference rate for the diagnosis in question. It reflects aggregate errors and usually includes differences in both directions, which may be partly off-setting. The net difference rate is the difference between "b" and "c", divided by N. It is an estimate of the non-offsetting part of the gross error. A negative net difference rate indicates that the influence of false positives is greater than false negatives. 3/

Table 12. Calculation of Net and Gross Difference Rates in Designation of Principal Diagnosis

IOM re-abstracts coded as principal	Original abstracts coded as principal		
	Specified target diagnosis	Other	Total
Specified target diagnosis	a	b	a + b
Other	c	d	c + d
Total	a + c	b + d	N

$$\text{Percent with no discrepancy} = \frac{a}{a + c} \times 100$$

$$\text{Gross difference rate} = \frac{b + c}{N}$$

$$\text{Net difference rate} = \frac{b - c}{N}$$

Net and gross difference rates are useful in comparing the relative accuracy of different diagnoses and for measuring changes in the reliability of data over time. In interpreting them, however, the reader should note that a change in the frequency of occurence of a particular diagnosis in a population is not necessarily reflected in net and gross difference rates. The number of cases for which both assessments agree (cell "a") may change

3
U.S. Department of Commerce, Bureau of the Census, The Current Population Survey Reinterview Program: Some Notes and Discussion, Technical Paper No. 6 Washington, D. C.: U.S. Government Printing Office, 1963), pp. 8-9.

without altering net and gross difference rates. The implications for reliability of similar net and gross difference rates for diagnoses with dissimilar incidence rates may be quite different. Therefore, the proportion of cases for which there is concordance between the abstract and re-abstract must be taken into account.

The net and gross difference rates for target diagnoses in Appendix I are illustrative of the variation in reliability that may exist. They should be interpreted with caution, however, because of a downward bias in the estimates for cell "b", the false negatives. The rates presented in this report were calculated on the assumption that there were no false negatives among the non-sampled categories of discharges; however, this may be a false assumption. As an example with weighted numbers, the table total (cell "N") represents 3,569,620 Medicare and Medicaid discharges in hospitals eligible for inclusion in the study. An estimated 130,668 cases would have been classified on the original abstracts as cerebrovascular disease (cells "a" plus "c"). An estimated 45,280 additional cases would have been classified as cerebrovascular disease on the re-abstracts (cell "b"), although they had been classified in some other category on the original abstracts. Cell "d" contains an estimated 3,393,672 cases which would not have been classified as cerebrovascular disease on the original abstracts and presumably would not have been so classified on the re-abstracts. Of these cases, however, only 1,312,332 were actually sampled. The remainder had no chance for inclusion in the study because they did not fall within the sampling frames delineated in Chapter 3. Even though the sample was designed to detect as large a proportion of the false negatives as possible, the remainder of cell "d" may include some additional cases of cerebrovascular disease. This would mean that the figure in cell "b" actually under-estimates the number of false negatives. Furthermore, the highly variable weights result in instability of the estimates. Unfortunately, the extent to which these rates are biased cannot be determined with existing data sources. These caveats apply to Tables 13 and 14, discussed below, as well as to Appendix I.

If the concepts of false negatives and false positives are used in calculating admission rates and lengths of stay, the operational implications of net and gross difference rates are easier to understand. Table 13 contains estimates of the distributions of principal diagnoses. Because of the absence of a population-based denominator customarily used to calculate admission rates, a proxy measure was computed based on the number of abstracts for Medicare and Medicaid patients with a particular diagnosis divided by the total number of Medicare and Medicaid admissions in the universe of hospitals eligible for inclusion in the study. This is referred to as a "rate," although it is not, in the usual sense. The basic admission rates are based on the number of cases for which both the original and Institute of Medicine abstracts have the same principal diagnostic code (cell "a") divided by the total number of admissions. The abstract service admission rates are calculated by dividing the total number of original abstracts with a target diagnosis (including false positives) by the total number of admissions. The re-abstract admission rates are calculated by dividing the total number of re-abstracts with a target diagnosis (including false negatives) by the total number of admissions. The rates are analyzed to three and four digits. However the re-abstract rates are the same for both four and three digits because the cases for which the original and

Table 13. Influence of False Positives and Negatives on Proxy Diagnostic Specific Rates of Admission (times 1,000) Based on All Medicare and Medicaid Admissions in Study Population

Principal diagnosis	Four-digit basic admission rate a/N	Three-digit basic admission rate a/N	Four-digit abstract service admission rate $\frac{a+c}{N}$	Three-digit abstract service admission rate $\frac{a+c}{N}$	Re-abstract admission rate $\frac{a+b}{N}$
Cerebrovascular disease	34.66	37.30	36.60	39.24	47.35
Chronic ischemic heart disease	11.50	12.80	36.81	38.12	14.43
Acute myocardial infarction	15.63	21.49	17.00	22.88	25.60
Diabetes	8.67	9.70	13.24	14.28	14.16
Carcinoma of the breast	4.62	4.83	5.90	6.14	4.90
End stage renal disease	2.76	*	3.50	*	3.94
Cholelithiasis/ cholecystitis	11.84	13.53	12.80	14.49	15.90
Hernia without obstruction	9.17	9.57	9.92	10.32	9.77
Delivery	15.30	15.87	15.36	15.93	17.63
Hypertrophy of tonsils and adenoids	5.22	5.22	5.38	5.38	5.30
Fracture of neck of femur	12.77	15.36	13.10	15.70	16.56
Displacement of intervertebral disc	3.42	3.80	4.04	4.42	5.05
Cataract	29.61	31.19	29.80	31.39	31.60
Neuroses	5.07	5.16	6.02	6.10	6.96

*Three-digit comparison is not appropriate for end stage renal disease.

re-abstracts disagreed at only the fourth digit are shifted from cell "b" to cell "a" in the three-digit comparisons. The numerator (a + b) remains the same and, therefore, the rate does not change.

As one would expect, the basic rates increase as one moves from four to three digits. The abstract service admission rates are consistently higher than the basic admission rates for both three and four-digit comparisons, because they include the false positives. If the number of false positives is roughly equivalent to the number of false negatives, then the abstract service rates may be an acceptable approximation to the "actual" rates. However, the re-abstract admission rates, which include the false negatives, are frequently higher than the abstract service rates. If the re-abstract admission rates are assumed to be correct, then reliance on abstract service data at the three-digit level will under-estimate the number of admissions for cerebrovascular disease, acute myocardial infarction, cholelithiasis/cholecystitis, delivery, fracture of neck of femur, displacement of disc, cataract, and neuroses. For the remaining target diagnoses, the abstract service data will over-estimate the number of admissions--most notably for chronic ischemic heart disease. The discrepancies increase with four-digit comparisons.

The under-representation of certain diagnoses within the abstract service data can be partially explained with further analysis. When an abstract was included in the sample because a satellite diagnosis was listed as principal and the field team disagreed with that determination, in 45.7 percent of those cases the field team decided that the target diagnosis corresponding to the satellite diagnosis should have been coded as principal instead. This percent remained fairly constant for all six satellite diagnoses. When an abstract was sampled because a target diagnosis was listed as "other" on the original abstract and the field team disagreed with the determination of principal diagnosis, in 20.4 percent of those cases the field team decided that the target diagnosis originally coded as "other" should have been coded as principal instead. But the percent varied by diagnosis. The corresponding percents for individual diagnoses were: cerebrovascular disease, 50.1 percent; acute myocardial infarction, 81.3 percent; cholelithiasis/cholecystitis, 48.0 percent; fracture of neck of femur, 69.0 percent; intervertebral disc, 56.6 percent; and cataract, 88.4 percent. Thus, much of the under-representation of diagnoses in the abstract service data may stem from predictable false negative diagnoses, which the sampling plan was specifically designed to address.

The influence of false positives and false negatives on length of stay may also be examined if the number of days is divided by the number of abstracts in the appropriate groupings of cells, as shown in Table 14. Four-digit lengths of stay are not consistently different from three-digit, except possibly for basic lengths of stay, where some analyses requiring agreement to four digits are associated with slightly increased lengths of stay. Lengths of stay based on abstract service data (including false positives) are about equally likely to be higher or lower than the corresponding basic numbers for both three- and four-digit comparisons. The re-abstract lengths of stay (including false negatives) are usually lower than both the basic and abstract service lengths of stay for three- and four-digit comparisons. If the re-abstract data are assumed to be correct, lengths of stay based on abstract service data may overestimate the average stay.

Table 14. Influences of False Positives and Negatives on Diagnostic Specific Lengths of Stay (in Days) for All Medicare and Medicaid Admissions in Study Population

Principal diagnosis	Four-digit basic length of stay	Three-digit basic length of stay	Four-digit abstract service length of stay	Three-digit abstract service length of stay	Re-abstract length of stay
Cerebrovascular disease	15.2	14.8	15.4	15.0	15.0
Chronic ischemic heart disease	12.6	12.4	12.2	13.1	11.8
Acute myocardial infarction	19.8	18.5	19.2	18.2	17.9
Diabetes	12.4	12.9	13.0	13.3	11.3
Carcinoma of the breast	13.5	13.4	14.1	14.0	13.4
End stage renal disease	8.9	*	15.3	*	10.3
Cholelithiasis/ cholecystitis	13.9	13.5	14.1	13.6	13.5
Hernia without obstruction	7.9	7.8	8.2	8.2	7.9
Delivery	5.5	5.5	5.5	5.5	5.4
Hypertrophy of tonsils and adenoids	2.8	2.8	2.8	2.8	2.8
Fracture of neck of femur	23.6	23.5	23.5	23.5	23.3
Displacement of intervertebral disc	16.1	15.4	15.5	15.1	15.2
Cataract	6.2	6.3	6.2	6.3	6.3
Neuroses	14.2	14.2	13.3	13.3	13.2

*Three-digit comparison is not appropriate for end stage renal disease.

INFLUENCE OF HOSPITAL CHARACTERISTICS

Selected aspects of the abstracting process within the hospitals and interactions between hospitals and abstract services, as well as basic hospital characteristics, were examined to determine their relationship to the reliability of data for diagnoses and procedures. Each hospital or abstract process characteristic was cross-tabulated by the percentage of abstracts for which the abstract service provided correct diagnostic data compared to both third and fourth digit and correct data on procedures. "Correct" data include all abstracts for which there were no discrepancies between the original and Institute of Medicine abstracts, plus those where there was a discrepancy but the original abstract was correct. "Incorrect" data include all abstracts for which there were discrepancies and the correct data source was either the Institute of Medicine re-abstract, neither, or indeterminate. 4/ A chi-square test of significance was calculated to determine the independence of the two variables. 5/

The influence of structural and functional characteristics on the reliability of data is difficult to interpret because, for many characteristics, a statistically significant relationship was found for diagnoses compared to three or four digits or for procedures, but the influence was not consistent for all three categories. In other instances, statistically significant relationships exist, but the response alternatives for the structural characteristics do not fall into a meaningful pattern. Because the number of cross-tabulations is large, only a summary table is presented. The more important and consistent relationships are discussed below.

As shown in Table 15, the influence of the medical record department supervisor's training was relatively unimportant except for procedures. The accuracy of procedural information was highest in hospitals where the supervisor was trained as a Registered Record Administrator (RRA) or "Other." The latter category included persons with many years of experience, some of whom had taken correspondence courses and were eligible for certification as an Accredited Record Technician (ART) or were presently enrolled in ART training programs. The training of persons who entered the diagnostic or procedural code on the

---

5 Because of the instability of the weighted numbers, the chi-square was based on a re-distribution of the unweighted numbers according to the weighted percentages. A statistically significant relationship was assumed if the chance of its occurence was less than .05.

4 It would have been possible to re-distribute the abstracts for which the correct data source was indeterminate and augment the percent for which the original abstract was correct by applying the estimating procedures used in Table 8. But the calculations were extremely complicated. Re-distributions were used in a few cross-tabulations. However, the resultant changes were so slight that they did not affect the conclusions drawn from the unadjusted cross-tabulations. Therefore, the indeterminates were not re-distributed in the remaining tables, and the data presented here do not include the adjustments.

Table 15. Relationships between Hospital and Abstracting Process Characteristics and the Accuracy of Information on Diagnoses and Procedures*

Characteristics	Four-digit diagnoses	Three-digit diagnoses	Procedures
Personnel & Training			
Number of full time equivalents per record department	Significant but not meaningful	Significant but not meaningful	Significant but not meaningful
Training of supervisor	NS	NS	RRA & others= better data
Training of coder	Both RRA & ART= better data	Only ART= poorer data	NS
Training for new abstractors	NS	NS	On-the-job training with review=better data
Participation in training in 1974	No training= better data	NS	Training=better data
Communication with abstract service	Monthly or more frequent=better data	Monthly= better data	Monthly or more frequent=better data
Abstracting Process			
Completeness of record when abstracted	Coding when incomplete= better data	Coding when incomplete= better data; before & after completion= poorer data	NS
Time lapse between patient discharge and abstracting	Significant but not meaningful	Significant but not meaningful	Significant but not meaningful
Records abstracted per day	High volume= better data	Moderate volume =poorer data	NS
Review of accuracy of abstracting	NS	Total review= poorer data	NS

Table 15 continued

Characteristics	Four-digit diagnoses	Three-digit diagnoses	Procedures
Comprehensiveness of record review	Face sheet only =poorer data	Face sheet only =poorer data	Face sheet only =poorer data
Hospital Characteristics			
Control	For-profit= poorer data	For-profit= poorer data	For-profit= poorer data
Presence of house staff	Interns only= poorer data	Interns only= poorer data	Interns only= poorer data
Medical school affiliation	NS	NS	NS
Review and use of data by medical staff	Use=better data	Non-use=better data	Non-use= better data
Bed size	Significant but not meaningful	Significant but not meaningful	Significant but not meaningful
Geographic location	SMSA=better data	SMSA=better data	SMSA=better data

*NS = Not significant; significance is accepted if the likelihood of the relationship occurring by chance is less than .05.

abstract was not significantly related to the accuracy of information on procedures, but was to diagnostic information. Hospitals in which both RRAs and ARTs determined codes had the best data compared to four digits. For three-digit comparisons, hospitals in which only ARTs entered codes had poorer data than if both RRAs and ARTs coded or if ARTs and "others" coded.

On-the-job training of newly hired abstractors combined with a review of their work was associated with higher accuracy of procedural data than either on-the-job training without review or other types of training. Participation by medical record staff during 1974 in training programs specifically directed at abstracting and coding medical record information was positively related to the accuracy of information on procedures only. It was inversely correlated with the accuracy of diagnoses coded to four digits. Communication between the medical record department and abstracting service was positively associated with the accuracy of both diagnostic and procedural coding.

Several questions considered the process by which abstracting occurred in the hospital. The quality of diagnostic information analyzed to four digits was better when coding occurred before the record was complete. For three-digit comparisons it was worse when coding occurred either before or after the record was complete--perhaps indicating the lack of established abstracting procedures. The relationship between the accuracy of data and the number of weeks from patient discharge to completion of the abstract was statistically significant for all information, but there was no logical explanation for the relationship. Similarly, the number of records abstracted per day did not have a consistent relationship to the accuracy of data. For all three variables, the variance within response groupings approximated the variance between groupings, suggesting that the data are so unstable as to have little meaning. Review of the accuracy of abstracts by a supervisor, based on a total review, sample, problem cases only, or some other method was not related to the quality of data except for diagnoses compared to three digits. Here, total review was was associated with poorer data--possibly suggesting that review was needed but ineffective, or that the reported degree of supervision was somewhat exaggerated.

The thoroughness with which the medical record is reviewed before diagnoses and procedures are coded consistently influences the quality of data. In hospitals where only the face sheet is routinely reviewed, the quality of data is poorer for procedures and diagnoses compared to both three and four digits.

For-profit hospitals were consistently likely to have less accurate data than either government (non-federal) or non-government (not-for-profit). But the representation of for-profit hospitals in the sample is so small that this finding is tenuous. Hospitals with an internship program only, rather than a residency program, both, or neither, had consistently less accurate data. Hospital location within a Standard Metropolitan Statistical Area was associated with more accurate diagnostic and procedural data. The number of bed per hospital was significantly associated with the accuracy of information, but there is no apparent explanation for the relationship. Bed size is correlated with other variables, however, which probably influence the quality of data more directly.

