Improving

DATA QUALITY

A GUIDE FOR DEVELOPING COUNTRIES

WORLD HEALTH ORGANIZATION
REGIONAL OFFICE FOR THE WESTERN PACIFIC

IMPROVING DATA QUALITY: A GUIDE FOR DEVELOPING COUNTRIES

WORLD HEALTH ORGANIZATION
Regional Office for the Western Pacific
United Nations Avenue
1000 Manila, Philippines

Fax No : (63-2) 521-1036
Tel. No : (63-2) 528 8001
Email : postmaster@wpro.who.int
Website : http://www.wpro.who.int

ISBN 92 9061 050 6

Improving Data Quality

A GUIDE FOR DEVELOPING COUNTRIES

WORLD HEALTH ORGANIZATION
REGIONAL OFFICE FOR THE WESTERN PACIFIC

Contents

3 Quality in data collection   28

4 Clinical coding quality   38

5 Data quality of statistical reports   54

6 Data quality in public health records and reports   64

Acknowledgement

Introduction

Accurate, timely and accessible health care data play a vital role in the planning, development and maintenance of health care services. Quality improvement and the timely dissemination of quality data are essential if health authorities wish to maintain health care at an optimal level.

In recent years, data quality has become an important issue, not only because of its importance in promoting high standards of patient care, but also because of its impact on government budgets for the maintenance of health services.

Authorities at all levels of health care, including hospitals, community health centres, outlying clinics and aid posts, as well as ministries or departments of health, should be concerned about poor data quality and the impact it has on the quality of health care. In many countries, administrators are dogged by poor medical/health record documentation, large backlogs of medical records waiting to be coded and inconsistent coding, plus poor access to, and utilization of, accurate and accessible morbidity data.

These concerns not only relate to the quality of medical/health record documentation but also to the collection of health care statistics at all levels, from the largest hospital to the smallest clinic or aid post. At the hospital or clinic level, statistics are used to assess how much services are being used to enable the facility to make appropriate financial and administrative plans, and to conduct vital research. At state, province and national levels, governments use health statistics for planning health care services and allocating resources where they are needed most. The accuracy and relevance of the information processed, thus, are vital to the smooth running of the facility and also in assisting governments with decisions on the provision of health care services locally and nationally.

I n recent years, data quality has become an important issue, not only because of its importance in promoting high standards of patient care, but also because of its impact on government budgets for the maintenance of health services.

Many clinics and aid posts offer immunization programmes, counselling, family planning and other health-related services that are not always considered parts of "traditional" health care but are, in fact, important components within the community. As such, they must maintain accurate and reliable data.

To address these issues and improve the quality of data collected, as well as the information generated from that data, quality-control measures need to be taken.

Aims and objectives of these guidelines

The purpose of this booklet is to provide a set of guidelines to enable health care workers, health information managers and administrators at all levels to focus on improving the timeliness, accuracy and reliability of health care data. Although the emphasis might seem to be on hospitals and hospital medical records, these guidelines have been designed to address all areas in health care where data are collected and information generated. The guidelines describe the activities that should be considered when addressing the question of data quality in health care, regardless of the setting. The reader is guided to assess and, where necessary, improve the quality of data generated in the environment within which they function, regardless of size, remoteness or sophistication.

Target readership

These guidelines are aimed at government policymakers and health care administrators at primary, secondary and tertiary levels of health care, as well as doctors, nurses, other health care providers and health information managers. All of these share responsibility for the documentation, implementation, development, and management of health information services.

Using the guidelines

It is important that the reader treats these guidelines not as a set of definitive rules applicable in every situation but as an *aide memoire,* to help ensure that important activities have been considered and addressed to improve data and ensure data quality.

Although the term "patient" is generally used in the following pages, it is important to note that health services in many countries also use terms such as "consumer" or "client."

The book's structure

The text has been divided into six chapters. The first two chapters deal with defining data and data quality. In Chapter 1, data are defined and the difference between data and information explained. In Chapter 2, data quality is defined along with a description of the components of data quality. Data

standards and their implementation are also outlined together with quality control methods and performance improvement measures. Limitations in overcoming problems related to data quality are also discussed along with a description of some measures that can be used to solve identified problems of data quality.

Chapters 3, 4, 5 and 6 deal with specific areas of data quality. Chapter 3 deals with the type of data collected, how they are collected, causes and sources of poor data collection, and methods that could be used to overcome such problems.

Chapter 4 is dedicated to clinical coding quality and, again, issues relating to the causes and sources of poor quality data and methods to improve the quality of coded clinical information are highlighted. An outline of activities that could be used to measure coding productivity and conduct coding audits is included.

In Chapter 5, special attention is given to the quality of statistical reports, including the need for standardization of terms and formulae and the appropriateness and presentation of statistical information. The causes and sources of poor quality statistical data and methods to improve the quality of statistical reports are also covered.

Last, Chapter 6 deals with causes and sources of poor data quality in public health records and reports, along with possible methods to improve the quality of public health data collections.

1 Defining data and information

Whether we collect data on paper or in a computer, the data should be organized in such a way that we can understand and retrieve them when needed.

The starting points for health care information are data and the collection of data, whether maintained manually or electronically at a large teaching hospital, health centre or outlying clinic. Demographic and clinical data stored in a patient's medical/health record are the major source of health information and are of no value to medical science or health care management if they are not accurate, reliable and accessible.

The comparison of health care data between facilities, states or provinces, within a country or between countries, is vital to the growth and dissemination of health information throughout the world. This possible sharing is meaningless, however, without the use of standardized systems for data collection, disease classification and health care statistics.

Before discussing data quality, we need to define data, how data are transformed into information and how they are collected.

What are data?

Data may be defined as a representation of facts or concepts or instructions in a formalized manner, suitable for communication, interpretation or processing by manual or electronic means. An element of data is an item, idea, concept or raw fact (Abdelhak et al., 1996).

In health care, these facts describe specific characteristics of individual patients. Whether we collect data on paper or in a computer, the data should be organized in such a way that we can understand and retrieve them when needed (Davis and LaCour, 2002).

Primary and secondary data

There are two types of data – primary data and secondary data. In the context of health care, primary data and secondary data are distinguishable as follows (Abdelhak et al., 1996).

> *Primary data* are obtained from the original data source. That is, documentation in the patient's medical/health record collected by staff at either a hospital, clinic or aid post. Daily ward census reports collected in hospitals are also primary data.
>
> *Secondary data* are data sets derived from primary data. Secondary data are individual or aggregate health care data found in reports that are summarized from the source. At the hospital or health centre level, secondary data include the master patients' index, disease and procedure indexes, health care statistics and disease registries. At primary care level, they also include such aspects as the patients' name index and statistics.

Data and information

In order to interpret data – to make sense of the facts and use them – the data items need to be organized. Once organized, data become information.

Health care data

Health care data are items of knowledge about an individual patient or a group of patients. In health care, data are captured about a patient in paper or electronic format during his or her attendance at an outpatient clinic, community health centre, primary health care provider, or his or her admission to a hospital (Davis and LaCour, 2002). The data collected should include all relevant findings relating to the patient's condition, diagnoses, treatment, if any, and other events as they occur.

Whether the data are collected manually or in a computer, it is important to ensure that the information is correct at **the point of entry**.

To ensure data quality, two key principles are **data accuracy and data validity**. To communicate effectively, data must be valid and conform to an expected range of values. To be useful, data must be accurate.

As the recording of data is subject to human error, there needs to be **built-in control measures** to eliminate errors, both in manual recording and computer entry.

Health Information

Health information is health care data that have been organized into a meaningful format.

Health information may refer to **organized data** collected about an individual patient, **or** the summary of information about that patient's entire experience with his or her health care provider (Davis and LaCour, 2002).

Health information can also be the **aggregate information** about all patients that have attended or been admitted to a hospital, or attended a health centre, outlying clinic or a community immunization or health screening programme.

Health information, therefore, can encompass the organization of a limitless array and combination of possible data items (Davis and LaCour, 2002).

> *Providers of health care services need information not only at the point of service but also at the point of decision making in a format that maximizes the decision-making process.*

Health care information

As a by-product of a patient's contact with a health care provider, regardless of the setting, information should have value as a clinical review or management tool. Whether in a manual system or computer, information will not be valuable unless it is accurate, relevant, structured and presented in an easily useable form. Health care information should be capable of:

■ promoting excellent clinical care;

■ describing the types of individuals using services and the types of services they receive;

■ measuring efficiency of the contact, treatment, referral and interaction by health care professionals;

■ helping in the co-ordination of care between services provided;

■ providing meaningful statistics for determining the health status of the community;

■ measuring quality from a patient and provider perspective; and

■ meeting accountability requirements.

Correct and up-to-date information is critical, not only for the provision of high-quality clinical care, but also for continuing health care, maintaining health care at an optimal level, clinical and health service research, and planning and management of health systems.

Accurate information about resources used and services delivered at all levels of health care is essential for resource management, use of clinical evidence and measurement of outcomes to improve the effectiveness of health care services.

The collection and use of information should not impose a burden on the health system. It should be collected as a routine by-product of the health care process.

Importance of data quality

Accurate and reliable health care data are needed for:

- determining the continuing and future care of a patient at all levels of health care;
- medico-legal purposes for the patient, the doctor and the health care service;
- maintaining accurate and reliable information about diseases treated and surgical procedures performed in a hospital and within a community, as well as immunization and screening programmes, including the number and type of participants;
- clinical and health service research and outcomes of health care intervention, if required;
- accurate, reliable and complete statistical information about the uses of health care services within a community;
- teaching health care professionals; and
- working out staffing requirements and planning health care services.