A final question attempted to determine the extent to which specific abstract service reports were regularly reviewed and used by either medical record department personnel, hospital administration, or the medical staff. Almost without exception medical record departments reportedly used the reports, so responses were not cross-tabulated by the accuracy of data. Because of the sizable number of "not certain" responses for administrative use of reports, those data were not analyzed. In hospitals where the medical staff regularly used abstract service reports, however, diagnostic information compared to four digits was more accurate than where the reports were not regularly used by medical staff. For three-digit comparisons and information on procedures the relationship did not hold. One may speculate that the interest of the medical staff or its leadership was instrumental in increasing the fourth digit accuracy levels.

In an attempt to ascertain the relative influence of hospital characteristics on reliability of data, simple and multiple regressions were performed with the more meaningful variables. The extent to which association with a particular abstracting service might be influential could not be examined directly because of prior assurances of confidentiality to the services; however, abstract service was introduced into the regressions as a dummy variable. The only independent variable which was consistently associated with accuracy of data was location within a Standard Metropolitan Statistical Area. For all regressions, the amount of variance explained was low, reaching a maximum of .369.

Of all the structural and functional characteristics examined, only the thoroughness with which the medical record is reviewed, communication with the abstract service, medical staff review of abstracted reports, and possibly, hospital location within a Standard Metropolitan Statistical Area are consistently and meaningfully associated with improved accuracy of abstracted information. Because of the frequent difficulty encountered by the field team in determining which of several diagnoses should be regarded as principal, the care with which the physician completes the record must be considered an additional, influential structural variable, even though it was not measured directly.

Chapter 5

SUMMARY AND RECOMMENDATIONS

The study reported here was conducted by the Institute of Medicine to assess the reliability of hospital utilization data compiled by private abstracting services during calendar year 1974 and based on abstracts of medical records. This information was requested by the Office of Quality Standards, Department of Health, Education, and Welfare, to assist in identifying an existing and accurate source of data to serve as a baseline for measuring the impact of Professional Standards Review Organizations. However, the findings have much broader implications, since the data can be used for internal hospital management, health services research, and policy analysis, provided they are sufficiently reliable.

It should be noted that the study was not intended to compare the relative reliability of the participating abstract services, which were assured of anonymity in the analysis. The data do not permit such comparisons. Furthermore, all participating services have revised their procedures and edit systems since 1974. The combined data do permit the analysis of various aspects of data reliability and the generation of national estimates at one point in time.

The accuracy of seven information items from the original abstract was determined by comparing those items with the results of an independent abstracting of medical records by a trained field team and noting the frequency and type of discrepancies. In addition, selected characteristics of the abstracting process within hospitals and interactions between the hospitals and abstract services were related to the accuracy of abstracted data to assist in identifying areas for improvement by either hospitals or services.

The analysis showed that information on hospital admission date, discharge date, patient's date of birth or age, sex, and anticipated principal source of payment was highly reliable. However, for all principal diagnoses combined, when codes were compared to four digits, the original abstract and Institute of Medicine re-abstract agreed for only 65.2 percent of the records; for all procedures the level of agreement was 73.2 percent.

Interpretation of the findings for diagnoses and procedures and the generation of recommendations for the use of such data is extremely difficult. The information is highly technical, and almost every statement needs to be qualified. Ultimately, value judgments are required. Nevertheless, several factors help explain the lower reliability of information on diagnoses and procedures.

1. The degree of coding refinement and breadth of diagnostic groupings influence the reliability of data. For all diagnoses combined, three-digit comparisons were more reliable than four-digit. For both three and four-digit comparisons, the reliability of the data base increased when all diagnoses except target diagnoses were excluded from the analysis. (Target diagnoses are those conditions included in the PSRO Evaluation Plan, in part, because they were assumed to be diagnosed with greater accuracy.)

2. Individual diagnoses contribute to overall reliability. For some diagnoses, the data were quite reliable. As examples, no discrepancies were found on 89.1 percent of the abstracts for patients with hernia without obstruction, 81.0 percent for hernia with obstruction, 86.5 percent for delivery, 97.0 percent for hypertrophy of tonsils and adenoids, 94.3 percent for cataract, and 100 percent for personality disorders--all analyzed to four digits. On the other hand, discrepancies were detected on 43.7 percent of the abstracts for low back pain and 69.8 percent for chronic ischemic heart disease. If chronic ischemic heart disease is removed from the data base, the reliability of data for all diagnoses combined increases by about ten percent.

3. The reasons for discrepancies vary by diagnosis. For most diagnoses with high discrepancy rates and a likelihood of co-morbidity where the re-abstract was found to be correct (especially chronic ischemic heart disease and diabetes), discrepancies occurred primarily because of erroneous selection of principal diagnosis (an ordering discrepancy), rather than because of mistakes in assigning a code number (a coding discrepancy). That is, the person originally abstracting information from the medical record and forwarding it to the abstract service had difficulty determining from the information recorded which diagnosis should be regarded as the principal diagnosis responsible for hospital admission.

4. In some instances, the field team detected a discrepancy between their work and the original abstract, but after re-examining the medical record, they were unable to state with certainty which was correct. Instead, they concluded that it was a matter of judgment and that the correct data source could not be determined. Again, the problems stemmed from difficulty in determining which diagnosis should be regarded as principal. For 10.7 percent of all diagnoses combined and for 16.3 percent of all procedures, the correct data source could not be determined. The comparable percents for individual diagnoses ranged from 18.0 percent for low back pain to 0.3 percent for fracture of neck of femur.

5. If one examines the number of cases that should have been coded as target diagnoses but were not, the influence of false positive and false negative diagnoses on admission rates and lengths of stay can be assessed.

Proxy admission rates based on the Institute of Medicine re-abstract data (including false negatives) suggest that reliance on abstract service data with three-digit comparisons will under-estimate the number of admissions for cerebrovascular disease, acute myocardial infarction, cholelithiasis/cholecystitis, delivery, fracture of neck of femur, disorders of disc, cataract, and neuroses. For the remaining target diagnoses, abstract service data will over-estimate the number of admissions--most notably for chronic ischemic heart disease. The discrepancies increase with four-digit comparisons. Similarly, lengths of stay based on abstract service data (including false positives) may over-estimate the average stay for both three and four-digit analyses.

6. Some characteristics of hospitals and the abstracting process are consistently and meaningfully associated with improved accuracy of information. These include the thoroughness with which the medical record is reviewed before code numbers are assigned, frequency of communication between the hospital and abstracting service, regular medical staff review and use of abstracted reports, and possibly, hospital location within a Standard Metropolitan Statistical Area. Because of the frequent difficulty encountered by the field team in determining which of several diagnoses should be regarded as principal, the care with which the physician completes the record must also be considered an influential structural variable, even though it was not measured directly.

The apparent contribution of many factors to the accuracy of abstracted medical record information suggests that additional study is needed before definitive prescriptions for using the data can be offered. The limitations of this study must also be acknowledged. A very restricted set of information was examined which is relevant primarily for utilization review. Very little is known about the reliability of abstracted history, physical examination, or laboratory findings that may be extremely important for medical care evaluation. Though the data are generalized to all Medicare and Medicaid patients who had one or more of the study diagnoses and were treated during 1974 in hospitals subscribing to one of the four participating abstract services or larger hospitals with internal data systems, the unweighted frequencies come from only 50 hospitals. The sampling plan resulted in an under-representation of false negative diagnoses. The weights used in the analysis may have introduced some instability. Furthermore, there may have been instances in which the field work was not reliable. These limitations notwithstanding, the committee offers the following recommendations for using abstracted hospital discharge data.

RECOMMENDATIONS

1. One must assume that abstracted hospital data contain errors and use them with caution. The seriousness of the error depends on the purpose to which the data are applied.

2. Existing abstracted data are adequate to describe general utilization patterns such as age or sex differentials or to compare average lengths of stay among hospitals. The influence of false positives and false negatives on the selection of principal diagnosis and the associated coding errors probably are

not sufficiently great to distort the use of such data for general program management and monitoring purposes. However, if such data are to be used for research or evaluation and, in particular, to assess the effects of specific changes in the health care delivery system such as the imposition of utilization review or PSRO programs on patterns of patient care, then more stringent precautions should be taken.

3. Some adjustments may be made to increase the reliability of existing and future information (assuming that the same kinds of errors persist). The accuracy of diagnostic information increases when diagnostic codes are analyzed to three digits, rather than four. For determining basic utilization trends and, in some instances lengths of stay, three-digit analyses are sufficient. For frequently occuring heart conditions, including chronic ischemic heart disease, groupings which are even broader than three digits may suffice, since the high volume of cases may off-set the increased statistical variance. With careful research, it may be possible to devise multiple diagnostic groupings to combine conditions that frequently occur simultaneously (particularly in the elderly), thereby avoiding problems associated with determining principal diagnosis. However, the increased reliability resulting from less precise coding must be balanced against the loss of ability to detect the presence of complications (usually denoted by the fourth digit) which may significantly affect length of stay.

4. Further improvements in reliability may be obtained by confining analyses to diagnoses which appear to be more accurately abstracted and either excluding the less reliabile diagnoses or intensifing efforts to improve their accuracy.

5. If quality assurance programs discontinue the current practice of reviewing all patients and physicians and move to a more targeted review of cases likely to be associated with poor quality (as recommended in the basic study report), in many cases this will require improving the data base in order to detect changes in utilization patterns. As an example, it is quite likely that criteria for hospital admission and continued stay for a patient with uncomplicated diabetes mellitus (code 250.0, using the H-ICDA system) would be quite different from criteria for a diabetic with acidosis or coma (code 250.1). In order to evaluate the effect of review, it is essential that diagnostic information be accurately coded to the fourth digit.

6. One consideration for selecting cases for targeted review might be the likely reliability of data on diagnoses and procedures. This should not be the only criterion, however, since it might result in eliminating from review those diagnoses or conditions for which both the quality of care and data are questionable. In such cases it would be important to intensify review efforts, but also to improve medical recording and diagnostic coding.

7. Whenever abstracted data are used to measure changes in utilization patterns, the amount of error, including the influence of false negative and false positive diagnoses, must be assessed at each time that measurement occurs. This is essential in order to determine whether perceived changes are truly associated with altered utilization or, instead, with changes in the reliability of data.

8. Periodic assessments of the reliability of PSRO data should be included in the information reporting requirements of the PSRO Management Information System. The Bureau of Quality Assurance should develop guidelines to assist in such assessments so that the reliability of both local and nationally aggregated data will be known.

9. Physicians should be encouraged to use more care in completing the medical record, to clearly indicate the diagnosis responsible for hospital admission, and to become more involved in analyzing reports based on abstracted medical record information. This is particularly important if the profile analysis component of PSRO review is to identify meaningful areas for more intensified concurrent review and medical care evaluation.

10. To improve the quality of abstracted information, record department personnel should review the body of the medical record and not just the face sheet before abstracting information on diagnoses and procedures. If a total review of all records is not practical, sampling techniques should be devised that would at least require a review of the complete record for patients with diagnoses for which abstracted information is known to have low reliability. Although training is not consistently associated with improved data, additional training which specifically addresses the abstracting and coding of information on diagnoses and procedures should improve the quality of data.

11. Abstract services should take a more active role in training hospital medical record personnel, monitoring the quality of their work, and furnishing estimates of the amount of error in data they provide. If error is introduced by hospital personnel, an intensified technical assistance effort by the services might help to make the data more useful to their clients.

12. Data recording and reporting guidelines should continue to require that diagnoses are coded to four digits. This is the only way to assure that the resulting data base will have sufficient flexibility to meet a variety of data needs--particularly in health services and epidemiological research and evaluation. For less precise requirements, the data can easily be analyzed to three digits only.

The steering committee believes that uniform, comprehensive, and reliable data bases are essential in order to plan for programs to meet the health needs of the nation and evaluate their effectiveness. This report is intended to assist in improving one potential data base on which crucial planning, management, and evaluation activities may proceed.

Appendix A

BIBLIOGRAPHY

Brook, Robert H. "Changes in a Disease Classification System Needed for Assessing Quality of Care," in Proceedings of the 15th Annual Public Health Conference on Records and Statistics. DHEW Pubn. No. HRA-75-1214.

Cooperative Health Information Center of Vermont, Inc. Notes on Data Quality: A Review of the Reliability of a Hospital Discharge Abstract System for Use in Quality Assurance and Regional Planning. Public Health Service Grants PHS-RMO 303 and HSO 1197. September 1974.

Diamond, Earl L., and Lilienfeld, Abraham M. "Effects of Errors in Classification and Diagnosis in Various Types of Epidemiological Studies." American Journal of Public Health 52 (July 1962): 1137-44.

Doyle, Donald N. "Accuracy of Selected Items of Blue Cross Claims Information." Inquiry 3 (September 1966): 16-27.

Duggar, Benjamin C. Evaluation of the Uniform Hospital Discharge Abstract Site Test. Springfield, Va.: National Technical Information Service, March 1973.

Hendrickson, Leslie, and Myers, Jeffrey. "Some Sources and Potential Consequences of Errors in Medical Data Recording." Methods of Information in Medicine 12 (January 1973): 38-45.

Ensign, James M., Hodgson, David A., and Kucken, Lawrence E. Uniform Hospital Discharge Data Demonstration--Summary Report. U.S. Department of Health, Education, and Welfare, Public Health Service, Pubn. No. HRA-74-3102, July 1973.

Lindberg, D.A.B. "Collection, Evaluation, and Transmission of Hospital Laboratory Data." Methods of Information in Medicine 6 (July 1967): 97-107.

Loup, Roland J. "Consistency of Final Diagnosis Coding in a Hospital Patient Record Data System." Paper presented at the Seventh Annual Interdisciplinary Conference on Health Records, Chicago, Ill., June 1976.

Marienfeld, C.J. Analysis of Discharge Diagnosis--Second Progress Report. Contract No. HRA 106-74-18, 1 February 1975.

Appendix A

U.S. Department of Health, Education, and Welfare. Quality Control in the Hospital Discharge Survey, by K. W. Harris and K. L. Hoffman. Vital and Health Statistics, Series 2: Data Evaluation and Methods Research, No. 68. Washington, D.C.: Government Printing Office, December 1975.

Wagner, G. "Quality Control and Error Checking." In Information Processing of Medical Records, pp. 222-26. Edited by J. Anderson and J. M. Forsythe. Amsterdam, The Netherlands: North Holland Publishing Co., 1970.

Appendix B

SAMPLE DESIGN FOR RE-ABSTRACTING STUDY*

The details of the sample design used in the assessment of the accuracy of abstracted medical record information for Medicare and Medicaid patients are presented in this paper. In planning the sample, several factors had to be considered so that an optimum probability design would result. The Institute of Medicine re-abstracting study was partially intended to assist in planning the national evaluation of Professional Standards Review Organizations (PSROs). This led to the decision that the re-abstracting study sample should be a subsample of a sample of hospitals that could be used in evaluating the PSRO program. The specific objectives of both activities had to be accommodated within the sampling plan.

The re-abstracting study focused on the quality of data produced by abstracting services, based on hospital discharge information abstracted from medical records. Quality of the abstracted information was determined by comparing results of an independent re-abstracting of a sample of records with the original abstract. When differences were found the medical record was reviewed to determine which abstract was correct and why the discrepancy occurred.

The national evaluation of PSROs requires the collection of baseline data to assist in determining the effects of PSROs on utilization of services over time. Baseline data for many hospitals could be compiled from existing information produced by abstracting services, although not all hospitals subscribe to these services. Any sample of hospitals designed for developing baseline data needs to take into account the existence of available abstracted data for the baseline year and the relative costs for use of such data, together with the costs for original abstracting in hospitals that had not had data abstracted for that year. The relative frequency of subscription to each of the several abstracting services by type of hospital and volume of discharges abstracted is an important factor in the tie-in between the design of a subsample for assessing quality in the re-abstracting study and the overall sample of hospitals for a national PSRO evaluation.

GENERAL DESIGN CRITERIA FOR EACH STUDY

The starting point for the design of the PSRO evaluation sample was the statement of some design considerations and precision goals for important statistics to be derived. The subsample for the re-abstracting study was to

*A report submitted by Survey Design, Inc., Silver Spring, Maryland, May 15, 1976.

be designed to achieve maximum precision within budget and administrative constraints.

PSRO Evaluation Study

Design considerations were:

1. The data obtained should be able to provide separate estimates of equal precision for each of the four Census regions.

2. The effects of sampling variability in measuring change between the baseline period and any subsequent year should arise only from the sampling required to obtain base year data. All data for subsequent years should be available from 100 percent tabulations required by PSRO regulation.

3. Stratification variables (in priority order) were:

- nine census divisions;
- hospital bed size: under 50, 50-99, 100-199, 200-299, 300-499, 500-999, and 1,000 beds and over;
- population density: SMSA, or non-SMSA; and
- ownership: government, voluntary, and proprietary.

Other stratification variables considered but not used were: hospital beds per 1000 population, extended care beds per population 65 years of age and older, medical school affiliation, and JCAH accreditation. For the first two variables, the available data were aggregated by state and therefore not sufficiently useful. Use of the latter two variables would have made the "controlled selection" process more involved than was desirable or feasible. 1/

4. A change in average length of stay of one-half day should be at the threshhold of significance at a 90 percent confidence level.

5. A change in the number of admissions of 10 percent should be at the threshhold of significance at a 90 percent confidence level.

6. The distribution of Medicare and Medicaid patients by state should be taken into account in the probability of selection of hospitals for the evaluation study.

1
R. Goodman and L. Kish. "Controlled Selection - A Technique in Probability Sampling," Journal of the American Statistical Association 45 September 1950): 350-72.

Appendix B

Re-abstracting Study

Design considerations were:

1. There was an overall budget limit of about $17,000 in direct costs to cover costs of the field work, including salaries, per diem, travel time, and travel costs.

2. Net and gross difference rates should be determined for each of 14 specified diagnoses as best as possible overall.

3. The sample for re-abstracting in each hospital should be of sufficient size so that a study examining gross difference rates and characteristics of hospital record abstracting (training of abstractors, length of service, qualifications, etc.) would be technically feasible.

4. There should be reasonable representation of hospitals serviced by each of the abstracting services.

5. Each re-abstractor would carry out his/her work by independently re-abstracting data required for the sample of discharges specified by the central staff, comparing re-abstracted data with (previously confidential) original abstract data, and determining the final re-abstracted result by reconciliation between original and re-abstract, where results differed.

6. The quality of the re-abstracting would be measured by independently checking a subsample.

DETAILS OF DESIGN FOR THE BASIC PSRO EVALUATION

It was initially decided that the universe of hospitals for the PSRO evaluation should include all short term general medical and surgical hospitals other than those controlled by a federal government agency. 2/ The National Center for Health Statistics (NCHS) 1973 Master Facility Inventory (MFI) tape was used to define the universe of interest as 5,944 general hospitals.

The design was importantly affected by the fact that not all hospitals subscribe to abstract services. Therefore, the basic PSRO evaluation study design had to take into account the differences in cost for accessing hospital record information for the base period. Since separate regional estimates were required, equivalent sample sizes (except for finite correction factors) were required in each region. The stratification objectives were considered to be well met if:

2
Average stay of inpatients is less than 30 days.

Appendix B

- Each of the nine census divisions were used as strata (within their regions); and,

- Controlled selection was used for hospital bed size crossed by population density and hospital ownership.

Therefore, the design required the use of two-way controlled selection applied to the 5,917 general hospitals that remained after excluding the 27 hospitals with over 1,000 beds which were included in sample with certainty. The data on Medicare-Medicaid patients were introduced as a proportionality factor within the overall regional allocation of sample size of number of hospitals.

The precision requirements were considered next, in the context of available information on precision and estimates of cost.

Variance and Cost Components

Some estimates of variances for the design objectives were available from the Hospital Discharge Study (HDS). 3/ (The general character of the HDS seemed consonant with the design character of the PSRO evaluation study. The research effort prior to establishment of the HDS had considered then available evidence on variance component ratios and estimates of cost in establishing an optimum probability design for HDS.) The relative error for the estimated 17 million discharges was 2.7 percent, based on HDS 1973 data. Similarly, the relative error on average length of stay was 3.4 percent. 4/

As noted earlier, costs of accessing abstracts would vary, depending on whether a new abstracting effort was required or whether abstracts were directly available from an abstract service. It was speculated that about 65 to 70 percent of hospitals were covered by the five or six major abstracting services and that about 75 percent of the discharges were abstracted by these major services. Based on costs of the ongoing HDS (and making allowance for introductory visits to the new hospitals) it was estimated that $3.20 would be the unit cost of an abstract prepared especially for the PSRO evaluation. On the other hand, it was estimated that access via computer to existing abstracted information would cost $0.10, with an additional $0.10 per abstract for sample selection regardless of whether the hospital subscribed to an abstracting service.

3
National Center for Health Statistics, series 2, no. 39, 1970, "Development of the Design of the NCHS Hospital Discharge Survey;" Unpublished data for 1973 provided by Division of Health Resources Utilization Statistics, NCHS.

4
Based on data from 424 hospitals and about 225,000 discharges.

Appendix B

Double Sampling - Optimization

The object of collecting variance and cost components was to acquire the data to determine an optimum allocation of effort between hospitals which subscribe to an abstracting service and those that do not, subject to the constraints of the precision requirements.

Available data on variance was examined in relation to precision goals. (The examination used the HDS sample design as a standard of reference, but with the assumption that for the PSRO evaluation, the sample would be restricted to the target populations and therefore would achieve for the same size sample of the target population, the equivalent precision for that population.) Of the two precision goals, the precision requirement for estimating change in the average length of stay was determined to be more demanding. A single phase design would have implied a need for some 225,000 abstracts, with about 373 the optimum number per hospital in about 600 hospitals.

To better determine the appropriate mix between and within the two groups of hospitals, however, double sampling was needed, since the cost per abstract would be substantially less in hospitals where discharge information had been abstracted previously. Using standard formulae in double sampling 5/, the unit costs involved in each type of abstracting and the proportion of each type led to the determination of the following design specifications:

1. In each region, the starting point for the sampling, subject to the effects of a finite correction factor within bed-size strata, was calculated to be a sample of 1,072 hospitals and 400,000 abstracts. Hospitals without a major abstracting service were to be subsampled at the rate of 1 in 4 for the second phase sample.

2. Application of the finite correction factor within bed-size strata by region reduced the required numbers of hospitals as follows:

	Total	Certainty	Non-certainty
Northeast	589	11	578
North Central	648	5	643
South	672	10	662
West	621	1	620

3. Within each sample hospital, the number of abstracts for the target population should be about 373 for the base year.

---

5 Morris H. Hansen, William N. Hurwitz, William G. Madow, Sample Survey Methods and Theory, vol. 1: Methods and Applications (New York: John Wiley and Sons, 1953), pp. 464-75.

4. It was recognized that considerations of cost or precision could dictate a reduced sample of hospitals. Therefore, the planned controlled selection of hospitals was undertaken so that the non-certainty sample in each region would be selected in 10 representative probability subsamples. Any random combination of the probability subsamples, plus the 27 certainty hospitals, can be used as a representative probability sample of the planned hospital universe. The effect on precision of using a subsample can be determined by multiplying the standard error of the full sample by the square root of the reciprocal of the proportion of the sample used.

Use of Medicare-Medicaid Data

The target population for both the PSRO evaluation study and the re-abstracting study is the universe of patients covered by Medicare or Medicaid. The number of admissions during 1973 for Medicare and Medicaid combined was used to distribute the total regional sample sizes of hospitals to each of the states.

Controlled Selection

A two-way controlled selection procedure was used to select the sample of hospitals in order to increase the precision of study estimates beyond those that would be obtained with simple stratification using the same stratification variables.

The general ideas of controlled selection are:

- Assigned probabilities for each sampling unit are maintained.
- The probabilities of selection of some (or all) preferred combinations of n out of N units are larger than in simple stratified sampling.
- The probabilities of selection of some non-preferred combinations are smaller than in simple stratified sampling.

Two-way controlled selection was executed separately for each of the nine geographic divisions, using the aggregate state within geographic division allocation of sample hospitals as one dimension of control. The other dimension within each of six bed-size strata was hospital control within type of area (SMSA and non-SMSA).

The 1973 MFI sampling frame of 5,944 hospitals was classified by the nominal variables to be used in the controlled selection (state, bed-size, hospital control, and type of area). The data on number of hospitals and aggregate numbers of beds for this universe of hospitals were tabulated. The data on number of beds were used to allocate sample size controls, by state, to the six bed-size strata.

The data on number of hospitals by type of control and area, within state and bed size stratum, were used to allocate sample size controls to these additional nominal variables. These data established the expected values of sample sizes of hospitals for each of the cells in each of the controlled-selection strata for each of the ten subsamples. These results determined a set of tables of joint probabilities of selection of elements that are the preferred combinations.

A set of alternative samples (patterns) was developed which satisfied the constraints of the expected values for each cell. Associated with each pattern was the probability of selection of the pattern. All patterns in aggregate exhausted the probability of selection for each cell. Then random numbers were used, systematically, to select the ten patterns for this sample and determined the exact sample count of hospitals for each cell in each of the ten patterns. These exact sample counts, in proportion to the universe counts, established within cell sampling ratios by cell. A systematic sample, after a random start, in each cell selected the actual sample of hospitals in the cell for each of the ten patterns.

Basic Weights

The inverse of the probability of the selection of a hospital is its basic weight. For the 27 certainty hospitals, this weight is one. The basic weights for one of the ten patterns were calculated. If more than one pattern is used, these basic weights should be divided by the number of patterns used for the given sample.

DETAILS OF DESIGN FOR RE-ABSTRACTING STUDY

The re-abstracting study sample design fits within the structure of the PSRO evaluation design.

Subsample of Panels

The re-abstracting study concentrated on hospitals subscribing to one of four abstracting services. One design consideration was the desirability of achieving reasonable representation of hospitals served by each abstracting service. The smallest service had but three or four subscriber hospitals per panel. Therefore, two panels from the larger sample of hospitals were randomly selected for inclusion in the re-abstracting study.

Sample of Hospitals

The re-abstracting study sample design was established primarily to satisfy the first and third design considerations--overall budget level and sample size needs in each sample hospital--subject to the constraints of the

other design considerations. Several combinations of numbers of sample hospitals and re-abstracts per hospital were considered which satisfied the budget constraint. The second design consideration--best possible information for each of 14 specified diagnoses, overall--indicated a need for as large a hospital sample as possible. This assumes that differences for any given diagnosis would be clustered in some hospitals and relatively infrequent in others. A limited amount of information concerning this assumption was available from a prior study. 6/ The third consideration suggested that the within-hospital sample be relatively large for the 14 specified diagnoses. In hopes of balancing the objectives, the decision was reached to aim for about 50 cooperating hospitals and about 80 abstracts per hospital. About six to eight sample hospitals each was the goal for the HUP, QUEST, and CHDC abstracting services, and the balance of about 30 hospitals for CPHA. In addition, several certainty hospitals (those with over 1,000 beds) were to be included. The sample was increased by about 20 percent to allow for unwillingness to cooperate.

For the two patterns of hospitals selected for study and the certainty hospitals, the first step in sample selection required matching with lists of hospitals serviced by the four abstracting services. After matching, the hospitals were identified by state, bed-size, area, and type of control. Overall subsampling ratios for each abstract service were established so that the sample size goals for each would be satisfied. An approximation to controlled selection was used in the systematic probability subsampling of the matched hospitals by abstract service.

During the course of the recruitment effort for the study, 70 hospitals were asked to participate. Five of these had internal systems; two agreed to cooperate. Of the remaining 65, 63 were asked to cooperate initially and two subsequently. Of the 63 initial sample hospitals, two were not serviced by their abstracting service in the base period and 13 others were unwilling to cooperate. The two invited to participate subsequently were selected to replace the hospitals that were ineligible during the base year, but were unwilling to cooperate. In total, 50 hospitals agreed to and did particpate in the study.

Sample of Abstracts

The statistical objectives of the re-abstracting study were to produce both net difference rates and gross difference rates for the 14 "target" diagnoses of interest. In order to do this, abstracts that satisfied the following conditions were used as basic sampling frames:

6
Bio-dynamics, Inc., "Evaluation of the Uniform Hospital Discharge Abstract Site Test," Final Report, Contract no. HSM 110-71-224, March 31, 1973.

- Abstracts for which the "target" diagnosis is the primary diagnosis.

- Abstracts for which the "target" diagnosis is listed but is other than primary. Twelve of the 14 diagnoses were considered in this frame.

- Abstracts for which a "satellite" diagnosis is the primary diagnosis, regardless of whether a "target" is listed as other than primary. "Satellite" diagnoses are those which are often mistakenly coded as primary in place of the correct target diagnosis; six "satellite" diagnoses were identified.

These major groupings and the diagnoses within them yielded 32 independent sampling frames. In each study hospital, the number in the universe in each of these 32 frames was established. For each of the 14 primary "target" diagnosis frames, a systematic probability sample was selected to yield four sample abstracts. For each of the remining 18 frames, one abstract was selected at random. Thus, a maximum of 74 abstracts were to be included in the sample for each of the sample hospitals. Fewer (all) were to be included when the universe for the given frame had less than the desired sample size goal.

Weighting

The two sampling processes--hospital sampling and abstract sampling--each played a role in defining the weighting fractions used in the analysis. The product of two weighting factors was required for determining the weight associated with each sample abstract.

The hospital weights took into account three factors:

- the basic probability of hospital selection (both for the PSRO evaluation and the subsampling for the re-abstracting study);

- adjustments for changes in universe designations (after sample hospitals were identified); and

- adjustments for non-cooperating hospitals (by abstracting service).

The abstract weights took into account all available abstracts. The weight used is the reciprocal of the joint probability of selection associated with an abstract. The total set of abstracts in a given hospital was sampled independently several times. Each "file" when sampled was not mutually exclusive of all other files. Therefore, the weight of each sample abstract reflects this sampling process in order to produce unbiased estimates. Several conditions exist in connection with a given abstract and were refelcted in the weight:

Appendix B

o An abstract included in sampling for a given "target" diagnosis as a primary diagnosis may have had one or more chances of being included in the sampling of 11 or 12 diagnoses listed as other than primary;

o An abstract included in sampling a file where any one of six "satellite" diagnoses is primary may have had one or more chances of being included in the sampling of any of the 12 "target" diagnoses listed as other than primary;

o An abstract included in sampling a specified one of 12 "target" diagnoses as other than primary, may have had a chance of being included in the sampling of any one of 13 other "target" diagnoses as primary, or one or more chances for 11 other "target" diagnoses listed as "other than primary", or a chance of being included in the sampling of any one of the six "satellite" diagnoses.

In the within-hospital sampling process the joint probabilities are additive of the individual file sampling probabilities when multiple selection chances are found to occur. If a given abstract was selected t times, its weight was t times as large as if it were selected but once.

Evaluation of the Accuracy of the Field Work

The sixth objective of the re-abstracting study required an evaluation of the quality of the field work. This evaluation was to be done independently for a subsample of the abstracts. The work of each re-abstractor was subdivided into two time periods: early work and late work. Within each time period a subsample of one hospital was selected for each re-abstractor. The overall goal was to obtain a probability sample of half the volume of re-abstracts to be quality checked, of which no more than one-half would be those with differences between the re-abstracting and original data and the balance were without differences.

Appendix C

NATIONAL ACADEMY OF SCIENCES
2101 CONSTITUTION AVENUE
WASHINGTON, D. C. 20418

INSTITUTE OF MEDICINE
OFFICE OF THE PRESIDENT

SAMPLE LETTER TO HOSPITAL ADMINISTRATORS

December 5, 1975

Dear _____:

The Institute of Medicine has contracted with the Department of Health, Education, and Welfare to conduct an evaluative study of systems for monitoring the quality of medical care. This year-long project is a partial response to a congressional request, which was included in Section 4(a) of the Health Maintenance Organization Act of 1973 (P.L. 93-222).

As part of this broad study, the Institute has been asked to determine the extent to which hospital utilization data obtained through the major abstracting services, including the (_____abstract service) to which your hospital subscribes, can be used for research purposes related to the assessment of the quality of medical care. This task is of particular interest to the Office of Professional Standards Review, DHEW, and the National Professional Standards Review Council, which are responsible for generating baseline data to evaluate the impact of the PSRO program, created under Public Law 92-603.