Accurate and reliable health care data are used by:

- **doctors, nurses and other health care professionals** treating a patient admitted to a hospital or seen in an outpatient department or emergency room, and in community health centres, outlying clinics or general practitioners' offices. They use the medical/health record as a means of communication during an episode of care and treatment of a patient, and as an *aide memoire* for continuing care of that particular patient. Doctors also use health care data to evaluate the services provided;
- **nursing staff** in hospitals to evaluate data and develop critical pathways and patient care plans for admitted patients;

- **health insurers** who require information to reimburse the patient and/or health care facility for services rendered whether for an inpatient or ambulatory patient;
- **legal representatives and courts** as documentary evidence of a patient's care and treatment by a health care worker in a hospital, health centre or clinic. They are also used to protect the legal interests of the patient, doctor and other health care professionals, the health care facility, and the public;
- **the state and national ministry of health** to review vital statistics and the incidence and prevalence of disease in a city, state or country. The provision of accurate and reliable aggregate data is important for public policy development and funding of health care services;
- **quality assurance committees and medical staff** as a basis for analysis, study and evaluation of the quality of health care services rendered to patients;
- **researchers**, to analyse and interpret data to determine causes, prevention methods and treatment for diseases, injuries and disabilities;
- **health care facility accreditation and licensing agencies** to review medical/health records to provide public assurance that quality health care is provided; and
- **national governments** who use the information to develop health care policy and provide and regulate funds.

If data are not correct and valid, the usefulness for the above purposes will be jeopardized.

Summary

To maintain data quality, health care data must be (Abdelhak et al., 1996):
- **appropriate** – for the overall needs of the patient and health care services;
- **organized** – to enable it to be understood and used when required;
- **timely** and **available** – to users when required;
- **accurate** and **complete** – otherwise it could be difficult to make appropriate decisions; and
- **Cost-effective** – if the cost of collecting and disseminating information exceeds its value, the usefulness of the data collected must be addressed.

In the following chapters, the importance of data quality will be discussed in more depth along with guidelines on how to maintain quality in all areas of documentation, including the collection of data at the original source. Data quality in clinical coding, statistical collections and reports, as well as the dissemination of information to health care authorities, will be covered.

References

Abdelhak M, Grostick S, Hankin MA, Jacobs E. *Health Information: Management of a Strategic Resource.* Philadelphia, WB Saunders Company, 1996.

Davis N, LaCour M. *Introduction to Information Technology*. Philadelphia, WB Saunders Company, 2002.

Johns ML. *Health Information Management Technology*. Chicago, American Health Information Management Association, 29, 2002.

2 Data quality

The original data must be accurate in order to be useful. If data are not accurate, then wrong impressions and information are being conveyed to the user.

As coded health care data are being increasingly used in the health care environment, it is important to ensure that the **original source data** are accurate and timely, which in turn, will produce reliable and useful information.

Regardless of whether in a hospital, health centre or an outlying clinic, the quality of health care data and statistical reports has come under intensive scrutiny in recent years. Thus, all health care service employees, including clerical staff, health professionals, administrators, and health information managers, need to gain a thorough knowledge and understanding of the key components of data quality and the requirements for continuous data improvement.

Data quality — what it means

In general terms, quality, as defined by Donabedian (1988), consists of the ability to achieve desirable objectives using legitimate means. Quality data represent what was intended or defined by their official source, are objective, unbiased and comply with known standards (Abdelhak et al., 1996). Data quality includes:

- **accuracy and validity** – of the original source data;
- **reliability** – data are consistent and information generated is understandable;
- **completeness** – all required data are present;
- **legibility** – data are readable;
- **currency and timeliness** – data are recorded at the time of observation;
- **accessibility** – data are available to authorized persons when and where needed;
- **meaning or usefulness** – information is pertinent and useful; and

■ **confidentiality and security** – both important particularly to the patient and in legal matters.

The central theme is that data quality is proportionate to the attainment of achievable improvements in health care.

Importance of data quality in health care

Health care data are maintained for the present and future care of the patient regardless of the level at which the service is provided. The quality of that data is crucial, not only for use in patient care, but also for monitoring the performance of the health service and employees. Data collected and presented must be accurate, complete, reliable, legible and accessible to authorized users if they are to meet the requirements of the patient, doctor and other health professionals, the health care facility, legal authorities, plus state, province and national government health authorities.

Components of data quality

Accuracy and validity

The original data must be accurate in order to be useful. If data are not accurate, then wrong impressions and information are being conveyed to the user (Davis and LaCour, 2002). Documentation should reflect the event as it actually happened. Recording data is subject to human error and steps must be taken to ensure that errors do not occur or, if they do occur, are picked up immediately.

Example of accuracy and validity in a manual medical record system
■ The patient's identification details are correct and uniquely identify the patient.
■ All relevant facts pertaining to the episode of care are accurately recorded.
■ All pages in the health record are for the same patient.
■ The patient's address on the record is what the patient says it is.
■ Documentation of clinical services in a hospital is of an acceptable predetermined value.
■ The vital signs are what were originally recorded and are within acceptable value parameters, which have been predetermined and the entry meets this value.
■ The abstracted data for indices, statistics and registries meet national and international standards and have been verified for accuracy.
■ The codes used in hospitals to classify diseases and procedures conform to pre-determined coding standards.

In a manual system, processes need to be in place to monitor data entry and collection to ensure quality. In a computerized system, a computer can be instructed to check specific fields for validity and

alert the user to a potential data collection error. Computer systems have in-built checks such as edit and validation checks, which are developed to ensure that the data added to the record are valid. Edits or rules should be developed for data format and reasonableness, entailing conditions that must be satisfied for the data to be added to the database, along with a message that will be displayed if the data entry does not satisfy the condition. In some instances, the computer does not allow an entry to be added if it fails the edit. In other instances, a warning is provided for the data entry operator to verify the accuracy of the information before entry.

Examples of edits and validity in a computer-based system

- In a hospital system, a patient must have a unique number because it is the key indexing or sorting field.
- The patient's number must fall within a certain range of numbers or the computer does not allow the data entry operator to move to the next field or to save the data.
- For hospital patients, the date of admission must be the same as or earlier than the date of discharge.
- A laboratory value must fall within a certain range of numbers or a validity check must be carried out.
- Format requirements such as the use of hyphens, dashes or leading zeros must be followed.
- Consistency edits can be developed to compare fields – for example a male patient cannot receive a pregnancy test.

Reliability

Data should yield the same results on repeated collection, processing, storing and display of information. That is, data should be consistent.

Examples of reliability

- The diagnosis recorded on the front sheet of the hospital medical record is consistent with the diagnosis recorded in the progress notes and other relevant parts of the medical record.
- Surgical procedures recorded on the front sheet of the hospital medical record are the same as recorded in operation reports in the body of the medical record.
- The age of the patient recorded on the first sheet of a medical/health record is the same as that recorded on other pages.
- The correct name of the patient is recorded on all forms within the medical/health record at the point of care or service given.

Completeness

All required data should be present and the medical/health record should contain all pertinent documents with complete and appropriate documentation.

Examples of completeness

- The first sheet of the medical/health record contains all the necessary identifying data to uniquely identify an individual patient.
- For inpatients, the medical record contains an accurately recorded main condition and other relevant diagnoses and procedures and the attending doctor's signature.
- Also for inpatients, all progress notes — from date of admission to discharge or death — are complete with signatures and date of entry.
- Nursing notes, including nursing plan, progress notes, blood pressure, temperature and other charts are complete with signatures and date of entry.
- For all medical/health records, relevant forms are complete, with signatures and date of attendance.

Legibility

All data whether written, transcribed and/or printed should be readable.

Examples of legibility

- Handwritten demographic data are clearly written and readable.
- Handwritten notes are clear, concise, readable and understandable.
- In all medical/health records, undecipherable codes or symbols are **not used** in either manual or electronic patient records.
- If abbreviations are used, they are standard and understood by all health care professionals involved in the service being provided to the patient.

Timeliness

Information, especially clinical information, should be documented as an event occurs, treatment is performed or results noted. Delaying documentation could cause information to be omitted and errors recorded.

Example of timeliness

- A patient's identifying information is recorded at the time of first attendance and is readily available to identify the patient at any given time.
- The patient's past medical history, a history of the present illness/problem as detailed by the patient, and results of physical examination, is recorded at the first attendance at a clinic or admission to hospital.
- On discharge or death of a patient in hospital, his or her medical records are processed and completed, coded and indexed within a specified time frame.
- Statistical reports are ready within a specified time frame, having been checked and verified.

Accessibility

All necessary data are available when needed for patient care and for all other official purposes. The value of accurately recorded data is lost if it is not accessible.

Examples of accessibility
- Medical/health records are available when and where needed at all times.
- Abstracted data are available for review when and where needed.
- In an electronic patient record system, clinical information is readily available when needed.
- Statistical reports are accessible when required for patient-care committees, planning meetings and government requirements.

Development and implementation of standards for health care documentation

The primary purpose of recording data is communication. To make sure that communication is achievable, it is important to ensure that the language of health information, regardless of the level of service and information concepts, conform to a recognized standard.

Prior to introducing procedures to monitor data quality, the health care facility, whether a hospital, health centre, or outlying clinic, should look at implementing standards for health care documentation, processing and maintenance.

What are standards and how are they developed?

A standard is a specimen or specification by which something may be tested or measured.

In health care, standards are defined as:

Best practice principles and guidelines for collection and storage of health care data (Abdelhak et al., 1996).

Standards that define demographic and other identifying data elements suited for data capture and use in patient identification in health care settings and provide guidelines for their application should be developed and adhered to by staff.

Clinical data standards

Clinical data standards should also be developed to ensure that data collected in one facility means the same in another, allowing for comparison between facilities. Two or more persons looking at the content of a medical/health record may have different opinions about its adequacy. An assessment about the adequacy of patient care information in the absence of any explicit standards is a value judgement.

To be useful, standards should be developed to promote data quality at the time of documentation.

In many countries, to achieve standardization of data, steps such as the development of **Minimum Data Sets** and a **data dictionary** have been instituted.

Minimum Data Set

A minimum data set is a core set of data elements agreed to by a National Health Information Group for the mandatory collection and reporting at a national level.

It is reliant on a national agreement to collect uniform data and to supply it as part of the national collection, but does not preclude a health care facility from collecting additional data to meet its own needs.

Australia and the United States of America are two countries that have national minimum data sets. These are used to identify data elements that should be collected by health care facilities for each patient and provide uniform definitions for common terms. In the United States of America, the list of patient-specific data items used is the **Uniform Hospital Discharge Data Set** (UHDDS) (Johns, 2002).

An example of UHDDS Data Elements (Johns, 2002)	
Data element	**Definition/Descriptor**
01. Personal identification	The unique number assigned to each patient within a hospital that distinguishes the patient and his or her hospital record from all others in that institution
02. Date of birth	Month, day and year of birth. Capture of the full four-digit year of birth is recommended
03. Sex	Male or female

Data dictionary

A data dictionary is a set of common standards for data collection.

> "*A data dictionary contains a set of core uniform definitions and data items relating to the full range of health services and a range of population parameters*" *(AIHW, 2001).*

It is used to promote uniformity, availability, reliability, validity, consistency and completeness of data at all levels of health care service. The data dictionary is useful especially in promoting national standard definitions and should be readily available to all individuals and organizations involved in the generation, use and/or development of health and health service information.

The National Health Data Dictionary, Version 10, produced by the Australian Institute of Health and Welfare (AIHW), Canberra, is used extensively by health care services in Australia to ensure the production of **standardized data collections**. Data elements are based on the ISO/IEC Standards 11179 *Specification and Standardization of Data Elements.* The dictionary is used to help ensure that data elements are collected uniformly from all health care facilities to improve the quality of information for community discussion and public policy debate on health issues in Australia (AIHW, 2001).

Data dictionaries can also exist at state or province level and should be consistent with the appropriate national standards.

A sample entry in the data dictionary

Health Outcome

Identifying and definitional attributes :

Data element type	:	Data element concept
Definition	:	A change in the health of an individual, or group of people or a population, which is wholly or partially attributable to an intervention or a series of interventions.
Context	:	Admitted patient and non-admitted patient health care.

Some countries also produce standards on documentation associated with **Accreditation Guidelines for Health Care Facilities,** including hospitals and community health care centres and clinics. For example, medical/health record documentation determined by the Joint Commission on Accreditation of Healthcare Organizations (JCAHO, 1995) in the United States of America include standards such as:

Example of standards produced by JCAHO

IM.7.5.1	Records of emergency visits contain time and means of arrival.
IM.7.5.2	Records of emergency visits contain conclusions at end of treatment, including final disposition, condition and instructions for follow-up.

An example of a standard related to medical records published in the First Edition of the Joint Commission International Accreditation Standards for Hospitals (JCIASH, 2000) is as follows:

Standard MOI.2.1

The clinical record contains sufficient information to identify the patient, support the diagnosis, justify the treatment, document the course and results of treatment, and promote continuity of care among health care providers (MOI.2.1 JCIASH).

Measurable elements of MOI.2.1 include:
1. patient clinical records contain adequate information to identify the patient;
2. patient clinical records contain adequate information to support the diagnosis;
3. patient clinical records contain adequate information to justify the care and treatment;
4. patient clinical records contain adequate information to document the course and results of treatment;
5. patient clinical records promote continuity of care; and
6. the specific content of patient clinical records has been determined by the organization.

The intent of this standard is that the clinical record of each patient needs to present sufficient information to support the diagnosis, justify the treatment provided, and document the course and results of treatment. A standardized format and content of a patient's clinical record helps promote the integration and continuity of care among the various providers of care to the patient.

The Australian Council on Healthcare Standards also provides a guide to health care facility accreditation in which standards for medical records are included. An example of a standard is:

Standard 4.2   Data collection, aggregation and use

Patient/consumer care, management of services, education and research are facilitated by the collection and aggregation of data.

Criteria – data management

4.2.1   Relevant, accurate, quantitative and qualitative data are collected and used in a timely and efficient manner for delivery of patient/consumer care and management of services.

Information created by the organization needs to be timely, accurate, clear, concise and presented in a way that is appropriate to the users' needs. Excessive quantities or poorly constructed information can deter a person from reading or using it. Regular review of the organization's reporting systems can assist in identifying how well the information systems meet users' needs and whether any action is required to improve the system. Those responsible for providing information can monitor their performance through measuring production delays, rework and frequency of inaccurate information. Feedback from staff may also provide suggestions for improving the systems for managing information.

4.2.2   The collection of data and reporting of information comply with professional and Australian standards and statutory requirements.

Ensuring consistency is an important part of the data collection process. Data consistency is promoted by complying with protocols, definitions and standards. It is also promoted by the use of standard definitions, symbols and abbreviations throughout the organization. Periodic evaluation of data consistency and quality can help ensure the data collected by the organization and its services are appropriate and contribute to management and patient care.

4.2.3   Indexing of data facilitates information retrieval.
Effective indexing of data assists the storage and retrieval of information.

4.2.4   Reference and research information is collected and managed for use by staff and patients/consumers in achieving the organization's goals.

Clinical records

4.2.5   The organization uniquely identifies all patients/consumers, including newborn infants. The organization needs to consider the efficiency and effectiveness of its system of uniquely identifying patients/consumers.

Other

Other segments of this standard include:

4.2.6   Every patient/consumer has a sufficiently detailed record to assist with the continuity of care, education, research, evaluation and medico-legal and statutory requirements.

4.2.7   Care providers document details in the clinical record. All entries are legible, dated and signed with designation.

4.2.8   In circumstances where part of the clinical record is available electronically, appropriate mechanisms exist to alert staff to the existence of these data.

4.2.9   Clinical record data are coded and indexed to ensure the timely production of quality patient/consumer care information.

Benchmarking

Another method of setting standards is benchmarking, which may be defined as *"the process of systematically searching for and incorporating international best practice into an organization"* (AIHW, 1996). A benchmark may be achieved by comparing an organization, a standard or peer groups (Johns, 2002). It is a quality improvement technique used by one facility to compare its practices with that of another with superior performance by reviewing the process that has proven effective (Davis and LaCour, 2002). Benchmarking, however, may not always be achievable. Some problems that may impede benchmarking include agreement on a benchmark, finding partners to participate and training staff.

Leadership in data quality

Many health care administrators already recognize that quality improvement is the way to add value to the services offered and that the dissemination of quality data is the only way to demonstrate that value to health care authorities and the community.

- **Health care administrators/managers** should be leaders in the move to improve the quality of data collected in the health care facility as they are responsible for the overall management of the facility and the quality of the information produced.
- **Senior doctors** should take the lead in ensuring data quality by taking time to ensure the more junior doctors record clinical data accurately and in a timely manner. Doctors should play an important role in maintaining data quality and should understand the need for accurate and timely data in the care of patients.
- **Senior staff** in departments such as **laboratory, radiology and pathology** and **community nurses**, **allied health professionals such as physical therapists, occupational therapists and social workers,** should be responsible for the provision of accurate and timely reporting and better checks on the quality of the content of reports.
- **Senior medical record staff** should ensure that the medical/health record is complete, available and accessible when needed. They should also take responsibility for accurate and timely abstracted data and statistical reports. In a hospital setting, they should ensure that discharge summaries are accurately prepared and that diagnoses and procedures are coded accurately and within a specified time frame. A major part of data quality emanates from the medical record department through the timely preparation of the medical record, quality checks for accuracy and completeness of recorded data, and the availability of the medical/health record at all times when needed for patient care.

Routine data quality monitoring

With standards in place, procedures relating to data collection and monitoring data quality should be carried out on a routine basis.

Two monitoring procedures for inpatients that have been undertaken for many years in some countries are quantitative and qualitative analysis of medical records.

Quantitative analysis of medical records

To evaluate the quality of documentation and, subsequently, patient care, the medical record must be complete. In a quantitative analysis, medical records should be reviewed to check that all documentation has been included (Huffman, 1963), for example:

- Patient identification is accurate and all details are complete.
- The history, physical examination, all progress and nursing notes are present, and all relevant reports such as pathology, X-rays, etc., are included.

- If the patient had surgery an operation and anaesthetic report is present.
- All entries are signed and dated.

In other words, all relevant documentation must be present and authenticated.

> *In the assembly of medical records, the required standard is that all medical records are assembled within 24 hours of patient discharge with 100% accuracy.*

When a manual system is used in a hospital, it is often difficult to verify data unless certain steps are built in. These could include routinely performing quantitative analyses on all medical records of discharged patients with verification by a second person. However, this would affect efficiency, as it would delay the processing of the medical record. Alternatively, a regular review could be conducted. The thorough training of staff in the majority of situations, however, would assist in eliminating many errors and improve the data collection.

The medical record department personnel can verify the completeness of the medical record by undertaking a formal medical record review using a random sampling method on a regular basis.

The implementation of such a system would enable staff to review data collected, identify deficient areas, provide a system of notification to individuals responsible for the collection of data, and provide follow-up for correcting deficiencies (Abdelhak et al., 1996).

A sample medical record review form can be found in Appendix 1.

Again, this procedure could be used in a non-inpatient environment with minor adjustments to the form, so that data collected by health professionals and important statistical data can be verified.

In a computer-based patient record, most data are captured electronically with built-in checks and edits, as previously discussed, which should reduce the margin of error.

Qualitative analysis of medical records

In a qualitative analysis of medical records, the information pertaining to patient care is reviewed for accuracy, validity and timeliness (Huffman, 1990). This includes:

- reviewing medical records to ensure that all clinically pertinent data have been accurately recorded; and
- checking the front sheet to ensure that the patient's diagnosis and treatment have been recorded and are supported by documentation in the body of the medical record.

> *The quality of health care is measured by the quality of the data in the medical record.*

In an ideal situation, a staff member trained in quality assessment should perform a qualitative analysis on every medical record of discharged patients. This procedure takes a significant amount of

time and in most situations, there are insufficient staff with the time available to complete the job effectively. The administrator or manager of a health care facility is generally responsible for determining what type of record to review and how often they should be reviewed.

> *To be an effective tool in data quality, a qualitative analysis should be performed at three- to four-month intervals on a random sample of patient records. It could be on 30 or 5% of records of discharged patients, or on a specific diagnosis or procedure performed.*

In addition to the above, other procedures that could be implemented to ensure data quality are listed below.