The study is being conducted by an expert steering committee and project staff under the chairmanship of Robert J. Haggerty, M.D., Roger I. Lee Professor of Health Services and Child Health at Harvard University. It has been approved by the governing board of the National Research Council. A brief summary of the goals and activities of the Institute of Medicine is enclosed.

The purpose of this letter is to request your assistance in conducting the study. More specifically, we are asking you to permit a member of our field team to abstract approximately 75 pre-selected medical records within your medical record department. This information will form the basis for our analyses, which will concentrate on establishing the reliability of hospital discharge abstract data. The field work will be conducted by three highly skilled Registered Record Administrators (RRAs) with extensive research and administrative experience, who have been specially trained by Faye Brown, a former president of the American Medical Record Association. Because of their proven competence, they will be able to complete their work with no significant disruption of your record department's regular routine, although some staff time will be required to pull the records in advance of their visit.

Appendix C

Since the study does necessitate the review of medical records, we are very concerned about confidentiality and have taken special precautions in this regard. The unique patient number will only be used to access the relevant record and to assist in compiling the abstracted information. In no instance will the name of the patient or physician be recorded. The resulting report will contain only statistical summaries, which will not permit the identification of patient, physician, or hospital.

(_____abstract service) has agreed to participate in this study and believes that the findings will be very valuable. However, (_____abstract service) obviously will not make available to us any information on discharges from your hospital without your explicit approval.

Enclosed is an agreement whereby you may authorize the (_____abstract service) to release the information needed for our analyses. We hope that you will review this agreement and sign it on page three in the space provided for "hospital." Once signed, this form should be returned to (_____abstract service administrator) in the pre-addressed envelope enclosed.

It may be worthwhile noting that this request is for a single study only and does not involve ongoing or repetitive activities. Your participation is strictly voluntary, of course; and there is no penalty for non-participation. However, we sincerely hope that it will be possible for you to join us in this important undertaking. The results of this study will have broad applicability for health services research and evaluation in general. They should also be valuable from the hospital's viewpoint in determining the utility of such data for internal management information. Your hospital was selected through a very carefully designed sampling procedure, and your participation is essential if the results of the study are to be truly generalizable. We will, of course, be most happy to share with you the results of the completed study.

Within a few days Sarah Brown of the Institute of Medicine staff will contact you to see if you have any questions regarding the study. Needless to say, we hope that you will respond to this request in the affirmative.

Sincerely,

David A. Hamburg, M.D.
President
Institute of Medicine

cc: Medical Records Room Supervisor
Administrator - Abstract Service

Appendix D

TARGET AND SATELLITE DIAGNOSES AND CORRESPONDING DIAGNOSTIC CODES

The sample of records for the re-abstracting study focused on 14 diagnoses, referred to as "target" diagnoses, which correspond to the conditions in the PSRO Evaluation Plan for which changes in utilization of services will be evaluated. Records were selected when the target diagnosis was was listed as the "principle diagnosis" and also when it appeared as one of the "other diagnoses." The target diagnoses are:

Cerebrovascular disease
Chronic ischemic heart disease
Acute myocardial infarction
Diabetes mellitus with and without acidosis
Malignant neoplasm of the breast
End stage renal disease
Cholecystitis/cholelithiasis
Inguinal hernia without obstruction
Delivery with and without complications
Hypertrophy of tonsils and adenoids
Fracture of femur
Displacement of intervertebral disc
Cataract
Neuroses

"Satellite diagnoses" were also defined. These are conditions which frequently are coded erroneously as the principal diagnosis instead of the correct target diagnosis. The six satellite diagnoses and the target diagnoses with which they often are confused are listed below:

Target Diagnoses	Satellite Diagnoses
Chronic ischemic heart disease and Acute myocardial infarction	Subacute ischemic heart disease
Inguinal hernia without obstruction	Inguinal hernia with obstruction
Hypertrophy of tonsils and adenoids	Acute tonsillitis

Appendix D

Target Diagnoses	Satellite Diagnoses
Fracture of neck of femur	Fracture of other and unspecified parts of femur
Displacement of intervertebral disc	Low back pain
Neuroses	Personality disorders

The two disease classifications used in the study (ICDA-8 and H-ICDA-2) are not completely interchangeable. Each was developed with a slightly different purpose in mind which is reflected in the codes. It is generally true that while the two systems may be roughly comparable at the three-digit level, fourth digit comparability is rare. In drawing the sample, it was not always possible to create identical diagnostic categories in both systems. In a typical diagnostic category, one coding system would exclude a particular diagnosis that was included in the other. For example, the broad category of chronic ischemic heart disease is assigned the general code of 412 in both coding systems. However, ICDA-8 includes arteriosclerotic cardiovascular disease within the 412 category, while H-ICDA-2 places it in 440.9. To make the two systems equivalent, one might consider adding H-ICDA-2 440.9 to the 412 group. But 440.9 includes so many diagnoses in addition to arteriosclerotic cardiovascular disease that such a revision still would not produce the desired comparability. Examples of important differences between the codes are indicated by footnotes in the list that follows:

Target Diagnoses	ICDA-8	H-ICDA-2
Cerebrovascular disease (C.V.A.) a/	430.0	430.0
	430.9	430.1
	431.0	431.0
	431.9	431.1
	432.0	431.2
	432.9	431.3
	433.0	432.0
	433.9	432.1
	434.0	433.0
	434.9	433.1
	435.0	434.0
	435.9	434.1

a
H-ICDA includes some types of malignant hypertension; ICDA-8 specifically excludes malignant hypertension.

Appendix D

Target Diagnoses	ICDA-8	H-ICDA-2
Cerebrovascular disease (continued)	436.0	435.0
	436.9	435.1
	437.0	436.0
	437.9	436.1
	438.0	437
	438.9	438.0
		438.1
		438.2
		438.3
		438.4
		438.5
		438.9
Chronic ischemic heart disease b/	412.0	412.0
	412.9	412.1
		412.9
Acute myocardial infarction	410.0	410.0
	410.1	410.1
	410.9	410.2
		410.3
		410.4
		410.5
		410.6
		410.7
		410.9
Diabetes with and without acidosis c/	250.0	250.0
	250.9	250.1
		250.2
		250.3
		250.4
		250.5
		250.6
		250.7

b
H-ICDA includes arteriosclerotic cardiovascular disease in 440.9, whereas ICDA-8 includes it in the 412 category. To adjust for this difference, one might sample from H-ICDA 440.9 as well. But 440.9 includes so many diagnoses in addition to arteriosclerotic cardiovascular disease that it would not produce the desired comparability.

c
Chemical diabetes mellitus is coded 250.7 in H-ICDA-2. ICDA-8, however, makes no specific mention of chemical diabetes, although one may assume that it is included in the 250 category.

Appendix D

Target Diagnoses	ICDA-8	H-ICDA-2
Malignant neoplasm (carcinoma) of the breast	174	174.0
		174.1
		174.2
End stage renal disease d/	753.1	753.1
	590.0	590.0
	590.1	590.1
	580	590.2
	582	580
		582
Cholelithiasis and cholecystitis	574.0	574.0
	574.1	574.1
	574.9	575.0
	575	575.1
		576.0
Inguinal hernia without mention of obstruction	550	550.0
		550.1
		550.2
Delivery with and without complications	650	650
	651.0	651.0
	651.1	651.1
	651.9	651.2
	652	651.3
	653	651.4
	654.0	651.5
	654.1	651.9
	654.2	652.0
	654.3	652.1
	654.9	652.2
	655	652.9
	656.0	653.0
	656.1	653.9
	656.2	654.0
	656.3	654.1
	656.4	654.2
	656.8	654.3

d
The H-ICDA-2 and ICDA-8 codes for end stage renal disease (ESRD) are not equivalent, especially for renal disease occuring with pregnancy. The ICDA-8 codes exclude ESRD arising during pregnancy or the puerperium; H-ICDA-2 includes some pregnancy-related ESRD.

Appendix D

Target Diagnoses	ICDA-8	H-ICDA-2
Delivery with and without complications (continued)	656.9	654.4
	657.0	654.9
	657.1	655.0
	657.2	655.1
	657.9	655.2
	658.0	655.9
	658.1	656.0
	658.2	656.1
	658.9	656.2
	659	656.3
	660.0	656.4
	660.1	656.5
	660.9	656.6
	661.0	656.8
	661.1	656.9
	661.2	657.0
	661.3	657.1
	661.8	657.2
	661.9	657.3
	662	657.4
		657.5
		657.6
		657.7
		657.8
		657.9
		658.0
		658.1
		658.2
		658.3
		658.9
		659.0
		659.1
		660.0
		660.1
		660.2
		660.3
		660.4
		660.5
		660.6
		660.7
		660.9
		662.0
		662.1
		662.2
		662.9
		663.0

Appendix D

Target Diagnoses	ICDA-8	H-ICDA
Delivery with and without complications (continued)		663.1 663.2 663.3 663.4 663.9 664.0 664.1 664.2 664.3 664.4 664.5 664.6 664.7 664.8 664.9
Hypertrophy of tonsils and adenoids	500	500
Fracture of neck of femur	820.0 820.1 820.2 820.3 820.4 820.5 820.9	820.0 820.1 820.2 820.3 820.4 820.5 820.9
Displacement of intervertebral disc	725.0 725.1 725.8 725.9	725.0 725.1 725.2 725.3 725.8 725.9
Cataract e/	374.0 374.1 374.8 374.9	374.0 374.1 374.2 374.8 374.9

e
ICDA-8 code 374 specifically excludes diabetic cataract (which falls within the 250 diabetes category), whereas H-ICDA-2 includes diabetic cataract within the 374 code.

Appendix D

Target Diagnoses	ICDA-8	H-ICDA-2
Neuroses	300.0	310.0
	300.1	310.1
	300.2	310.2
	300.3	310.3
	300.4	310.4
	300.5	310.5
	300.6	310.6
	300.7	310.7
	300.8	310.8
	300.9	310.9
Satellite Diagnoses		
Subacute ischemic heart disease	411.0	411
	411.9	
Inguinal hernia with obstruction	552	552.0
		552.1
		552.2
Acute Tonsillitis	463	463
Fracture of other and unspecified parts of femur	821.0	821.0
	821.1	821.1
	821.2	821.2
	821.3	821.3
	821.9	821.9
"Low back pain" f/	728.0	728.5
	728.5	789.0
	728.7	789.1
Selected personality disorders	301.0	311.0
	301.1	311.1
	301.2	311.2
	301.3	311.3
	301.4	311.4
	301.5	311.5
	301.6	311.6
	301.7	311.7
	301.8	311.8
	301.9	311.9

f
H-ICDA-2 789.1 includes low back pain of a psychogenic origin, whereas ICDA-8 places this diagnosis in 305.1.

Appendix E

IOM RE-ABSTRACTING FORM

Hospital Patient I.D. Number ☐☐☐☐☐☐☐☐☐☐ 1-10

OMB Number 68-S75124

Date of Admission

** CONFIDENTIAL—All information which would permit identification of the individual will be held in strict confidence, will be used only by persons engaged in and for the purposes of the study, and will not be disclosed or released to others for any other purpose.

Abstract Service I.D. Number ☐☐ 11-12 | Hospital I.D. Number ☐☐☐ 13-15 | Coder I.D. Number ☐ 16

1	2	3	4	5	6
Code:	Do abstract and re-abstract agree?	If no, enter information from abstract	. . . If no, which is correct?	Reason for discrepancy	Reconciled result (if neither is correct)
ADMIT DATE M M D D Y Y ☐☐ ☐☐ ☐☐ 17-22	☐ 1= Yes. ☐ 2= No → 23	M M D D Y Y ☐☐ ☐☐ ☐☐ 24-29	☐ 1 Abstract ☐ 2 Re-abstract ☐ 3 Neither ☐ 4 Indeterminate 30	☐ 1= Clerical ☐ 2= Completeness ☐ 3= Procedural 31	M M D D Y Y ☐☐ ☐☐ ☐☐ 32-37
DISCHARGE DATE M M D D Y Y ☐☐ ☐☐ ☐☐ 38-43	☐ 1= Yes. ☐ 2= No → 44	M M D D Y Y ☐☐ ☐☐ ☐☐ 45-50	☐ 1 Abstract ☐ 2 Re-abstract ☐ 3 Neither ☐ 4 Indeterminate 51	☐ 1= Clerical ☐ 2= Completeness ☐ 3= Procedural 52	M M D D Y Y ☐☐ ☐☐ ☐☐ 53-58
SEX ☐ 1= Male ☐ 2= Female ☐ 3= Other ☐ 9= Not recorded 59	☐ 1= Yes. ☐ 2= No → 60	☐ 1= Male ☐ 2= Female ☐ 3= Other ☐ 4= Not recorded 61	☐ 1 Abstract ☐ 2 Re-abstract ☐ 3 Neither ☐ 4 Indeterminate 62	☐ 1= Clerical ☐ 2= Completeness ☐ 3= Procedural 63	☐ 1= Male ☐ 2= Female ☐ 3= Other ☐ 4= Not recorded 64
DATE OF BIRTH (for HUP and CHDC) M M D D Y Y ☐☐ ☐☐ ☐☐ 65-70	☐ 1= Yes. ☐ 2= No → 71	M M D D Y Y ☐☐ ☐☐ ☐☐ 72-77	☐ 1 Abstract ☐ 2 Re-abstract ☐ 3 Neither ☐ 4 Indeterminate 78	☐ 1= Clerical ☐ 2= Completeness ☐ 3= Procedural 79	/ / CARD 2 M M D D Y Y ☐☐ ☐☐ ☐☐ 11-16
AGE (for PAS and QUEST) ☐☐ ☐ 1= Days ☐ 2= Mos. ☐ 3= Yrs. 17-19	☐ 1= Yes. ☐ 2= No → 20	☐☐ ☐ 1= Days ☐ 2= Mos. ☐ 3= Yrs. 21-23	☐ 1 Abstract ☐ 2 Re-abstract ☐ 3 Neither ☐ 4 Indeterminate 24	☐ 1= Clerical ☐ 2= Completeness ☐ 3= Procedural 25	☐☐ ☐ 1= Days ☐ 2= Mos. ☐ 3= Yrs. 26-28
ANTICIPATED PRINCIPAL SOURCE OF PAYMENT ☐ 1= Medicare ☐ 2= Medicaid/ Medi-Cal ☐ 3= Other ☐ 9= Not recorded 29	☐ 1= Yes. ☐ 2= No → 30	☐ 1= Medicare ☐ 2= Medicaid/ Medi-Cal ☐ 3= Other ☐ 4= Not recorded 31	☐ 1 Abstract ☐ 2 Re-abstract ☐ 3 Neither ☐ 4 Indeterminate 32	☐ 1= Clerical ☐ 2= Completeness ☐ 3= Procedural ☐ 4= Change in status 33	☐ 1= Medicare ☐ 2= Medicaid/ Medi-Cal ☐ 3= Other ☐ 4= Not recorded 34

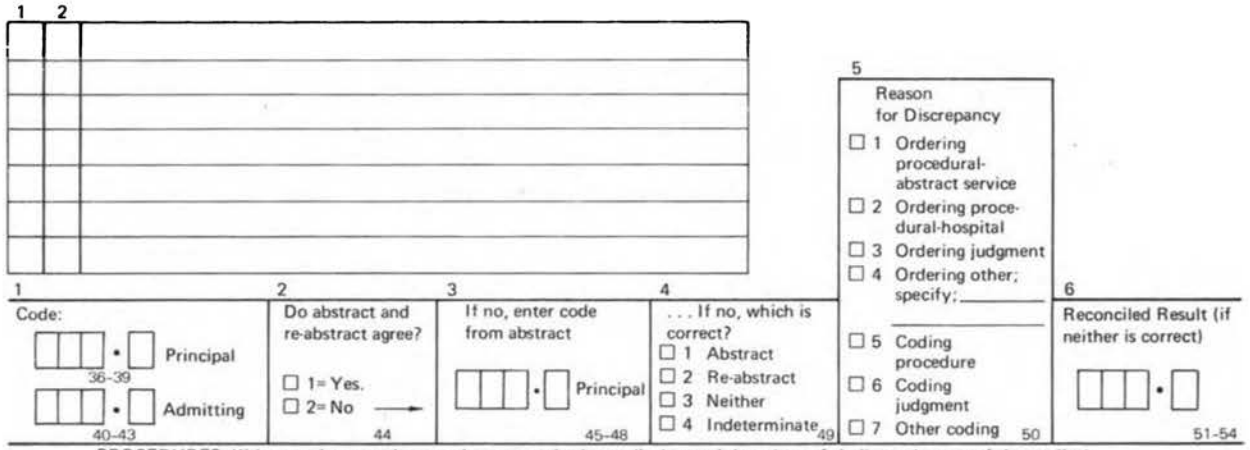
DIAGNOSES: Please check appropriate box to indicate the nomenclature used: 35 ☐ H-ICDA-1 ☐ H-ICDA-2 ☐ ICDA-7 ☐ ICDA-8