Data entry checks

In a manual system, steps to check the entry of data in the medical/health record should be taken at the point of entry.

- Check the collection of demographic data by clerical staff in the admission office, emergency department and outpatient department reception, and health centre and clinic reception/registration area prior to the provision of health care services. The accuracy of this data is crucial for the identification of the patient during the present visit as well as their admission or future encounters with the health care service. Regular checks should be in place to prevent incorrect data being entered.
- Doctors and other health care professionals at all levels of health care should check the accuracy of the data for which he or she is responsible. In most medical/health records, data are recorded by a variety of persons, all of whom are responsible for the **accuracy of his or her documentation**. That is, the responsibility for **accurate and timely data entry rests with the professionals** involved and regular checks should be undertaken.
- When the medical/health record is returned to the medical record department (or the file room after attendance at a centre or outlying post) after discharge or death or outpatient attendance, staff are responsible for checking for completion before filing. They are responsible for monitoring the quality of the data and ensuring that health information generated from the medical record is timely, complete, accurate and accessible. The person responsible for the health record services, regardless of the type and level of health care, must manage those services in a manner that promotes quality information.

Checks on the quality of abstracted data

If a qualitative analysis is not undertaken on the complete medical/health record, a data quality check should be carried out on abstracted data. For most inpatient health care services, an abstract, that is

the abstraction of information from a document to create a summary, is prepared at the end of a patient's hospital stay by the attending doctor. It is necessary to ensure the quality of the abstract information, for accuracy, validity, completeness and timeliness.

- A staff member, other than the initial clerk should **routinely check the abstracts**. To do this, the patient's record is retrieved from the database, or manually if a non-computerized record system is maintained, and the data elements are verified. In most cases, random samples are undertaken, errors noted are corrected and documented, and the staff member responsible for the original abstract re-trained.
- In a paper record, if a doctor forgets to sign an order, the order is not authenticated and cannot be carried out until the doctor signs.
- In a computer-based or electronic patient record, the authentication of an order is captured by a key word or code, and entered by the doctor when he or she has completed the order.

There are various methods of capturing authentication data — they are referred to as electronic signatures. An electronic signature does not capture a person's actual signature into the computer, what happens is that the computer recognizes a unique code that only the author has in his or her possession. (Davis and LaCour, 2002) The computer can be programmed to reject orders that do not contain appropriate authentication. *"The computer would look for both the existence of the authentication [for data completeness] and the correct authentication [for data validity]"* (Davis and LaCour, 2002).

- This procedure can also be used when abstracting data on non-inpatient services offered at a clinic or aid post. The type of data extracted in this situation is usually of a statistical nature and should be checked for quality on a regular basis.

Audit of accessibility of medical/health records

To help assess the accessibility and the quality of the medical/health record services, study programmes should be instituted. For example, the medical record retrieval process should be studied using pre-determined criteria and data collection form (see Appendix 2).

Medical/health record audit

Similar to a medical/health record review is a medical/health record audit, which is also a retrospective review of selected medical/health records or data documents to evaluate the quality of care or services provided compared with predetermined standards. To validly assess the completeness, accuracy, consistency, and legality of the medical record, an evaluation of the adequacy of medical record content can be conducted. A sample data collection form can be found in Appendix 3.

Some steps identified by Jones (1993) following an audit of hospital medical records are listed below.

- The health facility administration and senior doctors should be asked to seek improvement in medical record documentation by assisting with development and design strategies to enhance data collection formats.

■ Provision should be made for the allocation of sufficient resources to adequately monitor data quality.

■ Support should be obtained from the administration to work with clinical departments and senior clinicians to examine strategies for the provision of adequate patient data.

■ A comprehensive training programme on documentation practices for junior doctors should be developed with the support of the hospital administration and senior medical staff.

■ Continuous auditing of documentation practices should be carried out and findings monitored regularly.

■ A multi-professional forum should be set up to address documentation and other issues and consider using a total quality management approach for improving the quality of data.

Development of an on-going quality assessment plan

If a health care facility or ministry of health are serious about data quality, they need to develop a plan aimed at improving and maintaining the quality of data and the information generated from that data.

To develop a quality plan, several structural components need to be in place. These include:

■ commitment by top-level management to support the programme, which would involve the appointment of a quality co-ordinator with adequate clerical support; and

■ staff responsible for quality control should be involved and deal with quality reports properly by reading and acting upon recommendations in a timely manner (Schofield, 1994).

Remember that data validity is a measure of data quality. Only valid data entered within a legitimate range of values for that type of data should be collected. For example, in a computer system, when collecting dates of birth, the computer is looking for an eight-digit number — if more or less, the entry is invalid.

To develop a data quality assessment plan, whether in a hospital, health centre, clinic or aid post, the administration should take certain initial steps, which include:

■ assign responsibility – **a specific staff member should be assigned to audit aspects of documentation contained in the patient's medical/health record;**

■ **identify important aspects of data collection** – such as accuracy and validity, reliability, completeness and timeliness;

■ **determine indicators of data quality** for each documentation component;

■ **set a threshold** – that is, determine an acceptable error rate;

■ **develop an organized method for collecting data** according to quality indicators previously developed;

- **assess actions** taken to improve documentation; and
- **communicate the results** of the review/audit to those affected.

Quality assessment should be undertaken to ensure health information management functions are working effectively within the standards previously determined.

Performance improvement techniques

Along with a continuous quality assessment plan, steps should be taken to institute performance or quality improvement. This is a process by which a health care facility reviews its services to ensure quality (Davis and LaCour, 2002).

- Staff responsible for health care services should be encouraged to not only meet a certain standard but should also seek to **improve their performance**.
- **Performance improvement** should not only include the staff of medical record services but also the entire staff of the facility and should be multi-disciplinary. To improve data quality, all persons involved in completing, checking and using data should be involved in insuring that the data are correct, valid, timely and relevant.
- Employees should use teamwork to improve a process. By instituting **performance improvement in data collection,** the outcome – patient care – will ultimately be improved.

A popular method of performance improvement used in the United States of America is the "Plan, Do, Check and Act" (PDCA) method developed by Walter Shewhart (Abdelhak et al., 1996). The steps are:

- **the plan phase** consists of data collection and analysis to propose a solution to a specific problem;
- **the do or implementation phase** tests the proposed solution;
- **the check phase** monitors the effectiveness of the solution over a period of time; and
- **the act phase** formalizes the changes that have proven effective in the "do" and "check" stages.

The important point to remember is that a process is being improved. Concentration should be placed on the process and not on the employee performing the job. To improve the process, **ALL** persons involved must be part of the team. As outlined by Abdelhak et al. (1996), the PDCA plan is as follows:

Plan

- Co-ordinate a team
- Investigate the problem – gather facts
- Discuss potential solutions
- Decide on a plan of action

Do
- Test the plan of action
- Educate employees on the new process
- Pilot the new process

Check
- Monitor the new process during the pilot study
- Did the plan of action work the way the team intended?
- Make necessary adjustments and continue the pilot

Act (When certain the process is an improvement)
- Change the policy
- Educate and train all affected employees
- Implement the new process

By carrying out a data quality improvement plan, the review of medical/health records provides information to patient care committees, doctors, administrators and outside agencies.

Other steps to assist with data quality improvement

- **Performance indicators** can be developed as a guide to monitor and evaluate the quality and appropriateness of care. They are reliable and could be used to detect change such as health outcomes.
- **Policies** should be defined concerning the facility's overall position on quality. The policies should reflect a commitment to the highest standard of care for patients, including accurately and competently documented demographic and clinical information, and an opportunity for input to the programme from all staff, with full support from the administration.
- A **quality review committee** should be established and charged with the responsibility of overseeing the quality activities programme.
- A **quality coordinator** should be responsible for the day-to-day co-ordination of the programme. This person must understand quality and its implementation and must be an effective communicator with the ability to impart knowledge to others. That is, a resource person who needs to network within and outside the organization.
- The **data collection system** needs to be **simple** and **user-friendly**.
- Quality activities need to **focus on practice** and not individual workers.
- **Confidentiality** needs to be maintained in all programmes.
- **Staff education** requires that all staff clearly understand what quality means to the facility, how the programme is managed, what is expected of staff and what they can expect to achieve.

Limitations in overcoming problems related to data quality

Data quality can be hampered by a number of issues, including the following.

- **Lack of uniformity of data** – without predetermined standards and uniform data sets, problems relating to the quality of health care data are difficult to solve.
- **Poorly designed data collection forms** – if forms are not well designed, the collection of data could be affected, resulting in poor quality data.
- **Limitations to doctors' capacity to communicate** – some doctors find it difficult to record data in a clear and concise manner, resulting in poor information. They also often use non-standard abbreviations and are "too busy" to complete medical records once the patient has been discharged from the facility or does not require further treatment. Limited education in documentation requirements of medical staff is a major factor in poor data quality.
- **Limitations to information transfer from different parts of the facility** – sometimes information being transferred from the laboratory to the ward or a clinic does not contain the correct patient's name and medical/health record number. Such errors make it difficult to ensure that all data pertaining to an individual patient are filed in that patient's medical record. The transfer of information from one department to another or from a hospital to a clinic or aid post is often slow or information is lost in transmission.
- **Limited education of processing staff** – the processing of medical/health records requires staff that can understand the need for accuracy and completeness. If they are not properly trained, the production of quality data is threatened.
- **Lack of planning** by administrative staff to ensure data quality control programmes are in place. All data collection and abstracting staff should be properly trained; and doctors should be educated in the requirements for accurate and timely documentation of patient care details.
- **No single record** – a problem of quality arises if more than one medical record is kept on each patient. Some facility staff, such as in cardiology, oncology and social work, insist on keeping their own records, thus limiting the overall collection of meaningful data about an individual patient.
- **Data discrepancies** – arising when errors occur at the point of collection and plans are not in place to check the entry and verify the data.