(Form is continued on the next page)

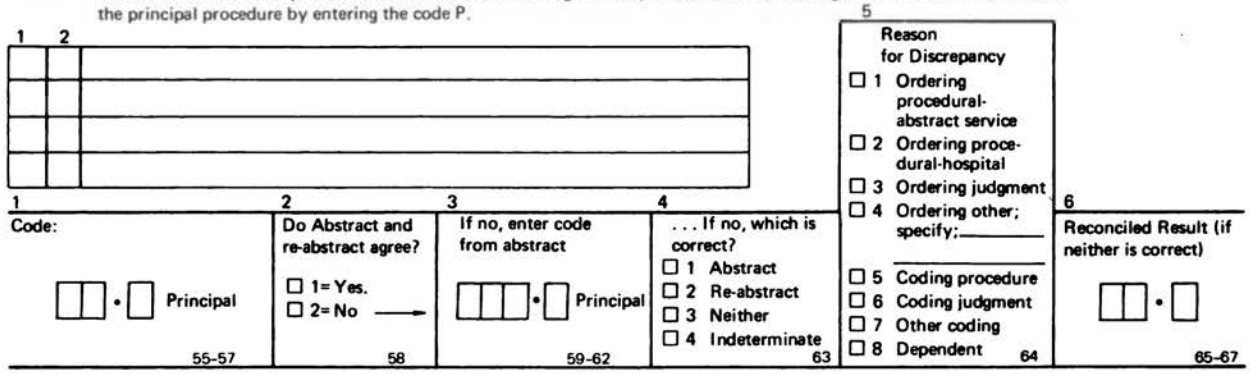
Appendix E

(IOM Re-abstracting Form continued)

Write out all diagnoses as they appear in the medical record. In Column 1 below, indicate the part of the record from which each diagnosis is abstracted: F = Face Sheet; D = Discharge Summary; O = Operation Report; P = Pathology Report; C = Consultation. In Column 2, indicate principal and admitting diagnosis by entering the appropriate codes: P or A.



PROCEDURES: Write out the procedures as they appear in the medical record. In column 1, indicate the part of the medical record from which each procedure was abstracted according to the symbols listed for the diagnoses. In Column 2, indicate the principal procedure by entering the code P.



Appendix E

GENERAL INSTRUCTIONS FOR FIELD TEAM IOM RE-ABSTRACTING FORM

Each field team member will be provided with a master list prepared by the Institute of Medicine of all patient identification numbers and their respective admission (or discharge) dates for each hospital visited. Upon entering the record department, the master list should be compared with the records which have been previously pulled by department personnel in order to ascertain whether all required records are available. At this time, any missing record should be requested from the supervisor of the record department. The master list will be attached to a sealed envelope containing the information provided by the abstract service. The envelope should not be opened until all records have been abstracted.

Before beginning the actual abstract process, the field team member should obtain the MEDICAL RECORD DEPARTMENT QUESTIONNAIRE from the department supervisor and review it for completeness and clarity of response. In particular, items left blank should be noted and an effort to obtain missing information made. This must be done prior to the abstracting process, since it will provide the field team member with background information regarding abstracting procedures within a particular medical record department. At this time the field team member should also review the format of the medical record with the department supervisor to detect any unusual practices which are unique to that hospital. Special attention should be given to the location of information about patient financial status.

For each case to be abstracted, the field team member will be given an IOM re-abstracting form listing the hospital's patient identification number and date of admission (or discharge) for the hospital episode under study. A coder, hospital, and abstract service I.D. number will also be preprinted on the IOM form.

In completing the form, the field team should review the face sheet of the medical record, the discharge summary, operative report, pathology report, X-ray report (if appropriate), consultation notes, laboratory report, and EKG (if appropriate). The form will be used throughout the five steps of the re-abstracting process, as follows:

1. The IOM re-abstracting form will be used to abstract information from the medical record for the specificied admission date. All medical records to be studied at a particular hospital must be re-abstracted before proceeding to the next step. If any records cannot be retrieved (after previously asking the supervisor to locate it, as discussed above), do not substitute. Instead, simply return the incomplete re-abstracting form to the IOM with a note indicating that the record was not available. The items to be re-abstracted

and definitions for each are given in the specific instructions which follow.

The field team member may make changes in the information recorded on the IOM re-abstract form during the initial re-abstracting process. However, after the IOM re-abstract is completed and the comparison and reconciliation with the original abstract has begun, no changes may be made in the re-abstracted information.

2. After all records in a given hospital have been re-abstracted, open the appropriate sealed envelope which will contain copies of the abstracts provided by the abstract service. Compare the information on each newly completed IOM re-abstract with the information from the appropriate abstract. Indicate whether or not the two abstracts agree by checking the appropriate "yes-no" response. If the items do not agree, record the response provided by the abstract service in the appropriate place--column three, which is labeled "If no, enter information from abstract." After all abstracts have been compared, proceed to the next step for cases in which differences are found.

3. In each case, for those items in which there is a difference between the information re-abstracted and that provided by the abstract service, search the medical record to determine which abstract is correct. The correct abstract should be indicated by checking one of the four alternatives in column four: abstract, re-abstract, neither, or indeterminate. "Abstract" refers to the information provided by the abstract service, whereas "re-abstract" refers specifically to the IOM re-abstract form. "Neither" means that both abstracts are in error. An option of "indeterminate" should be used only if, in the opinion of the field team member, there is no "correct" response and the information on either abstract is equally acceptable.

4. After the correct abstract has been identified, refer to the item definitions to determine the reasons for discrepancy, which will fall into three major categories: general, ordering, and coding (see pp. 5-9 of "Specific Instructions"). In the event that both abstracts are in error (i.e., "neither" was checked in the fourth column) the reason for discrepancy should refer to the original abstract provided by the abstract service.

5. After the reason for discrepancy has been determined, and only if neither abstract was judged to be correct in the fourth column, record the correct information in the last column, referred to as "reconciled result."

6. In column 2, which is labeled "Do abstract and re-abstract agree?":

(a) If yes has been checked, information will not be recorded in columns 3, 4, 5, and 6, i.e, leave the rest of that row BLANK.

(b) If no has been checked, information must be entered appropriately in columns 3, 4, and 5. Six is to be completed only if neither abstract nor re-abstract is correct; otherwise leave it BLANK.

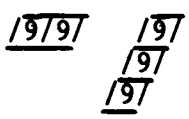
7. The following instructions refer to procedures for handling missing data. They apply first to the problem of the IOM re-abstractor who is unable to complete an item in column 1 due, apparently, to data missing from the medical record. They apply also to the problem of missing data on the abstract provided by the abstract service, which is recorded in column 3 in the reconciliation process if the abstracts do not agree. Procedures for each item are as follows:

(a) Admit and discharge date: if data are missing, enter 9's in the appropriate boxes; be sure to fill each box.

(b) Sex: check box labeled "not recorded" if data are missing.

(c) Date of birth: this item--the entire row--will be left BLANK if age is being recorded as in PAS and QUEST hospitals. However, if date of birth is being sought and the information is missing, enter 9's in the appropriate boxes.

(d) Age: this item--the entire row--will be left BLANK if date of birth is being recorded as in HUP and CHDC hospitals. However, if age is being sought and the information is missing, enter 9's in all five boxes:



(e) Anticipated Principal Source of Payment: check box labeled "not recorded" if data are missing.

(f) Principal Diagnosis: by necessity, there will be no allowance for missing data for principal diagnosis in either the abstract or the IOM re-abstract.

(g) Admitting Diagnosis: as noted in 8(c) of the Specific Instructions: if no admitting diagnosis was easily identified, enter code 999.9 in the appropriate boxes.

(h) Procedures: When coding column 1, 9's should be entered if the IOM field team member finds no significant procedures to be coded from the medical record. These 9's will not indicate that data

are "missing," but rather that there were no procedures to be coded. If the abstract provided by the abstract service notes no coded procedures, column 3 would be completed with 9's in a similar fashion. Of course, column 3 is only filled out if the abstracts do not agree and "no" has been checked in column 2; if "yes" has been checked, the rest of the row is to be left blank.

NOTE: please refer to Items 9 and 10(c)3 of the Specific Instructions for additional guidance on completing this item.

Appendix E

SPECIFIC INSTRUCTIONS FOR IOM RE-ABSTRACTING FORM

In general, the Institute of Medicine field team should abstract medical records using the item definitions of the Uniform Hospital Discharge Data Set, (UHDDS); these definitions are attached. Instances in which the objectives of this study require deviation from the UHDDS definitions are discussed below.

1. The Hospital Patient Identification Number is the number assigned to the patient by the hospital and serves to identify both the hospital record and original medical abstract. Only the number will be used, and in no case will the patient's name be recorded.

2. The Coder Identification Number refers to that number assigned to each member of the IOM field team.

3. The Hospital Identification Number is a three digit number assigned to participating hospitals by the IOM for purposes of this study.

4. Admission Date and Discharge Date are the same as that in the UHDDS except that the hour of admission or discharge will not be recorded. For example, October 1, 1974 should be coded.

M M	D D	Y Y
/1/0/	/0/1/	/7/4/

You will note that the boxes for "day" have been darkened. This is to remind you that the appropriate recording sequence is month, day, year. However, abstracts from the PAS system will array all date information in a day, month, year sequence; the field team will have to pay special attention to this difference in the reconciliation process.

5. Date of Birth refers to the patient's date of birth as recorded on the medical record and will be coded for HUP, CHDC, and some certainty hospitals which will be identified later. When coding for "date of birth," leave the row for "age" information BLANK. This date should be coded as previously explained: month/day/year.

For example: /10/ /01/ /47/ October 1, 1947

6. Age (rather than date of birth) will be abstracted in hospitals using the PAS or QUEST system and in selected certainty hospitals. When coding "age," leave the row for "date of birth" BLANK. Age should be computed to correspond to the admission of interest in this study.

Appendix E

If age is greater than or equal to 98, code as 98. In addition to recording the numbers, check the appropriate box to indicate whether age is recorded in days, months, or years. Only one box may be checked.

7. Anticipated Principal Source of Payment refers to the major source of payment anticipated at time of admission. If both Medicare and Medicaid are indicated, only Medicare should be coded. Payment sources that are not reported as Medicare or Medicaid/Medi-Cal will be coded as Other.

   For this item only, the response alternatives to be used in establishing the reason for discrepancy have one added possibility. If a discrepancy appears to stem from the fact that the anticipated payment differs from actual, and that the actual payment was reported to the abstract service and hence reflected on the abstract, the reason for discrepancy should be recorded as "change in status." For example: a patient is admitted to a hospital claiming that Medicare will cover his hospital bill and this is noted in the chart as the expected source of payment. An abstract is completed on this patient and "Principal Expected Source of Payment" is properly coded as Medicare. Several weeks after discharge, the billing office learns that in fact this patient is covered by Medicaid, rather than Medicare, and a correction notice is sent to Medical Records. In many hospitals, a correction would then be sent to the abstract service, thereby changing the expected source item on the abstract to Medicaid. The IOM re-abstractor, however, has been instructed to code the anticipated source of payment which, for the patient in question, is still Medicare. Note that the abstract provided by the abstract service will state Medicaid due to the later correction. This apparent discrepancy should be accounted for as a "change in status" (column 4), provided there is adequate evidence of this sequence of events in the chart; if documentation is missing, other discrepancy explanations should be considered.

8. (a) All Diagnoses are to be written on the re-abstract form, following the UHDDS definition. If more space is required please continue on the back of the form. In addition, the portion of the record from which each diagnosis was abstracted should be indicated in Column 1, as follows:

   F = face sheet
   D = discharge summary
   C = consultation
   O = operative report
   P = pathology report

   (b) In Column 2, place a "P" next to the diagnosis which is the primary diagnosis, according to the UHDDS definition. If an admitting diagnosis is clearly specified on the face sheet, place

an "A" by the appropriate diagnosis. Do not waste time searching for this item, and include it only if it is readily apparent. In many cases primary and admitting diagnoses will be the same, so that both codes will appear in Column 2--i.e., P/A.

(c) Enter the ICDA-8 or H-ICDA code for the principal and admitting diagnoses in the boxes provided. If principal and admitting diagnoses are the same, insert the number twice. If no admitting diagnosis was identified, enter code 999.9 in the approxpriate boxes.

(d) All coding outlined in (c) above should be done in ICDA-8 except for PAS hospitals, which require H-ICDA-2. Exceptions to this rule for certain hospitals will be indicated to individual field team members where appropriate. If the code manual indicates only a three digit code for a given diagnosis--such as the H-ICDA-2 code for hypertrophy of tonsils and adenoids which is 550--enter a "0" in the decimal position, i.e., 500.0.

9. (a) All Procedures are to be written on the re-abstract form. In addition, the portion of the record from which each procedure was abstracted should be indicated in Column 1 according to the symbols listed in 8(a) above.

(b) In Column 2, place a "P" next to the procedure which is the principal procedure according to the UHDDS definition. The field team will have to exercise some discretion in assigning a principal procedure when only a minor one--such as "manual arts therapy"--has been noted in the chart. Do not be overzealous in coding, but on the other hand procedures of clear significance should definitely be recorded.

(c) Enter the appropriate code for the principal procedure in the boxes provided.

(d) Do not code as a principal procedure any of the E or Y codes in either ICDA-8 or H-ICDA-2. However, the A and R codes in ICDA-8 are permitted as a principal procedure for the purposes of this study. If this instruction causes any problems in coding the abstract, please write a note on the back of the form explaining the difficulty.

(e) When no procedure is to be coded at all, 9's should be entered in all the appropriate boxes. The General Instructions provide further information for coding missing data. See 7(h).

Appendix E

10. The following codes will be used in explaining a reason for a discrepancy (see the Column 5 of the abstract form):

A. General: These categories apply only to the following items in the data set: admission date, discharge date, sex, date of birth or age, and anticipated principal source of payment.

1. Clerical: Discrepancy attributable to human error, mistakes of a particular clerk, errors in transcribing numbers, etc. (Example: obvious transposing of numbers on admission date or age.)

2. Completeness: Incomplete information or inaccurate information on the abstract or re-abstract due to an inadequate review of the chart. (Example: item missing from the admitting sheet, but clearly stated in the discharge summary.)

3. Procedural: Discrepancy resulting from a particular identifiable practice in the hospital, which produces a systematic difference between the original abstract and the IOM re-abstract. (Example: routinely and consistently different or misinterpreted definitions of the data set such as routinely inverting day and month in admission date.)

NOTE: for anticipated principal source of payment, please remember the additional reason (i.e., change in status), discussed on pp. 2-3 of the Specific Instructions.

B. Ordering Discrepancies: Ordering discrepancies, as well as coding coding discrepancies (see C, below), apply only to diagnosis and procedure items. However, the possibility of an ordering discrepancy must be considered and eliminated before considering the possibility of a coding discrepancy.

An ordering discrepancy will be used to reflect an inconsistency between the original abstract and IOM re-abstract which stems from uncertainty over whether a diagnosis or procedure should be regarded as "principal" or one of the "other" diagnoses or procedures. Specific types of ordering discrepancies and their codes follow:

1. Ordering Procedural - Abstract Service: Discrepancy in ordering of principal diagnosis and/or procedure because of a difference between definitions of the abstract service and UHDDS definitions. (Example: A patient is admitted for an open fracture reduction and, while on the operating table, suffers an acute MI which keeps him in the hospital three months. The UHDDS definition for principal diagnosis would be fracture because, as the definition requires, fracture is the diagnosis best explaining cause of admission. If, however, the hospital participates in QUEST, acute

MI probably would be the principal diagnosis recorded, since QUEST loosely defines principal diagnosis as "the most important diagnosis" (best accounting for the days of care).

2. Ordering Procedural - Hospital: Discrepancy in selecting of principal diagnosis and/or procedure because of systematic differences between the definitions within a given hospital and the definitions of both the abstract service and UHDDS definitions. (Example: coding the principal diagnosis routinely as the first listed by the physician rather than selecting principal diagnosis or procedure by a clear definition after studying the full record.)

3. Ordering Judgment: Discrepancy in selection of principal diagnosis or procedure which does not represent actual error, but rather, an honest difference of opinion in interpreting the medical record.

4. Ordering Other: A discrepancy in ordering of principal diagnosis and/or procedure other than the above. If this response is used, please write a note to indicate the reason.

C. Coding Discrepancies: These would apply only to the actual coding of the principal diagnosis and the principal procedure, after the possibility of an ordering discrepancy has been eliminated. They would not apply to other elements of the basic data set.

1. Coding Procedural: Discrepancy caused by routine and systematic misuse or misunderstanding of the coding system, resulting in discrepancy. (Examples: reliance on index without reference to tabular listings, failure to heed inclusion and/or exclusion advice from tabular listings, double coding when single coding is indicated, single coding when double coding is required, etc.)

2. Coding Judgment: Discrepancy caused by absence of complete word-for-word correspondence between the recording of the diagnosis or procedure and the wording in the H-ICDA or ICDA-8 manuals. That is, an honest difference of opinion over the correct code when it is not clear from the coding manual what the numbers should be. (Example: diagnosis listed as recurrent and it is unclear whether "acute" or "chronic" is actually the more appropriate qualifier for coding purposes and these are the only two options available.)

3. Coding other: Discrepancy in coding not due to either of the above. If this response is used, please write a note on the back of the form explaining the problem.

Appendix E

One discrepancy explanation in particular should be included in this category: a difference in the degree of detail in procedure coding. For example, the field team and the hospital may have a different sense of how significant a procedure must be to be coded as a principal one. Accordingly, it may happen that the abstract will have coded a very insignificant procedure as "principal" while the IOM abstract will have noted no principal procedure--or vice versa. When this discrepancy is clearly explained by a difference in how "insignificant" the principal procedure is allowed to be, check this discrepancy explanation and write a note on the back of the form explaining the situation. You will notice that an extra box has been provided in Column 3 of the procedures coding section. This is to allow you to code a four digit procedure code (such as Y 10 o 1 appearing on an abstract provided by the abstract service should this be necessary in the reconciliation process. In Column 6 only three boxes / / / / are provided as we have decided that the "correct" principal procedure will by definition not be a four digit code.

D. Dependent Discrepancy: Dependent discrepancies apply only to procedures. This reason will be used to reflect a discrepancy which results from a prior discrepancy in a related item. Usually, this situation will occur only when an earlier discrepancy in selection of the principal diagnosis results in a dependent discrepancy in selecting the principal procedure.

Appendix E

UHDDS DEFINITIONS FOR IOM STUDY*

1 Admission and Discharge Dates.

a. "Admission Date includes month, day, (and) year...of admission.

b. Discharge Date includes month, day, and year of discharge."

2. Sex. Male or Female.

3. Date of Birth. Month, day, and year of birth. (Age stated in days, months, or years is not part of UHDDS. However, because PAS and QUEST record age on their abstracts, use the definition described in item 6 of the Specific Instructions.)

4. A. Principal Diagnosis.

"The condition established after study to be chiefly responsible for occasioning the admission of the patient to the hospital for care."

B. Other Diagnoses.

"All conditions that coexist at the time of admission, or develop subsequently, which affect the treatment received and/or the length of stay. Diagnoses that relate to an earlier episode which have no bearing on this hospital stay are to be excluded."

5. Procedures.

"All procedures performed in operating rooms are to be reported... In addition to these procedures, all other significant procedures are to be recorded. A significant procedure is one which carries an operative or anesthetic risk or requires highly trained personnel or special facilities or equipment. Some examples of such procedures are cardiocatheterization, angiography, endoscopy, and super-voltage radiation therapy.

"When more than one procedure is recorded the principal procedure is to be designated. In determining which of several procedures is the principal, the following criteria apply:

---

*U. S. Department of Health, Education, and Welfare, Public Health Service, Health Resources Administration, Uniform Hospital Abstract: Minimum Basic Data Set, Vital and Health Statistics, Series 4: Documents and Committee Reports No. 14, DHEW Publication No. (HRA) 75-1451, August 1974.

Appendix E

(1) The principal procedure is one which was performed for definitive treatment rather than one performed for diagnostic or exploratory purposes, or was necessary to take care of a complication.

(2) The principal procedure is that procedure most related to the principal diagnosis."

Hospital I.D. Number

Appendix F

OMB Number 68-S75124

MEDICAL RECORD DEPARTMENT QUESTIONNAIRE TO BE USED IN THE INSTITUTE OF MEDICINE STUDY OF THE RELIABILITY OF ABSTRACTED MEDICAL RECORDS INFORMATION

PLEASE NOTE: In completing this questionnaire, please be assured that all information which would permit identification of individual hospitals or patients will be held in strict confidence, will be used only by persons engaged in and for the purposes of the study, and will not be disclosed or released to others for any other purpose. The information collected will be used only to prepare statistical summaries, and the questionnaires containing raw data will be destroyed after the analysis is completed.

GENERAL INSTRUCTIONS: Please record the requested information by entering the appropriate numerical value, check (✓) mark, or brief description.

If a question is *not applicable* to your particular hospital, please write in "not applicable" in the left hand margin next to that question. If you have any questions about a particular item, you may discuss them with a member of the IOM field team during the site visit, or you may call Sarah Brown, M.P.H., Professional Associate, Institute of Medicine. Please call her collect at (202) 389-6978.

ALL INFORMATION IN THIS QUESTIONNAIRE SPECIFICALLY REFERS TO CALENDAR YEAR 1974.

I. MEDICAL RECORD DEPARTMENT PERSONNEL

A. On the average, how many persons (full-time equivalents) were involved in the completion of the medical record abstract form during 1974 (excluding the supervisor)?

☐☐ Enter appropriate number(s).

B. 1. Of those persons, how many were:

☐☐ ART ☐☐ RN

☐☐ RRA ☐☐ *Other; specify* _____

2. and the supervisor of the record abstracting process?

RRA	ART	RN	other
☐	☐	☐	☐

If other, please specify _____

II. TRAINING ACTIVITIES FOR THE ABSTRACTING PROCESS

A. Did your medical record staff participate in any training programs (in 1974) specifically directed at abstracting and coding information from medical records? [Please check appropriate box(es).]

☐ 1. No
☐ 2. Yes; program offered by abstract service.
☐ 3. Yes; program other than that offered by abstract service.

Please describe such program(s) _____

_____

ALL INFORMATION SPECIFICALLY REFERS TO CALENDAR YEAR 1974

If your medical record staff did participate in any training programs (in 1974):

4. How frequently did the *supervisor(s)* attend such programs during 1974 per person?_____

5. How frequently were the training programs attended by the *coders* during 1974 per person?_____

B. When a new abstractor was hired in 1974, how was he/she trained into your abstracting system?_____

C. In general, how often was there contact between your medical record department and abstract service during 1974? We are interested in *any* communication, regardless of how it occurred or by whom it was initiated—for example, phone calls to discuss problem abstracts, site visits to your hospital by the abstract service, discussions over report formats and so forth. (Please check appropriate box.)

☐ (1) Never
☐ (2) Seldom
☐ (3) Monthly
☐ (4) More than once per month
☐ (5) Not certain

III. ABSTRACTING PROCESS DURING CALENDAR YEAR 1974.

A. For the average abstract, who actually entered codes for diagnoses and procedures on the medical records abstract?

ART ☐ RRA ☐ RN ☐ other ☐ *If other, please specify* _____

B. When were the codes for diagnosis and procedure entered on the abstract?

☐ (1) Before the medical record was completed
☐ (2) After the medical record was completed

C. In general, how long after a patient's discharge was the final abstract completed?

☐ Less than two weeks
☐ Two to four weeks
☐ Five to eight weeks
☐ Nine to 12 weeks
☐ 13 weeks or more

D. Which parts of the medical record were routinely referred to in order to complete the sections of the abstract regarding diagnoses and procedures? (Please check appropriate box.)

☐ (1) Face sheet only
☐ (2) Face sheet and only selected portions of the record; (*please specify*)_____
☐ (3) All parts of the medical record
☐ (4) No routine or standard practice
☐ (5) Not certain

E. Please estimate how many medical records were abstracted per day per person: _____

ALL INFORMATION SPECIFICALLY REFERS TO CALENDAR YEAR 1974

F. If a patient's actual source of payment turned out to be different from that initially recorded on the abstract as the anticipated source, was a correction statement for that record sent to the abstract service? (Please check the appropriate box.)

☐ (1) Yes ☐ (2) No ☐ (3) Not certain

Was this change in payment status recorded in the medical record? (Please check the appropriate box.)

☐ (1) Yes ☐ (2) No ☐ (3) Not certain

G. In 1974, was *every* abstract reviewed by a supervisor for accuracy? (Please check the appropriate box.)

☐ (1) Yes ☐ (2) No ☐ (3) Not certain

If no, indicate whether any of the following specific procedures were used for *selected* review?

☐ (1) A systematic or random sampling procedure
☐ (2) A review of problem cases only
☐ (3) Other; please specify _____

IV. USES OF THE REPORTS OF ABSTRACTED INFORMATION PREPARED BY THE ABSTRACT SERVICE:

Of those reports provided by the abstract service, please *check* the appropriate box to indicate which reports are regularly *reviewed and used* by the specified groups within the hospital. ("Yes" indicates that the report was regularly reviewed and used, "No" indicates that it was not regularly reviewed or used, and "NC" indicates that you are not certain of this information.)

	Medical Record Department			Hospital Administration			Medical Staff		
	Yes	No	NC	Yes	No	NC	Yes	No	NC
Discharge Summaries	☐	☐	☐	☐	☐	☐	☐	☐	☐
	Yes	No	NC	Yes	No	NC	Yes	No	NC
Diagnostic Summaries	☐	☐	☐	☐	☐	☐	☐	☐	☐
Operation and Procedures Summaries	Yes ☐	No ☐	NC ☐	Yes ☐	No ☐	NC ☐	Yes ☐	No ☐	NC ☐
	Yes	No	NC	Yes	No	NC	Yes	No	NC
Mortality Summaries	☐	☐	☐	☐	☐	☐	☐	☐	☐
	Yes	No	NC	Yes	No	NC	Yes	No	NC
Physician Summaries	☐	☐	☐	☐	☐	☐	☐	☐	☐
	Yes	No	NC	Yes	No	NC	Yes	No	NC
Other	☐	☐	☐	☐	☐	☐	☐	☐	☐

Appendix F

ALL INFORMATION SPECIFICALLY REFERS TO CALENDAR YEAR 1974

V. IN THE ABSTRACTING PROCESS, HOW WERE THE FOLLOWING ITEMS DEFINED BY THE ABSTRACTORS IN YOUR HOSPITAL IN 1974?

Principal or Primary diagnosis _____

_____

_____

Other diagnoses _____

_____

_____

Principal or Primary Procedure _____

_____

_____

Other Procedures _____

_____

_____

Length of Stay _____

_____

_____

_____

Signature of Person Completing Questionnaire

_____

Date

Appendix G

RELIABILITY OF FIELD WORK

Despite the care with which the field team was selected and trained and the thorough editing of data, an independent assessment of the reliability of the field team's work was performed. A consultant who had assisted in training the field team independently "re-re-abstracted" a subsample of records and compared her results to those initially compiled by the field team. This Appendix presents the methods and findings from that activity.

Methods

Six of the original 50 hospitals were chosen for inclusion in the independent assessment. The hospitals were first divided into three groups, depending on which field team member had visited each facility. The hospitals in each group were categorized according to whether they were visited during the first or last half of the abstractor's field work. From each of the resulting six groups, one hospital was chosen at random for inclusion in the study. Thus, all 50 hospitals had a possibility of selection. All four abstracting services were represented in the final six, although no hospital with an internal data system was selected. All six hospitals agreed to a second site visit by the consultant.

In each hospital approximately 35 records from the original sample of records abstracted in that facility were selected for review. These were chosen systematically and in proportion to: those records in which no discrepancy had been found between the original abstract supplied by the abstracting service and the IOM re-abstract; and those in which one or more discrepancies had been found between the abstract and re-abstract. One hundred ninety-seven records were available for analysis. Each new reabstract was assigned a weight to reflect the probability of selection of both abstract and hospital. The results can be generalized to the unweighted total of all IOM re-abstracts.

The forms and instructions used in the assessment are found at the end of this Appendix. The consultant did not know which member of the field team had done the initial abstracting or whether any discrepancies had initially been detected. After completing the independent abstracting, the consultant reviewed a form indicating for each item the code found on the original abstract and the code assigned by the IOM re-abstractor. Any discrepancies were reconciled in accord with a process described in the instructions. The goals of this process were to check whether the IOM re-abstractor made a reasonable judgment about the accuracy of the original codes and whether the field team's assessment of the reasons for discrepancy was plausible.

Appendix G

The independent "re-re-abstracting" does not answer definitively the question of the reliability of the field team's work. It simply lends added credibility to the initial judgments. Nevertheless, in a situation where the concept of data accuracy is tenuous at best (for many abstract items, there is no clear "right answer") the independent assessment was intended to determine whether the basic study data were reliable and whether the initial re-abstracting was based on sound judgment and careful attention to the research protocol.

Analysis

The analysis involved a comparison of three sources of data: that generated by the abstract service, the IOM re-abstractors, and the consultant. Special attention was given to determining whether the field team and consultant initially abstracted the medical record in a similar manner and, where there were differences, whether they agreed on the correct source of data and the reasons for discrepancies.

Table 1 shows that the general pattern of better reliability for items other than diagnosis and procedure (reported in Chapter 4) was also found in the assessment of the field work. The "no discrepancy" figures under-estimate the data reliability, however, because they do not include those cases where there was a discrepancy between the field team member and the consultant, but the consultant agreed with the re-abstractor's determination of correct data source. Often there was difficulty in determining a "correct" data source, particularly for age or date of birth, diagnoses, and procedures.

Table 1. Comparison of Data Abstracted by the Consultant and the Field Team (weighted percent)

		Agreement on correct data source where a discrepancy exists		
	No discrepancy	Agree	Disagree	Total
Admit date	99.2	0.3	0.5	100.0%
Discharge date	99.5	-	0.5	100.0
Age/date of birth	91.4	0.4	8.1	99.9
Sex	100.0	-	-	100.0
Expected principal source of payment	99.5	0.5	-	100.0
Principal Diagnosis*	86.1	2.3	11.6	100.0
Principal Procedure	87.0	8.2	4.8	100.0

NOTE: Unweighted N = 196 abstracts
*Compared to four digits

The instructions to both the field team and the consultant were to abstract directly from the medical record either date of birth or age, depending on the abstract service to which the hospital subscribed. The consultant

apparently did not follow these instructions consistently, which at least partially accounts for the lower reliability of data on age and (see Table 2) the higher use of the "Indeterminate" option for correct data source. As an example, if a hospital participated in the PAS system, the instructions were to code age directly from the medical record. In many cases the consultant would abstract the date of birth and compute age in the column of the assessment form, sometimes producing an age which was different from that which appeared in the medical record and on both the original abstract and the IOM re-abstract. In these cases, the consultant usually decided that the correct source of data was "Indeterminate." Although the consultant's work points out inconsistencies within the medical record, it was not performed in accord with the re-abstracting instructions and, therefore, may introduce an erroneous impression about the reliability of the field work.

Table 2. Correct Source of Data, According to the Assessment, When the Field Work and Assessment Disagreed (weighted percent)

Data item	Original abstract and assessment	Assessment	Indeterminate	Total
Age/date of birth	-	12.6	87.4	100.0%
Diagnosis	18.9	56.8	24.3	100.0
Procedure	8.4	29.7	61.9	100.0

The information on diagnoses and procedures supports the statements in Chapter 4 about the difficulty of determining these items with certainty. For 11.6 percent of the abstracts reviewed, the field team and consultant disagreed about the correct source of data for principal diagnosis. For 56.8 percent of those cases the consultant decided that her assessment was correct (see Table 2). (For 74.3 percent of those cases the original abstract and IOM re-abstract had agreed on a different principal diagnosis.) For 24.3 percent of the abstracts the consultant decided that the correct data source for diagnostic information could not be determined.