Summary

For the quality of data in health care delivery to be improved, administrators need to take the following steps:

■ **Policies** should be in place to cover all staff on data collection requirements**.** If the administration is committed to data quality, it needs to provide resources to establish a formal monitoring programme to ensure quality control. For example, if carrying out a re-abstracting study, a baseline measure of data quality needs to be established.

■ **Education** – programmes should be offered for medical, nursing and other health care staff as well as medical record staff on the importance of careful collection of original source data. Education could also be aided by the creation of a videotape on data quality for staff, including doctors, and the development of simple data quality guidelines.

■ **Tools** – sufficient equipment and technology support should be available for all staff. Access to a computer for patient identification and other data collection requirements would provide staff with the ability to edit work as it is being done and greatly reduce errors.

To ensure documentation meets the required standards previously determined, quality assessment, record review and improvement procedures should be developed. Poor quality data are a major hindrance to planning and decision-making, therefore, data quality is an important concern for health care delivery at anytime in any part of the world.

References

Abdelhak M, Grostick S, Hankin MA, Jacobs E. *Health Information: Management of a Strategic Resource.* Philadelphia, WB Saunders, 1996.

Australian Council on Health Care Standards. *The EQUIP Guide: Standards and Guidelines for the Evaluation and Quality Improvement Programme* (2nd edition). Sydney, ACHS, 1998.

Australian Institute of Health and Welfare. *National Health Data Dictionary*: Version 10. Canberra, AIHW, 183, 2001.

Davis N, LaCour M. *Introduction to Information Technology*. Philadelphia, WB Saunders, 2002.

Donabedian A. Quality Assessment and Assurance: Unity of Purpose, Diversity of Means. *Inquiry*: Vol. 25. 173–195, 1988.

Huffman E. *Manual for Medical Record Librarians*. Chicago, Physician's Record Company, 1963.

Huffman E. *Medical Record Management* (9th edition) (Finnegan R, Amatayakul M, eds.). Berwyn, Ill., Physician's Record Company, 1990.

Johns ML. Health Information Management Technology: An Applied Approach. Chicago, American Health Information Management Association, 2002.

Jones L. *The Quality of Clinical Documentation and Subsequent Effect on DRG Assignment: A Report on the Findings of the DRG Documentation Study*. Canberra, Case Mix Development Programme, Commonwealth Department of Health, Housing and Community Services, 1993.

Joint Commission on Accreditation of Healthcare Organizations. Comprehensive Accreditation Manual for Hospitals. Oakbrook Terrace. Ill., JCAHO, 1995.

Joint Commission on Accreditation of Healthcare Organizations: First Edition of the Joint Commission International Accreditation Standards for Hospitals, 2000.

Schofield J. Developing a Quality Plan. *Health Care Quality Activities in Practice in Australia*, 4th Edition. Sydney, Australian Council on Healthcare Standards, 7–15, 1994.

3 Quality in data collection

The collection of the original source data must have in-built procedures to ensure that they meet the required quality standards.

Building quality into the process of data collection right from the beginning is one of the major foundations of the overall move to data quality control.

The main source of health care data is the patient's medical/health record, which should contain essential data used in the health care decision-making process. It is the "who, what, when, where, why and how" of patient care (Abdelhak et al., 1996). A patient's medical/health record should provide accurate information on:

- who the patient is and who provided health care;
- what services were provided;
- when and where the services were provided;
- why the services were provided;
- how effective the services were; and
- what the outcome was of care and treatment.

Another important collection of data by hospitals is the bed census, which is collected daily and processed monthly and annually to produce statistics on the utilization of hospital services.

Health care facilities, regardless of type or size, collect data processed as statistics for specific external reporting to meet the needs of health care authorities, local and national governments, and funding organizations.

Statistical data are collected from clinics and outposts on the number, type and age of patients attending the facility, the types of conditions treated, tests conducted and referrals.

The collection of the original source data must have in-built procedures to ensure that they meet the required quality standards and also the requirements of the patient, health care provider, health care facility and government.

Type of data collected and how collected

The collection of accurate and timely data has a significant impact on the efficiency and effectiveness of the decision-making process in health care.

> *Data quality should*
> *start at this point.*

The type of primary health care data collected about a patient in either a manual medical/ health record or computer-based or electronic patient record at all levels of health care fall into two major categories (Abdelhak et al., 1996).

> *Administrative data*
> *and clinical data.*

Administrative data

Administrative data – can be divided into four sections: demographic, legal, financial and provider (Abdelhak et al., 1996).

- **Demographic and socioeconomic data** – personal data elements about a patient. It is essential that these data are accurately recorded at the **initial point of contact of the patient** and are sufficient to **identify an individual patient**. Demographic data include the name of the patient, sex, date of birth, place of birth, the patient's permanent address and a unique personal identifier. This is usually a number, whether for national identification, social security or another piece of information that will uniquely identify the patient.

- **Legal data** – should include a **signed consent** by the patient for health care services at a hospital or primary health centre and a signed and dated authorization for the **release of information**. A signed and dated authorization for specific procedures and treatment is also required. In addition, in the event of the patient's death, an authorization for an autopsy signed and dated by the patient's next of kin should be in the medical record.

- **Financial data** – data relating to the **payment of fees for services** rendered should include the **primary source of payment**, name and identification number of private health care insurance, if applicable, as well as the national health insurance number.

- **Data from the provider, that is the doctor or other health care professional** – the data collected on the **person providing care** should include the name of the principal medical officer and other health care workers providing services to the patient.

The above data are usually contained in the first section or pages of a patient's medical/health record. The rest of the record should contain clinical data, which are those elements specific to the patient's present and continuing care.

29

Clinical data

Clinical data on a patient whether admitted to a health care facility or treated in a primary health care environment must also be accurately recorded and can be divided into the following segments:

- **Main condition or principal diagnosis, other conditions or problems and surgical procedures** – this segment should contain an up-to-date summary of diagnoses and procedures performed, recorded on the first sheet of the patient's medical/health record, below the demographic data. It is dated and signed by the attending health care provider.

- **Patient medical history** – this should contain a detailed history of the patient's past health/illness including social history and habits. This section should also include a family medical history, presenting signs and symptoms, and the history of present illness or problem, recorded in the patient's own words.

- **Physical examination and assessment** – this segment should include the results of a physical examination with findings and objective observations recorded, along with a provisional or working diagnosis, signed and dated by the attending health care provider.

- **Orders and treatment plans** – these should include detailed data on the orders and treatment plans, with date and time of entry, and signature of the person writing the order, duration of the order, date and time of treatment plan, and results.

- **Diagnostic tests** – data in this segment should include the name of tests and/or X-ray, date and time the test was performed, reported findings and signature of the person completing the report, and date and time result was reported.

- **Medications** – this section should contain the name of the medication, date and time of order; dose and medication instructions; signature of the person ordering the medication; and for inpatients, the date and time when administered and signature of the person administering the medication.

- **Progress notes** – this segment forms the bulk of the medical/health record both for inpatients and patients treated at primary health care level and should give an ongoing chronological picture and analysis of the clinical condition of the patient. It serves as a means of communication and interaction between health care providers. All entries should be dated and signed at the time of entry.

- **Nursing notes** – these should contain data recorded by nurses during the continuing care of an inpatient, including a patient care plan, graphic charts for blood pressure, temperature and respiration, special observation charts, and signatures of attending nurses, plus date and time of each entry.

- **Operative procedures** – this segment should contain significant data elements relating to all procedures performed in an operating room or day surgery for diagnostic, exploratory or definitive treatment. Anaesthetic and recovery room reports should also be included and signed and dated by the attending surgeon and/or anaesthetist.

■ **Disposition (completion of episode of care, discharge or death)** – concluding data should be recorded on the completion of an episode of care at a clinic or aid post, or discharge or death of an inpatient. Such data should contain the date of completion of the service or time of discharge or death. If discharged, prognosis and follow-up instructions should be clearly stated along with a final diagnosis and appropriate signatures. In the event of the death of a patient, if an autopsy has been performed, a full autopsy report should be included.

Data collection can be carried out on paper or in a computer. In a computerized database, patient information is generally organized in a systematic, predetermined format. The collection of data in a manual medical record should also be systematically collected in a predetermined format.

Non-inpatient data

Data collected for an emergency patient, outpatient, clinic or aid post medical/health record should include:

■ patient identification as for inpatients;

■ relevant history of presenting illness and physical findings;

■ clinical observations;

■ reports of tests and procedures performed, such as immunization, health screening, etc.;

■ the outcome of the visit. For example, follow-up for further treatment, admission to hospital, no further treatment, etc.;

■ growth chart for children;

■ referral information – correspondence from a local doctor or community nurse; and

■ the signature of the health care provider seeing the patient to indicate their responsibility for the written information.

The same basic information should be collected regardless of the size or type of health centre or clinic.

The arrangement of the information should be convenient for those who must refer to it on a daily basis.

Collection of data in specialist outpatient clinics

In many countries, specialist clinics are held for patients who need to see a consulting specialist for a specific condition.

■ A specialist outpatient is often a patient with a chronic problem (hypertension, diabetes, etc.), a paediatric patient or a recent inpatient. There should be an appointment book for making appointments for each specialist.

- On the day of the clinic, the appointments should be noted as attended or did not attend. This information is needed to measure the workload of each clinic and determine the number of appointments made and not kept.
- At the end of the month, the number of patients who ATTENDED and DID NOT ATTEND should be counted for each clinic and included in the monthly report.
- Other statistics would be collected in the same way as for the general outpatients.

Patient-held health record

In some countries, outpatient medical records are not kept by the health care facility. In these situations, the health care worker documents the visit in a PATIENT-HELD HEALTH RECORD. The patient-held health record can consist of the maternal/baby health record, or patients can be asked to purchase an exercise book (sold by the hospital). The use of patient-held health records reduces the huge daily filing problem for general outpatient records. Problems associated with using PATIENT-HELD HEALTH RECORDS often outweigh their usefulness. Some of these problems include:

- the patient does not bring the health record to the outpatients;
- the health record has been lost; or
- the health record has been tampered with.