For 8.2 percent of the abstracts, the consultant and field team initially disagreed on the choice of principal procedure, but after completing the reconciliation process, they agreed on the correct source of data. This usually reflected a common selection of the "Indeterminate" option for correct data source. In most cases the discrepancy stemmed from the absence of definitive guidelines for deciding which procedures were important enough to warrant coding and the subsequent selection of the "Indeterminate" option. Of those abstracts where the consultant and field team disagreed on the correct data source, and the consultant decided that she was correct or the correct data source was "Indeterminate," about half were cases in which the original and IOM abstracts agreed.

Appendix G

Interpretation of these data is difficult. The sample size is very small. The weighting factors may have introduced distortions. Because of budget and time constraints, only one day per hospital was available for assessing the field work, which may have limited the thoroughness of the review. The problem in determining age or date of birth was noted above. A comparison of the amount of diagnostic and procedural information recorded in the initial field work and the assessment, for the same medical records, reveals considerable variability. (The instructions were to record all information on diagnoses and procedures before coding those that were principal.) This may have influenced the selection of principal diagnosis and procedure and the subsequent frequency of discrepancies between the initial field work and its assessment. These problems notwithstanding, additional analyses were made to determine the extent to which diagnostic grouping and the degree of coding refinement may have influenced the assessment of the field work.

As shown in Table 3, there is greater agreement between the initial field work and its assessment for all diagnoses combined and for the target diagnoses when diagnostic codes are compared only to three digits. After the abstracts for which there were no discrepancies are combined with those for which the field team and consultant agreed on correct data source, the level of agreement reaches 90 percent.

Table 3. Comparison of Data Abstracted by the Consultant and the Field Team at Different Levels of Coding Refinement (weighted percent)

	No discrepancy	Agreement on correct data source where a discrepancy exists		Total
		Agree	Disagree	
Diagnosis coded to four digits				
All diagnoses	86.1	2.3	11.6	100.0
Target	85.4	2.5	12.1	100.0
Other	87.5	2.4	10.1	100.0
Satellite	84.2	-	15.8	100.0
Diagnosis coded to three digits				
All diagnoses	88.1	2.3	9.6	100.0
Target	88.0	2.5	9.5	99.0
Other	87.5	2.4	10.1	100.0
Satellite	84.2	-	15.8	100.0

Appendix G

Diagnostic-specific discrepancy rates are not presented because the numbers of abstracts per diagnosis are so small. However, Table 4 shows the distribution of discrepancies between the field work and assessment by diagnosis. The assessment confirms the finding that reliability is lowest for chronic ischemic heart disease.

Table 4. Distribution of Discrepancies Between the Field Work and Assessment by Diagnosis

Diagnosis	Weighted percent of the total number of discrepancies for each diagnosis	Unweighted number of abstracts with discrepancies	Weighted percent of the total number of abstracts in the assessment for each diagnosis
Cerebrovascular disease	16.2	3	4.9
Chronic ischemic heart disease	28.2	7	6.9
Acute myocardial infarction	11.6	3	9.7
Diabetes	4.6	1	6.2
Carcinoma of the breast	-	-	5.1
End stage renal disease	5.9	2	5.2
Cholelithiasis/ cholecystitis	9.7	-	5.2
Hernia without obstruction	-	-	5.3
Delivery	-	-	3.5
Tonsils and adenoids	3.7	1	4.5
Fracture of neck of femur	-	-	7.5
Displacement of disc	-	-	5.2
Cataract	-	-	6.6
Neuroses	3.7	1	4.2
Other	13.1	4	15.0
Satellite	3.3	1	5.0
Total	100.0%	23	100.0%

Note: Unweighted N = 23 abstracts

The extent to which there was agreement between the field work and its assessment on reasons selected to explain discrepancies between the original abstract and the IOM re-abstract was also examined. The possibly subjective nature of this assessment and the need to apply judgment in determining which of seven or eight options was appropriate were noted in Chapter 3. Because of

the sizable number of options and the small number of abstracts reviewed in the assessment of the field work, only the general categories of reasons for discrepancy are considered here. Table 5 shows the extent to which the field team and consultant agreed that the reasons for discrepancy between the original and re-abstract stemmed from difficulties in deciding which diagnosis or procedure was principal (an ordering discrepancy) or from errors in assigning the proper diagnostic or procedural code number (a coding discrepancy). "General agreement" means that both were within the same general category, but may have selected different specific reasons. For example, one may have decided that the reason was "Coding-Procedural," while the other selected "Coding-Judgment." Where there was complete agreement, they selected the same general and specific reason. The field team and consultant were able to agree on 77 percent of the reasons to explain diagnostic discrepancies between the original and re-abstract (compared to four digits) and 89 percent of the reasons to explain discrepancies on principal procedure.

Table 5. Agreement Between the Consultant and Field Team on Reasons for Discrepancies when the Original Abstract and IOM Re-abstract Disagreed (weighted percent)

	Diagnosis (4 digit)	Procedure
Complete agreement	30.5	63.4
General agreement	47.0	26.0
Complete disagreement	22.5	10.6
Total	100.0	100.0
(Unweighted N)	(53)	(51)

Summary

The reliability of the Institute of Medicine field work was assessed by comparing data provided by the abstract service, the re-abstractor who performed the field work, and the consultant who performed the assessment.

The results of the assessment confirm both the findings and the caveats reported in Chapter 4. Data were best for information on hospital admission date, discharge date, sex, and anticipated principal source of payment. Some difficulty was encountered in conclusively determining which diagnosis or procedure should be regarded as "principal." The reliability of diagnostic data varied, depending on the level of coding refinement and the specific diagnosis. In some cases, the consultant decided that the correct source of data could not be determined. Reliability was lowest for chronic ischemic heart disease. Overall, the levels of agreement reached about 90 percent for diagnosis and 95 percent for procedures. Where differences were found between the original abstract and IOM re-abstract, the field team and consultant generally agreed on the reasons for discrepancies on principal diagnosis about 77 percent of the time and on principal procedure about 89 percent of the time.

These findings should be tempered by the limitations of the assessment. The sample size was very small. The weights may have introduced distortions. The time available for the assessment was limited, which may have hampered the thoroughness of the medical record review. There is some indication that the assessment may not have been conducted with strict adherence to the instructions followed by the field team.

Appendix G

Instructions for the Abstract to be Used in Assessing the Reliability of IOM Field Work

GENERAL NOTE:

These instructions are to be used in completing the abstract which will be used in assessing the reliability of the data collected earlier by IOM re-abstractors. Therefore, they require adherence to the definitions used by the IOM re-abstractors in initially coding the information on the abstract form. Specifically, this requires use of the UHDDS definitions, supplemented by the general and specific instructions for the initial IOM re-abstracting form, including correct data source (abstract, re-abstract, neither, and indeterminate) and reasons for discrepancy (clerical, completeness, procedural, change in status, ordering procedural-abstract service, ordering procedural-hospital, ordering judgment, ordering other, coding procedural, coding judgment, coding other, and coding dependent).

Since the general purpose is to determine the reliability of the field work, the abstract resembles the original IOM re-abstracting form, except that the new abstract consists of two pages. The first page will be used to code the new information independently abstracted from the medical record; the second sheet will reflect the decision-making process used by the field team to determine reliability.

Specific Instructions

The specific procedures for this assessment will be very similar to the initial re-abstracting process. The consultant will receive a master list prepared by the IOM of all patient identification numbers and their respective admission (or discharge) dates for each hospital. Upon entering the record department, the master list should be compared with the records which have been previously retrieved by department personnel in order to ascertain whether all required records are available. Any missing record should then be requested from the supervisor of the record department, but substitutions may not be made. At this time the consultant should also review with the department supervisor any unusual abstracting practices within the hospital which may have been revealed by the consultant's prior review of the completed medical record department questionnaire for that hospital.

For each case to be abstracted, an IOM re-abstracting assessment form will be provided that lists the hospital's patient identification number and date of admission (or discharge) for the hospital episode under study. A hospital and abstract service I.D. number will also be pre-printed on the form.

The first page of this IOM re-abstracting assessment form will be used in coding the information from the medical record which includes: the admission date, the discharge date, the sex of the patient, the date of birth (if it is a HUP or CHDC abstract), the age (if it is a PAS or QUEST abstract), the anticipated principal source of payment, the diagnoses, and the procedures.

Appendix G

Each record included on the master list must be independently abstracted and the information coded on the first page of the re-abstracting assessment forms before turning to the second pages.

The second page is designed to facilitate the consultant's decision-making process regarding the reliability of the original field work and the reasons for any discrepancies which are detected. The information on the second page is described below by first briefly reviewing the content of each of the seven columns and then by considering the steps of the decision-making process.

The first column on the second page contains information from the IOM re-abstracting form; the second column contains information provided by the abstract service. Both are pre-coded. The third column will be used in determining the correct data source--i.e. abstract, re-abstract, assessment (referring to the IOM re-abstracting assessment form), neither or indeterminate. The fourth column, which asks if the information from the IOM re-abstract agrees with that provided by the abstract service, will be used as a cue in starting the reconciliation process. In determining this, the consultant will compare the data in the first two columns. If there is a discrepancy between the IOM re-abstracting data and the data provided by the abstract service, then the reason for discrepancy should be noted in the 5th column. The reason for discrepancy should reflect the consultant's judgment on the first page of this assessment abstract. However, if the abstract and re-abstract data agree, but the validation abstract does not and is determined to be correct, then the reason for discrepancy must be noted in the 6th column. In the 7th column, the reconciled result will be entered for those cases where all three (abstract, re-abstract, and "validation abstract") are determined to be incorrect.

The steps involved in completing the second page are described below:

1. The consultant compares the data coded on the first page with the data in the first and second columns of the second page. After comparing these three sets of data, the consultant makes a determination as to the correct data source in column 3.

2. After making this determination, the consultant starts the reconiliation process by checking if the data in columns 1 and 2 agree and checking the appropriate box in column 4.

3. If the information in column 1 and column 2, disagree then the consultant is required to determine the reason for discrepancy. For some abstracts, this marks the end of the procedure.

4. However, if the information in columns 1 and 2 agree and the validation does not, the consultant will proceed to the 6th column and determine the reason for discrepancy.

Appendix G

5. When information from the abstract service, the re-abstracting form, and the assessment abstract are determined to be incorrect, the consultant must code the information determined to be the reconciled result in the 7th column.

The following examples will better explain the mechanics of this assessment procedure. The three examples considered are limited to the diagnosis item. The diagnoses of concern are hernia with an obstruction and hernia without an obstruction.

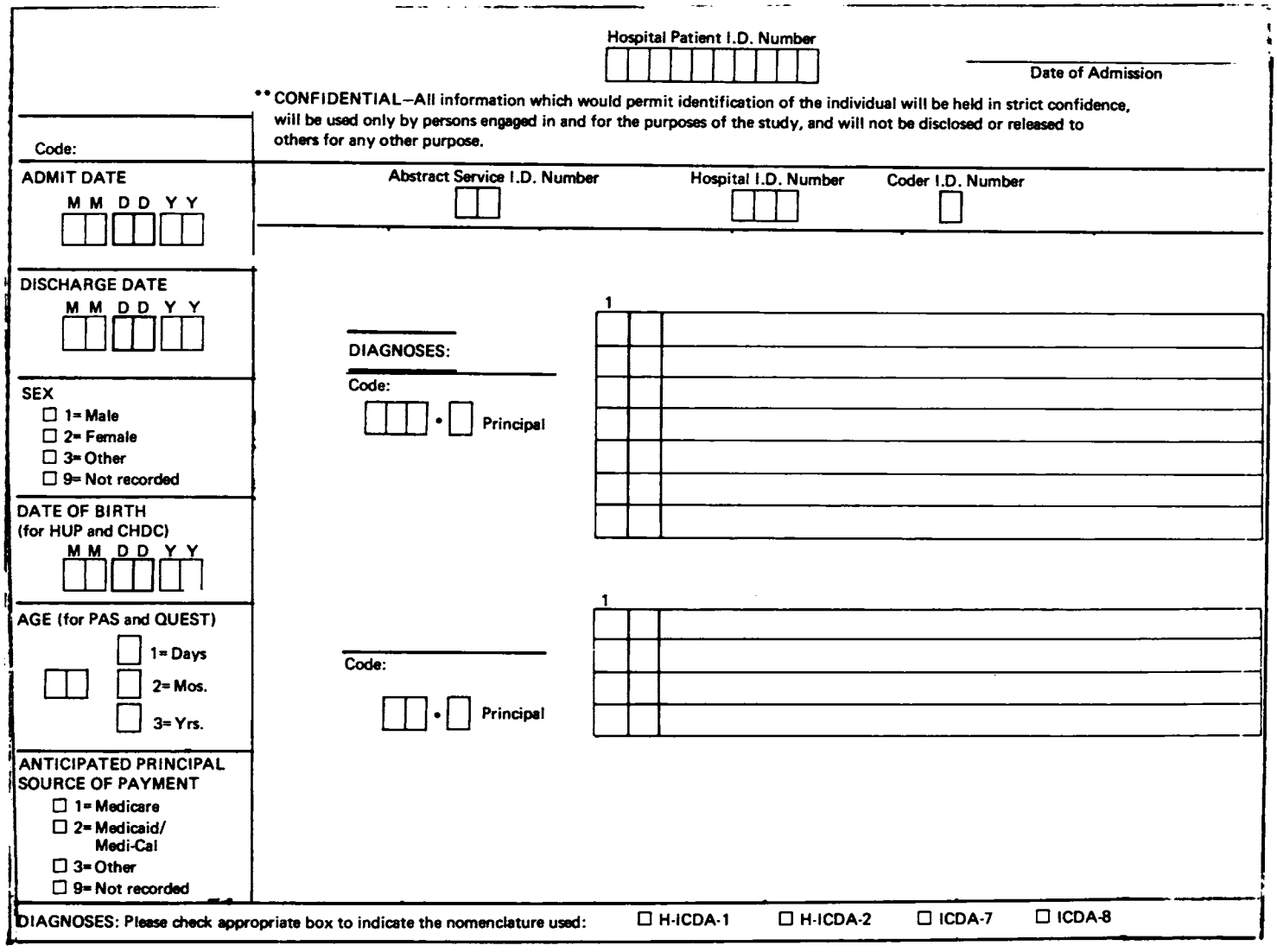
1. In the first example, the consultant has coded hernia with obstruction as principal diagnosis. The information from the IOM re-abstracting form and the abstract service also indicates hernia with obstruction as the principal diagnosis. In this case, data from all three sources agree and all are determined to be correct. For this example, the consultant checks in the third column that abstract, re-abstract, and validation are all correct. After checking "yes" in the 4th column to indicate that the information from the IOM re-abstracting form agrees with the information from the abstract service, this assessment precedure ends.

2. In the second example, hernia with obstruction was determined by the consultant to be the principal diagnosis. However, hernia without obstruction was determined to be the principal diagnosis by the IOM re-abstractor and the abstract service. Again, the consultant will be asked to determine the correct data source. If it is determined that both the abstract and re-abstract are correct, the consultant will proceed to the 4th column and check that the information from the re-abstract agrees with that of the abstract service. For this example, this assessment procedure ends. However, if the assessment abstract was determined to be correct, the procedure would continue in column 6 where the consultant would check the reason for discrepancy.

3. In the third example, the principal diagnosis was determined by the consultant and the IOM re-abstractor to be hernia with obstruction and by the abstract service to be hernia without obstruction. Again, the procedure includes having the consultant determine the correct data source(s). If the assessment and the re-abstract data are correct, the consultant would make note of these in column 3. In the 4th column, the consultant would check that the IOM re-abstracting data and the abstract data do not agree. The consultant would then proceed to the 5th column and determine the reason for discrepancy between the IOM re-abstract information and the abstract service. In this last example, if all three data sources are incorrect, then the diagnoses which was determined to be correct would be entered in column 7 (the reconciled result section of the abstract).

To assist the consultant in this assessment process, a few IOM re-abstracting assessment forms will be completed to serve as guides.

INDEPENDENT ASSESSMENT RE-ABSTRACTING FORM



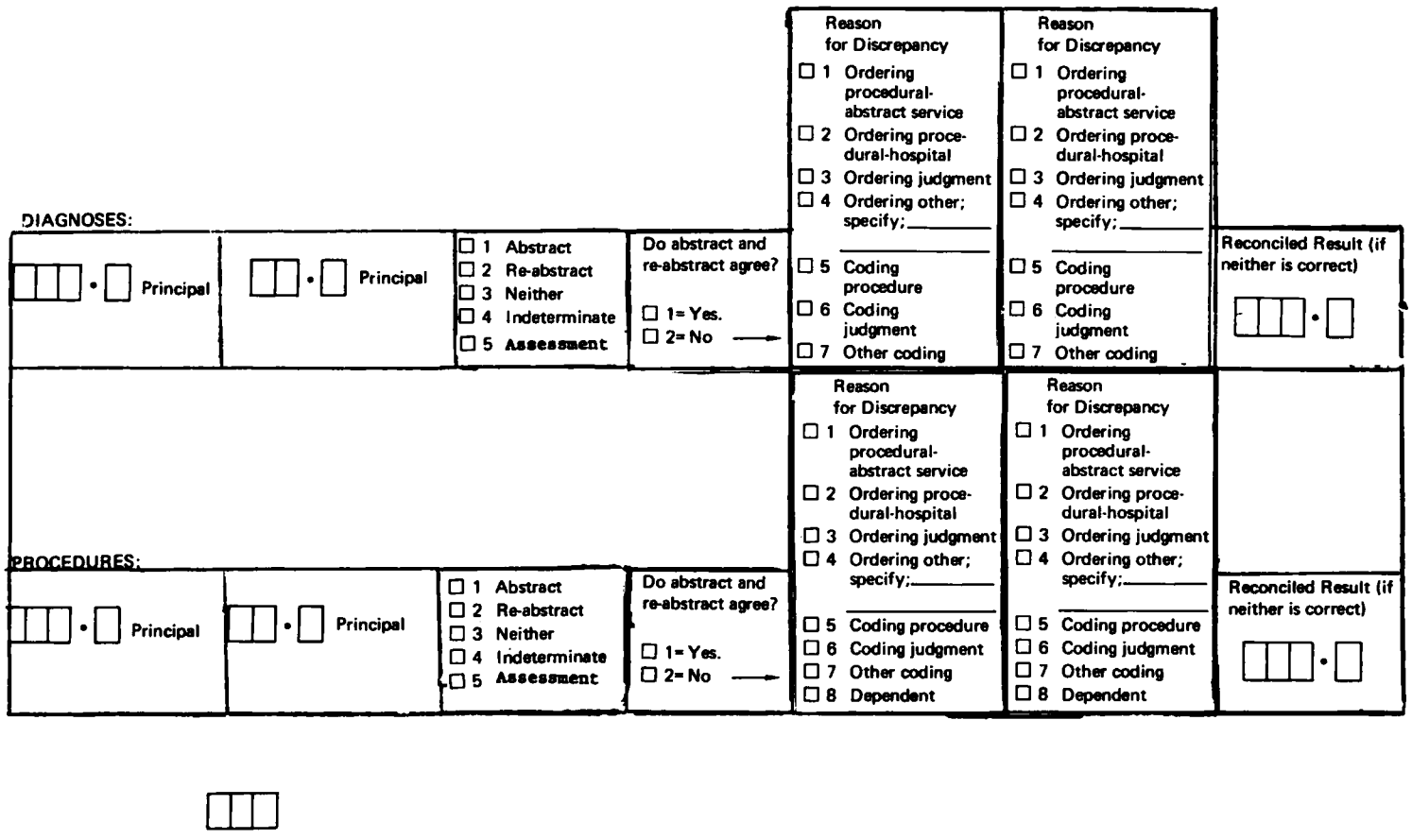
(Form is continued on the next page)

Appendix G

(Independent Assessment Re-abstracting Form continued)

Information from IOM re-abstract	Information from abstract service	Which is/are correct	Does information from the re-abstract agree with that of the abstract service?	If abstract and re-abstract are different, note reason for discrepancy.	If abstract and re-abstract agree but the validation abstract does not and is determined to be correct, note reason for discrepancy.	Reconciled result if all three are determined to be incorrect.
ADMIT DATE M M D D Y Y	M M D D Y Y	☐ 1 Abstract ☐ 2 Re-abstract ☐ 3 Neither ☐ 4 Indeterminate ☐ 5 **Assessment**	☐ 1= Yes. ☐ 2= No →	☐ 1= Clerical ☐ 2= Completeness ☐ 3= Procedural	☐ 1= Clerical ☐ 2= Completeness ☐ 3= Procedural	M M D D Y Y
DISCHARGE DATE M M D D Y Y	M M D D Y Y	☐ 1 Abstract ☐ 2 Re-abstract ☐ 3 Neither ☐ 4 Indeterminate ☐ 5 **Assessment**	☐ 1= Yes. ☐ 2= No →	☐ 1= Clerical ☐ 2= Completeness ☐ 3= Procedural	☐ 1= Clerical ☐ 2= Completeness ☐ 3= Procedural	M M D D Y Y
SEX ☐ 1= Male ☐ 2= Female ☐ 3= Other ☐ 9= Not recorded	☐ 1= Male ☐ 2= Female ☐ 3= Other ☐ 4= Not recorded	☐ 1 Abstract ☐ 2 Re-abstract ☐ 3 Neither ☐ 4 Indeterminate ☐ 5 **Assessment**	☐ 1= Yes. ☐ 2= No →	☐ 1= Clerical ☐ 2= Completeness ☐ 3= Procedural	☐ 1= Clerical ☐ 2= Completeness ☐ 3= Procedural	☐ 1= Male ☐ 2= Female ☐ 3= Other ☐ 4= Not recorded
DATE OF BIRTH (for HUP and CHDC) M M D D Y Y	M M D D Y Y	☐ 1 Abstract ☐ 2 Re-abstract ☐ 3 Neither ☐ 4 Indeterminate ☐ 5 **Assessment**	☐ 1= Yes. ☐ 2= No →	☐ 1= Clerical ☐ 2= Completeness ☐ 3= Procedural	☐ 1= Clerical ☐ 2= Completeness ☐ 3= Procedural	M M D D Y Y
AGE (for PAS and QUEST) ☐☐ ☐ 1= Days ☐ 2= Mos. ☐ 3= Yrs.	☐☐ ☐ 1= Days ☐ 2= Mos. ☐ 3= Yrs.	☐ 1 Abstract ☐ 2 Re-abstract ☐ 3 Neither ☐ 4 Indeterminate ☐ 5 **Assessment**	☐ 1= Yes. ☐ 2= No →	☐ 1= Clerical ☐ 2= Completeness ☐ 3= Procedural	☐ 1= Clerical ☐ 2= Completeness ☐ 3= Procedural	☐☐ ☐ 1= Days ☐ 2= Mos. ☐ 3= Yrs.
ANTICIPATED PRINCIPAL SOURCE OF PAYMENT ☐ 1= Medicare ☐ 2= Medicaid/ Medi-Cal ☐ 3= Other ☐ 9= Not recorded	☐ 1= Medicare ☐ 2= Medicaid/ Medi-Cal ☐ 3= Other ☐ 4= Not recorded	☐ 1 Abstract ☐ 2 Re-abstract ☐ 3 Neither ☐ 4 Indeterminate ☐ 5 **Assessment**	☐ 1= Yes. ☐ 2= No →	☐ 1= Clerical ☐ 2= Completeness ☐ 3= Procedural ☐ 4= Change in status	☐ 1= Clerical ☐ 2= Completeness ☐ 3= Procedural ☐ 4= Change in status	☐ 1= Medicare ☐ 2= Medicaid/ Medi-Cal ☐ 3= Other ☐ 4= Not recorded

(Independent Assessment Re-abstracting Form continued)



Appendix H

DIFFERENCES BETWEEN THE ORIGINAL ABSTRACT AND IOM RE-ABSTRACT WHEN THE CORRECT DATA SOURCE COULD NOT BE DETERMINED*

Diagnosis Listed on Original Abstract	Diagnosis Listed on Re-abstract
250.0 - Diabetes mellitus without complications	427.9 - Other and specified disorders of heart rhythm
250.0 - Diabetes mellitus without complications	516.9 - Other pneumoconioses and related diseases
250.0 - Diabetes mellitus without complications	599.5 - Other and unspecified diseases of urinary tract
250.0 - Diabetes mellitus without complications	450.0 - Pulmonary embolism and infarction
250.0 - Diabetes mellitus without complications	360.0 - Conjunctivitis and opthalmia
250.0 - Diabetes mellitus without complications	790.2 - Depression
250.0 - Diabetes mellitus without complications	573.9 - Diseases of the liver - other and unspecified
250.2 - Diabetes mellitus with or without acidosis	707.1 - Lower extremity, except decubitus ulcer
250.2 - Diabetes mellitus with or without acidosis	445.9 - Gangrene not elsewhere classified
250.2 - Diabetes mellitus with or without acidosis	443.9 - Other peripheral vascular disease unspecified
250.9 - Diabetes mellitus without mention of acidosis or coma	427.0 - Congestive heart failure

*Includes all such cases when either diabetes mellitus, acute myocardial infarction, or chronic ischemic heart disease was coded as principal on the original abstract, since these three diagnoses were the most frequently occurring of all "Indeterminates."

Appendix H

Original Abstract	Re-abstract
250.9 - Diabetes mellitus without mention of acidosis or coma	523.9 - Periodontal disease - other and unspecified
250.9 - Diabetes mellitus without mention of acidosis or coma	427.9 - Other and unspecified disorders of heart rhythm
412.0 - Healed myocardial infarction	402.0 - Hypertensive heart disease
412.0 - Healed myocardial infarction	411.9 - Other acute and subacute forms of ischemic heart disease without mention of hypertensive disease
412.1 - Chronic ischemic heart disease	434.1 - Cerebral embolism with paralysis
412.9 - Chronic ischemic heart disease - other and unspecified	427.4 - Atrial (auricular) fibrillation or flutter
412.9 - Chronic ischemic heart disease - other and unspecified	410.9 - Ischemic heart disease - site not specified
412.9 - Chronic ischemic heart disease - other and unspecified	712.4 - Rheumatoid (ankylosing) spondylitis
412.9 - Chronic ischemic heart disease - other and unspecified	782.9 - Shock without mention of trauma
412.9 - Chronic ischemic heart disease - other and unspecified	427.0 - Congestive heart failure
412.9 - Chronic ischemic heart disease - other and unspecified	427.0 - Congestive heart failure
412.9 - Chronic ischemic heart disease - other and unspecified	782.1 - Palpitation
412.9 - Chronic ischemic heart disease - other and unspecified	410.6 - Acute myocardial infarction
412.9 - Chronic ischemic heart disease - other and unspecified	427.0 - Congestive heart failure
412.9 - Chronic ischemic heart disease - other and unspecified	427.9 - Other and unspecified disorder of heart rhythm

Appendix H

Original Abstract	Re-abstract
412.9 - Chronic ischemic heart disease other and unspecified	427.9 - Other and unspecified disorder of heart rhythm
412.9 - Chronic ischemic heart disease - other and unspecified	436.0 - Acute, but ill-defined cerebrovascular disease without paralysis
412.9 - Chronic ischemic heart disease - other and unspecified	413.0 - Angina pectoris with hypertensive disease
412.9 - Chronic ischemic heart disease - other and unspecified	427.9 - Other and unspecified disorder of heart rhythm
412.9 - Chronic ischemic heart disease - other and unspecified	427.1 - Symptomatic heart disease - left ventricular failure
412.9 - Chronic ischemic heart disease - other and unspecified	411.0 - Ischemic heart disease with hypertensive disease
412.9 - Chronic ischemic heart disease - other and unspecified	411.0 - Ischemic heart disease with hypertensive disease
412.9 - Chronic ischemic heart disease - other and unspecified	427.4 - Atrial (auricular) fibrillation of flutter
412.9 - Chronic ischemic heart disease - other and unspecified	427.0 - Congestive heart failure
412.9 - Chronic ischemic heart disease - other and unspecified	427.0 - Congestive heart failure
412.9 - Chronic ischemic heart disease - other and unspecified	427.0 - Congestive heart failure
412.9 - Chronic ischemic heart disease - other and unspecified	427.0 - Congestive heart failure
174.0 - Malignant neoplasm of breast	805.2 - Fracture and fracture-dislocation of vertebral column without mention of spinal cord lesion-- Dorsal (thoracic) closed
174.0 - Malignant neoplasm of breast	199.0 - Multiple malignant neoplasm without specification of site

Appendix H

Original Abstract	Re-abstract
174.0 - Malignant neoplasm of breast	198.5 - Other secondary malignant neoplasm - other parts of nervous system
174.0 - Malignant neoplasm of breast	199.0 - Multiple malignant neoplasm without specification of site

Appendix I

NET AND GROSS DIFFERENCE RATES IN DESIGNATION OF PRINCIPAL DIAGNOSIS (BASED ON FOUR-DIGIT COMPARISONS)*

	Percent of abstracts with indicated principal diagnoses			Net and gross difference rates (times 1,000)	
IOM re-abstract	Original abstract			Net	Gross
	Cerebrovascular disease	Other	Total		
Cerebrovascular disease	3.5	1.3	4.8		
Other	0.2	95.0	95.2		
Total	3.7	96.3	100.0%	10.74	14.62
	Chronic ischemic heart disease	Other	Total		
Chronic ischemic heart disease	1.1	0.3	1.4		
Other	2.5	96.0	98.5		
Total	3.6	96.3	99.9%	-22.37	28.24
	Acute myocardial infarction	Other	Total		
Acute myocardial infarction	1.6	1.0	2.6		
Other	0.1	97.3	97.4		
Total	1.7	98.3	100.0%	8.58	11.37
	Diabetes	Other	Total		
Diabetes	0.9	0.5	1.4		
Other	0.5	98.1	98.6		
Total	1.4	98.6	100.0%	0.91	10.07

*See discussion on pages 35-37 of text.

Appendix I

IOM re-abstract	Original abstract			Net	Gross
	Carcinoma of the breast	Other	Total		
Carcinoma of the breast	0.5	0.0	0.5		
Other	0.1	99.4	99.5		
Total	0.6	99.4	100.0%	-1.04	1.58
	End stage renal disease	Other	Total		
End stage renal disease	0.3	0.1	0.4		
Other	0.1	99.5	99.6		
Total	0.4	99.6	100.0%	0.42	1.93
	Cholelithiasis/ cholecystitis	Other	Total		
Cholelithiasis/ cholecystitis	1.2	0.4	1.6		
Other	0.1	98.3	98.4		
Total	1.3	98.7	100.0%	3.11	5.02
	Hernia without obstruction	Other	Total		
Hernia without obstruction	0.9	0.1	1.0		
Other	0.1	98.9	99.0		
Total	1.0	99.0	100.0%	-0.15	1.34
	Delivery	Other	Total		
Delivery	1.5	0.2	1.7		
Other	0.0	98.3	98.3		
Total	1.5	98.5	100.0%	2.27	2.37

Appendix I

IOM re-abstract	Original abstract			Net	Gross
	Hypertrophy of t&a	Other	Total		
Hypertrophy of t&a	0.5	0.0	0.5		
Other	0.0	99.5	99.5		
Total	0.5	99.5	100.0%	-0.07	0.25
	Fracture of neck of femur	Other	Total		
Fracture of neck of femur	1.3	0.4	1.7		
Other	0.0	98.3	98.3		
Total	1.3	98.7	100.0%	3.43	4.13
	Displacement of intervertebral disc	Other	Total		
Displacement of intervertebral disc	0.3	0.2	0.5		
Other	0.1	99.4	99.5		
Total	0.4	99.6	100.0%	1.00	2.24
	Cataract	Other	Total		
Cataract	3.0	0.2	3.2		
Other	0.0	96.8	96.8		
Total	3.0	97.0	100.0%	1.78	2.17
	Neuroses	Other	Total		
Neuroses	0.5	0.2	0.7		
Other	0.1	99.2	99.3		
Total	0.6	99.4	100.0%	0.93	2.82