In addition, it is difficult to monitor the quality of such health records, so education on quality requirements for health care personnel writing in the record is essential.

Hospital census data

In addition to data collected in the patient's medical/health record, health care facilities also collect daily census data on inpatients, emergency patients and ambulatory patient attendance. These data must also be accurate and reliable to maintain the quality of information generated.

The type of inpatient data that should be collected by nursing staff is recorded on the daily bed census and processed by the admission office staff or a member of the medical record staff and should include the following:

- **daily inpatient census** – the count of all inpatients at the census taking time each day, that is, the number of inpatients present in the facility at a given time. Patients admitted and discharged the same day are also included;
- **daily number of admissions;**
- **daily number of transfers** in and out of wards; and
- **daily number of discharges and deaths**.

It is essential that the information is accurate and reliable so that it can be used by the facility with confidence to monitor the volume of patients treated on a daily, weekly, monthly and yearly basis, and to plan future services.

Secondary health care data collection

Abstracted data from the daily inpatient census are used to:

■ accurately compile **monthly and annual statistics** such as the length of stay of inpatients and the inpatient occupancy rate; and

■ calculate the gross and net death rates, autopsy rates, maternal, foetal, newborn and neonatal death rates, hospital infection rates, and others as required.

The accuracy of the information is important, as the facility uses it to review the utilization of health care facility services, the workload of doctors, nurses, and other staff, and to plan for future development and human resource management.

The facility, whether a hospital or primary health centre, must also abstract data:

■ on diseases treated and procedures performed using a classification system to complete government required morbidity statistics;

■ on vital statistics (births and deaths) for the production of national death and birth rates. Standard birth and death certificates should be used for the collection of these data; and

■ for special registries related to patients with a specific diagnosis, condition or procedure, for example, diabetes, HIV/AIDS, birth defects, infectious and contagious diseases, and organ transplants. Registries may be developed for any type of disease or condition for which specific data are collected and forwarded by the facility, local doctor or outlying clinic to the registry.

Cancer registries have been operating in many countries for some time and rely on accurate data being collected and forwarded by the facility or person treating a patient. Information gathered by such a registry has been used, and is still being used, to identify trends and changes in the incidence of cancer within the population it serves.

Most governments require facilities offering health care services to forward accurate and reliable morbidity, mortality and utilization statistics to the appropriate government department for processing. From this pool of information, nationwide statistics are produced to enable comparison between facilities, states or provinces, and from country to country.

Causes and sources of poor data collection

Whether collecting data to be stored in a paper medical/health record or in a computer-based or electronic patient record or for statistics or specific registries, data must be accurate, reliable and organized in such a way that they are understood and health information can be retrieved.

The first step should be to determine what data are needed and how they are to be collected.

Poor data collection occurs when data are not collected in a logical sequence, and when the instrument used to collect the data is deficient.

Some causes and sources of poor data collection are:

- **poorly designed data collection forms** with no logical sequence;
- **inefficient clerical staff, lack of training** in interviewing patients and recording details, and lack of understanding of the need for accurate data collection. Although trained in procedures, staff may not collect **ALL** the required information at the time of first attendance. They also may not understand the consequences of such action;
- **lack of professional judgement** by health care providers when recording data about a patient and his or her treatment;
- **delay in recording data** – when data are not recorded at the point of contact with a patient; and
- **lack of understanding of requirements of data collection and data quality** by medical officers, nurses and other health care professionals.

Examples of how inaccuracies in data collection arise

Major factors prevent clerical staff and health care providers from completing medical documentation adequately. These include:

- **lack of time** caused by pressure of work, poorly trained and insufficient staff. Pressure of work and poorly trained personnel play a major role in creating data collection problems. With many health care professionals under pressure to provide maximum services at minimum cost, documentation is not always a high priority. Clerks also have many pressures on them during the day, for example the number of patients waiting for outpatient/clinic registration or inpatient admission, or the number of medical records waiting completion, which may cause them to take "short cuts" or make errors in data collection;
- **limited resources** available for staff education and the implementation of data improvement measures;
- lack of understanding **as to the affect that inadequate data collection has on patient care and the quality of information generated by the facility;**
- **poorly abstracted data** from the original source;
- **insufficient information** is available at the time of attendance or admission to the health care facility;
- the **patient is unsure** as to whether he or she has been to the hospital previously and clerical staff do not thoroughly check the Patient Name Index, which could result in the duplication of patient registration and medical records and inability to find earlier information that could be important; and
- invasive procedures are ordered for which **no rationale or sufficient reason** can be determined from the data recorded in the medical record.

It is important to understand that poor, inaccurate and incomplete documentation has a substantial impact on data quality.

It also has an impact on the process of health care and has potential financial consequences for the health care facility.

Ways to overcome problems of poor data collection

Clinical decisions are only as good as the information on which they are based. To improve documentation in health care and subsequently the overall quality of care:

- data collected at all levels must be of an **appropriate standard**;
- procedures should be developed at all levels of health care to **monitor data collection**;
- methods should be used that **improve the quality of the collection** and avoid incorrect data being recorded;
- a programme should be developed to **educate medical and other health care professionals**, as well as clerical staff, on the need for data quality improvement;
- administrators and medical staff should give **high priority to** such a programme; and
- **support and commitment** must come from the top and there should be a systematic approach to quality rather than introducing ad hoc measures without a specific plan.

The achievement of quality, specifically data quality, takes time and commitment and will not occur unless there is careful planning that addresses problems as and when they occur.

More specifically, to improve data quality at the point of collection, the following steps should be taken:

Review of data collection forms

- Review data collection forms and redesign as required.
- Design forms to collect data in a logical sequence.
- Maintain simplicity of design.

Staff requirements and training

- Employ sufficient staff to meet the needs of the data collection area.
- Train clerical staff on requirements and importance of accurate data collection. This should include educating them in why data quality is important and the consequences of poor quality.

■ Educate all health care providers on the importance of timely and accurate documentation of patient care data, emphasizing why it is important and the consequences of poor data quality.

■ Information should be recorded at the time of consultation or procedure, or as soon as available.

■ Each entry should be clearly dated with the date and time where necessary.

■ The name and designation of the health care provider performing the task or consultation should be clearly legible and each entry should be signed and dated.

■ The treating health care provider should periodically review the record for correctness.

■ Alterations or deletions should not be made – if the original entry is incorrect it should be lined through so that the original entry remains readable. Such action should be explained and signed by the person making the notation.

Standards and checks

■ Develop and maintain standards for data collection.

■ Include edit checks in the computer system.

■ Institute routine checks on data for accuracy, validity, reliability, legibility, completeness and accessibility in both manual and computerized systems.

■ Institute an annual review of data collection procedures.

Improving data collection and documentation

Data collection and documentation in medical/health records can be improved by ensuring the following:

■ All patient identification data are accurately recorded on the first sheet of the medical/health record and the patient's name and medical/health record number are clearly shown on subsequent pages.

■ The main condition and other diagnoses, problems and procedures are clearly written on the front sheet, along with the signature of the attending health care provider.

■ The history of past and present illnesses/problems is recorded clearly, and the entry dated and signed.

■ Consent forms are signed, dated and witnessed.

■ Progress notes, whether for an inpatient or outpatient, are recorded daily or each time the doctor sees the patient and are clearly written, legible, signed and dated.

■ For surgical patients, either as an inpatient or at a day surgery, operation forms should be completed with all relevant information, as well as anaesthetic forms and recovery room report, signed and dated.

■ Nursing notes for inpatients should be completed daily, written clearly, and each entry dated and signed.

There should be an audit trail for electronic records. An audit trail is a programme that records the access and/or action that occurs in a computer record by logging the user identification and the file identification, recording date and time of access and action done (Davis and LaCour, 2002).

Why some data are collected and not used by planners, managers, etc.

To be useful, data must not only be correct, legible, reliable, complete and accessible but also organized in an appropriate form. If the data are deficient in any way, the information generated is not useful and may not be used by planners or managers because it cannot be relied upon to give the necessary information. In addition, if data are collected and the meanings are not standard, the information may be useless to the health care professional, patient, health care facility and public. That is, if the characteristics of the data are not up to the required standard, confidence in the information generated is diminished.

In addition, if the data collected are not current, that is, recorded at the time the patient is seen or treatment given, their usefulness is again diminished. Also, data could be accurate, complete, current and timely but not accessible. That is, if the medical/health record is not available, the information is useless and cannot be used for the purposes for which it has been collected.

> *Accessibility is an important component of data quality.*

Over the years, many health care facilities have collected data on patients in the form of medical/health records but have not developed up-to-date systems for maintaining the quality of the services provided. All too often, medical/health records are not available when needed and therefore not used.

Summary

It is important that all personnel, including clerical staff, medical officers, nursing staff, other health care professionals and administrative staff understand the standards and regulations governing the collection of data and adhere to them at all times.

Timely and accurate data collection will help the person responsible for patient care to ensure that quality of care is at the highest level possible. It will also help the administrative staff to plan for the future. Pressure of work often causes people to make errors or not collect all the data required. The continual monitoring of data collection is therefore vital.

References

Abdelhak M, Grostick S, Hankin MA, Jacobs E. *Health Information: Management of a Strategic Resource.* Philadelphia, WB Saunders Company, 1996.

Davis N, LaCour M. *Introduction to Information Technology*. Philadelphia, WB Saunders Company, 2002.

Johns ML. Health Information Management Technology: An Applied Approach. Chicago, American Health Information Management Association, 2002.

4 Clinical coding quality

The clinical coding process begins with the source document, which is the medical/ health record, from which pertinent clinical information relating to a patient is abstracted.

The comparison of health care data at all levels of health care services, states or provinces, within a country or between countries is vital to the growth and dissemination of health information throughout the world. This possible sharing is meaningless, however, without the use of standardized disease classification systems and accurate and reliable statistics. As a by-product of patient care, quality health information has value in the review of clinical services and as a management tool for both the health care service and the government.

Medical/health records are coded to enable the retrieval of information on diseases and injuries. This information is used at a national level for planning health care facilities; for determining the number of health care personnel required; and for educating the population on health risks within their country. It is used at an international level to compare the health status of countries in a region or globally.

Defining clinical coding

Coding generally is defined as "classifying data and assigning a representation for that data" (Abdelhak et al., 1996). Clinical coding is the assignment of a specific code to a narrative statement of diagnoses and procedures. Such coding was often considered limited to inpatient information and usually undertaken within a health care facility, but it is in fact carried out at most levels of health care. Whatever

classification system is used, similar diseases are grouped together under a single code. For the coding of health care data and production of morbidity and mortality statistics, most countries in 2002 use the *International Classification of Diseases and Related Health Problems*, 10th Revision (ICD-10), produced by the World Health Organization (WHO, 1993), or an adaptation. In a hospital, the code is entered into a manual or automated index system for statistical analysis and subsequent retrieval of medical records for research on specific diseases or procedures (Amatayakul, 1996). In primary health care, codes are usually recorded on a statistical data collection form to enable the collection of local and national non-hospital morbidity statistics.

In hospitals, health care data collected provide government authorities (e.g. ministry of health) with information required to not only review the services of all hospitals under their control, but also to plan for the future. In addition, the use of a disease classification system at primary health care level enables the government to collect data on the health status of the community and provide detailed national health statistics. In some countries, the ministry of health determines whether hospitals are required to supply information only on the main condition or on all diagnoses treated and procedures performed.

A decision is also made in each country whether to code using three-digit or four-digit codes from ICD-10. This decision should be made by a health statistician or epidemiologist in consultation with the ministry/department of health, and will be based on the level of specificity needed.

If the government has determined that only the principal diagnosis is to be coded, the person coding should find the code number for the main condition and record it on the FRONT SHEET in the correct place or on the statistical collection form in the case of primary health records. The definition of main condition varies from country to country and coders should check to ensure that they are using the correct definition.

As in all data collections, the quality of clinical coding is important, whether the main condition only is coded as for primary health care or all conditions treated are coded.

Steps need to be taken to:

■   identify problems, including causes and sources of poor coding; and

■   implement systems to improve the quality of coding, and subsequently the quality of information derived from the coded data.

Importance of accurate and timely coded clinical data

The clinical coding process begins with the source document, which is the medical/health record, from which pertinent clinical information relating to a patient is abstracted. The importance of using coded clinical data for reporting, compiling and comparing health care data cannot be disputed. The data compiled are needed for:

■   internal and external reviews of the appropriateness and timeliness of health care;

■   planning health care delivery services;

■   analysing payments for health services; and

■   conducting epidemiological and clinical and health service research.

The quality of these data should be given the attention they deserve. Staff responsible for the management of health information must be prepared to do their best to ensure data quality in the services they provide.

> *To be able to compare data between health care facilities, states, provinces and countries, it is essential that disease classification systems and coding procedures are standardized.*

Over the years, standardization of coding has tended to pose some problems as users of the information often have conflicting needs. For instance, medical researchers in hospitals require highly specific data on disease and surgical procedures, often requiring a large selection of codes. Health care planners, on the other hand, use disease classification data for statistical and evaluative studies, service reviews, and demographic and epidemiological studies. To enable meaningful analysis, they require diseases to be grouped into specific categories.

To be really useful, a disease classification system must reconcile the needs of medical researchers and health-care planners. Attempts have been made to do just that in the production of ICD-10 and recent adaptations of ICD-10.

Clinical coding procedure used to ensure quality

Whether undertaking simple coding for primary health care services or for more sophisticated hospital health care services, a thorough knowledge of the classification systems' key components of data quality – accuracy, validity, reliability, completeness and timeliness – are important (Hyde, 1992).

Each clinical coding component consists of:

■   selecting the condition or procedure to code; and

■   assigning a code and sequencing codes.

> *Accurate diagnostic and procedure coding cannot be achieved, however, without clear and complete documentation in the medical/health record.*

The clinical coding procedure should be composed of three parts:

- the analysis of the medical/health record to determine what items should be coded;
- the selection of the main condition/principal diagnosis, associated/other diagnoses (if applicable) and surgical procedures; and
- the allocation of correct codes.

These three activities are not independent and a **thorough reading of the medical/health record** is necessary to bring to light evidence regarding the patient that may effect the choice of codes.

Analysis of the medical record

The first step in the clinical coding procedure is to analyse the medical/health record for completeness and accuracy. As a minimum, the coder should:

- review the front sheet;
- read the discharge summary or abstract (if applicable), and review the operation report and histopathology report for tissue removal (if applicable);
- find out whether the main condition and associated diagnoses and surgical procedures have been recorded;
- check that the diagnoses are supported within the record with necessary evidence such as pathology reports, cultures, etc., to identify the bacteria or virus responsible for an infection such as pneumonia or gastroenteritis;
- check X-ray reports – to specify site and type of fracture/condition;
- review progress notes – to determine the main condition/principal diagnosis if not clear on the front sheet or discharge summary; and
- for hospital coding, check previous admissions, to ensure that they have been coded.

Selection of main condition

The term used by WHO to code the reason for health care intervention is "**main condition.**" However, many countries refer to the main condition as the "principal diagnosis." For these guidelines, the term **"main condition"** will be used. The selection of a "main condition" is important, not only in patient care but also in the collection of morbidity statistics and should be determined by the medical officer or health care worker responsible for the patient's care and/or treatment. The "main condition" as defined by WHO is as follows:

Main Condition

The condition, diagnosed at the end of the episode of health care, primarily responsible for the patient's need for treatment or investigation. If there is more than one such condition, the one held most responsible for the greatest use of resources should be selected. If no diagnosis was made, the main symptom, abnormal finding or problem should be selected as the main condition (WHO, 1993).

Allocation of correct codes

Whether coding at hospital or primary health care level, the following steps are required to ensure accuracy (WHO, 1993).

■ identify the type of statement to be coded and refer to the appropriate section of the alphabetical index;

■ locate the lead term in the diagnosis, i.e. the noun in the diagnosis, which is the name of the disease, NOT the site of the disease. In ICD-10, however, some conditions expressed as adjectives or eponyms are included in the index as lead terms. The coder must read and be guided by any note that appears under the lead term. For example: *Influenza – there is no specified virus*;

■ locate the lead term in the alphabetical index (Vol. 3) and then check the list of modifiers for any remaining words (adjectives) in the diagnosis. Carefully follow any cross-references found in the index. For example: *Influenza, virus not identified (J11)*;

■ refer to the tabular List (Vol. 1) and verify the code given in the index with the entry in the tabular list. In this step, the coder is checking for exclusion notes that may change the decision to use a particular code. For example: *Influenza, virus not identified* – this excludes *Haemophilus influenzae: infection NOS (A49.2): meningitis (G00.0): pneumonia (J14); and*

■ assign the code.

> *Specific basic guidelines* for coding are included in Vol. 2 of ICD-10 and coders should study and follow them to assist with ensuring coding quality. In addition, encoding software is available to assist the coder in the allocation of correct codes.

What to code?

The number of diseases and surgical procedures to be coded varies from hospital to hospital and country to country.

■ Large teaching hospitals often need to collect detailed information on the types of diseases treated for research and teaching purposes. Most of these hospitals require staff to code all diseases **TREATED** during admission and every operation performed.

■ Small hospitals may choose to require staff to code only the main condition on each admission.

■ At primary health care level, staff are usually required to code only the main condition or reason for attendance at the clinic or aid post.

■ WHO recommends that whenever possible, the doctor should be encouraged to list separately other conditions or problems dealt with during the episode of health care. "Other conditions" are defined by WHO as:

Other conditions

Those conditions that coexist or develop during the episode of health care and affect the management of the patient (WHO, 1993).

> *It is important to remember that conditions related to an earlier episode that have NO BEARING on the current episode should NOT be recorded and NOT coded.*

In the guidelines for coding outlined in Vol. 2 of ICD-10, the authors stress that *"the main condition and other conditions should have been recorded by the responsible health care practitioner, and coding is therefore usually straightforward, since the main condition stated should be accepted for coding and processing, unless it is obvious that the guidelines given have not been followed"* (WHO, 1993).

> *Remember: as a minimum, the main condition or reason for visit, and if applicable, the principal operative procedure should be coded.*

Mortality coding

In the event of the death of a patient, the same procedure should be followed. However, as well as the immediate cause of death, the "underlying cause of death" should be coded as this is used in the collection of mortality statistics. The WHO definition for the underlying cause of death is:

Underlying cause of death

(a) The disease or injury which initiated the train of morbid events leading directly to death; or

(b) the circumstances of the accident or violence which produced the fatal injury (WHO, 1993).

Importance of coding quality

Coding, however, is not an end in itself and the quality of coding is essential to be able to:

- retrieve accurate information on diseases for research by indexing the coded information. In hospitals, a disease and operation index lists diseases and conditions according to the code numbers assigned. When disease and operation codes are indexed, either manually or electronically, into a disease or operation index, the accuracy of the transfer of data is essential;

- determine the distribution patterns of disease within a health facility, community, city or country and for other research and epidemiological study;
- retrieve information for the collection of health care statistics. In many developing countries, disease and operation indexes are not maintained and only statistical data are collected. For the collection of statistics, the main condition and principal operation code numbers are recorded on statistical forms, along with other data, to enable the collection of information relating to morbidity and mortality.

As in all data collection, maintaining the quality of data being recorded is essential to be able to carry out these functions effectively.

Accurate and reliable coded diagnostic and procedure data are used to:

- report the type of diseases and causes of death occurring, within a health care facility, a local community and a country;
- retrieve medical records for essential clinical research and education of health care professionals;
- profile doctor and other health professional practice patterns contributing to licensing and re-licensing as members of a specific health profession;
- evaluate outcomes of patient treatment episodes for quality improvement;
- compile statistical data for health care policy formulation in areas such as funding new services, recommending practice protocols and targeting research;
- supply available data of expert systems for health care decision-making; and
- analyse a health care facility's case mix to enable informed decisions to be made on the need to introduce new products and services and discontinue products or services that are financial loses for the facility.

Case mix refers to the type of cases treated within a health care facility based on characteristics such as resource consumption, diagnosis or procedure, or reason for visit.

Accurate and reliable diagnostic and procedural coded data are used in some countries to assign a Diagnosis-Related Group (DRG) code as a key for financial reimbursement of health care providers.

A DRG is a patient classification scheme that provides a means of relating the type of inpatients a hospital treats (that is, its case mix) to the costs incurred by the hospital. It can be further defined as a system that classifies acute episodes of inpatient care into clinically meaningful groups with similar resource consumption.

Causes and sources of poor data quality in clinical coding

Data quality in clinical coding is essential and causes and sources of poor coding quality relate to a number of issues, which include:

The medical/health record

- **missing information** – resulting in the selection of an incorrect main condition as the medical/health record is incomplete;
- **incomplete information** or incorrect data about patients, for example age and date of admission;
- **poorly documented information** – where diagnoses and operations are not validated by content in the medical/health record; and
- **use of non-approved abbreviations** by health care workers leading to an incorrect decision made when selecting clinical codes.

The health care provider

- **poorly documented or incorrect information** – health care providers may incorrectly record the main condition, problem or diagnoses and/or procedures, for example recording an incorrect site of a fracture. The attending health care provider may not understand the local definition for main condition or requirements for recording data and therefore could make an incorrect entry. Or they may record an incorrect reason for the visit/attendance, which is not supported by the details in the medical/health record;
- **incomplete information** – the health care provider responsible for the care of the patient and for completing the medical/health record does not record all information as required within the pre-determined time frame; and
- **lack of knowledge** by health care providers of what codes are used for and the importance of accurate documentation.

The coder

■ **failure to review the entire medical/health record** when coding – the coder only uses details on the front sheet, the discharge summary or abstract and does not review the whole record. Too much reliance is placed on the diagnosis/problem recorded on the front sheet. Without reviewing the entire medical/health record, coders may also find it difficult to identify the main condition, which has not been correctly identified by the attending health care provider;

■ **coding from the front sheet only** and not verifying the diagnosis by reviewing the entire medical/health record;

■ **cheat sheets** – use of homemade lists or "cheat sheets" of codes instead of correctly using the coding books;

■ **selection of incorrect main condition** – misapplication of coding rules and the selection of the incorrect main condition or reason for visit. In morbidity coding, a frequent error is the selection of the underlying cause of the disease rather than the main condition. When multiple diseases are recorded, coders may not be able to identify the main condition;

■ **resequencing** – changes in the order of diagnoses. Substituting a secondary diagnosis for a correct main condition, which changes the case mix group into which a patient is categorized;

■ **lack of training** – for coders, particularly in relation to coding standards and the classification system;

■ **lack of managerial and clinical support** and recognition of coders' skills;

■ **basic human error** – recording the incorrect code and entering inaccurate codes into the database or on the front sheet of the medical/health record or abstract form;

■ **environment for the coder** – inadequate space to work, poor light and insufficient number of coding books;

■ **local variation in coding practice** within a facility and/or across the country;

■ **lack of knowledge** by coding staff of what codes are used for and the importance of accurate coding;

■ **time limitations** – pressure of workload and pressure to code "quickly" at the expense of quality;

■ **limited reference books** – insufficient number of reference books for use by coding staff, e.g. medical dictionary.

The classification system

■ **errors or problems in translation** when ICD has been translated into another language. For example, in the Chinese translation of ICD-9, it was found that some English terms needed to be translated into several Chinese expressions (Liu and Ma, 1992). This meant

that coders did not know which one to use in Vol. 2 of the Chinese version, resulting in coders having to try to look for different lead terms;

- **definitions and coding guidelines** not clearly defined;
- **the classification books are cumbersome** to use and there is a tendency to code only from the alphabetical index rather than from both the index and the tabular list.

General

- **local variations and differences in terminology/language** over time in how a term is defined;
- **health care provider not educated** in the need to be explicit when recording diagnoses, procedures, problems, etc.

Methods that could be implemented to improve data quality in clinical coding

Development of clinical coding standards

As with the collection of all health care data, the development of standards for clinical coding is essential.

Coding standards help ensure that providers, clinical coders and users of the data all apply the same definitions, guidelines and rules. Coding guidelines will vary between countries, therefore national coding standards are essential.

Many countries have already developed their own set of clinical coding standards. An example of Australian Coding Standards for ICD-10-AM (NCCH, 1998) for hospitals developed by the National Centre for Classification in Health is as follows:

Suspected conditions

If a diagnosis documented at the time of discharge is qualified as probable, suspected, possible or any other qualifying expression indicating uncertainty about the final diagnosis, then the following steps should be taken:

(a) Refer the case to the clinician to establish whether a definitive diagnosis can be made in preference to the query.
(b) If the clinician confirms that the final diagnosis is uncertain, assign a code based on the following criteria.

If investigations were undertaken but no treatment for the suspected condition was initiated, assign a code for the symptom(s), e.g. Patient admitted with a headache. The patient is discharged with a diagnosis of meningitis – investigations did not confirm meningitis and no treatment given – **code as main condition: Headache.**

If treatment is initiated, and investigation results are inconclusive, assign a code for the suspected condition. **Code main condition: Meningitis, unspecified**.

(ICD-10-AM Vol. 5, Australian Coding Standards, 1998).

Development of coding quality indicators

In addition to coding standards, coding quality indicators have been found to be an excellent tool for maintaining data quality in clinical coding.

> *A coding quality indicator is a measure of a particular aspect of coding practice. In other words, it is an objective measure in quantitative terms of coding quality.*

Coding quality indicators are designed to be flags that provide an alert to possible problems with coding practice, which can then be further investigated, using routine quality management techniques. They do this by providing a method for comparison between facilities, states, provinces and countries and, as such, act as benchmarks.

Coding quality indicators need to be:

■ **internationally applicable**, that is, to be able to be applied if working with versions of classifications such as ICD-9, ICD-9 CM or ICD-10;

■ **measurable and timely** in the ability of results to affect or change coding practice;

■ able to **concentrate on issues that affect a high volume of cases** and tie in with WHO health goals with regard to morbidity, mortality and disease prevention;

■ able to **focus on diagnostic coding,** as procedure classifications employed differ markedly around the world.

To prepare an indicator, one needs to look at:

■ the **TYPE** – whether a general or specialty-specific indicator;

■ the **AREA** – the major area of coding practice being addressed;

■ the **INDICATOR TOPIC** – a more specific aspect of the area of practice;

■ a **RATIONALE** – the reason and purpose for the development of the indicator; and

■ the **THRESHOLD** – levels of acceptability or unacceptability need to be developed.

Performance Indicators for Coding Quality (PICQ) developed by the National Centre for Classification in Health in Australia (NCCH, 2001) is comprehensive and includes a detailed training programme.

Other methods for improving data quality in clinical coding

As in all data collections, elements that must be evaluated when looking at the quality of clinical coding include accuracy, reliability, completeness and timeliness. Steps to overcome problems in coding quality are listed below:

- **establishing a time frame for completing the medical/health record** to help ensure that the case is fresh in the mind of the health care provider;
- **revising the structure of the front sheet** to include identified spaces for recording the main condition and other diagnoses/problems as well as definitions of what each of these terms means;
- **developing a policy** that specifies the requirements/standards for clinical documentation;
- **educating health care providers in documentation requirements**. All health care workers should be educated about the importance of coding and the reliance on the preparation and maintenance of an accurate and timely medical/health record. As mentioned previously, one of the main problems in obtaining coding quality is poor documentation. A set of documentation guidelines should be prepared and health care providers should be encouraged to use them. Examining the process may mean finding out from professional staff just what makes it difficult for them to present information in the required format. If the way information is presented in the source document – the medical/health record – is preventing data being reliably and consistently collected, then examining the process may mean looking at ways to change the medical/health record;
- **educating coding staff**. Education sessions for coders should be planned in addition to basic coding instructions, which can be offered through conducting coding workshops, independent study or more formal training programs. Coders should also learn anatomy and physiology, and medical terminology. To obtain the optimum level of specificity of coding and abstracting, clinical coders should be instructed to review the entire medical/health record and not just use the discharge list, front sheet, discharge summary or attendance list;
- **developing a strategy to alert all health care administrators** of the need for staff development procedures for clinical coders;
- ensuring that the **accuracy of coding non-clinical data** is treated as importantly as coding clinical data;
- **reviewing and updating coding rules** continually to ensure that the rules are meaningful and enforceable. Specific rules and guidelines for coding should be current;
- **developing, implementing and maintaining coding standards and indicators**;

■ **employing sufficient staff to code and abstract data**. The number of staff required will vary from facility to facility and country to country. A formula developed in the United States to determine the number of coding staff required, as outlined by Abdelhak et al. (1996), is:

$$\frac{\text{Coding time per record x number of records to be coded for the period}}{\text{Number of hours worked per coder for the same period}}$$

For example, for inpatients, if it takes 15 minutes (0.25) to code a medical record, there are 900 discharges per month, and coders each work 150 hours per month, the calculation 0.25 x900/150 would indicate that 1.5 coders are needed;

■ **reviewing procedures for measuring clinical coding productivity**. Too often, productivity is simply defined as the ratio between input and output, which considers quantity and efficiency as the baseline factors in determining the level of productivity. But quality and effectiveness are equally important. Quality is concerned with the accuracy with which a task is completed while quantity is concerned with the number of task units produced per staff hour. Therefore, in measuring coding productivity, we should determine the number of coded medical/health records produced per employee per hour **that meet the established levels of quality**;

■ **implementation of regular and ongoing coding audits**. Many countries already conduct audits on a regular basis using the re-coding method described below. Ideally, for inpatients, all medical records should be coded within 48 hours of discharge with 98% accuracy. In primary health care situations, all health records should be coded by the end of the day after attendance. A coding audit should be conducted monthly to review performance and compliance should be noted and problems actively addressed to prevent recurrence.

■ **a coding audit could be conducted as follows**:
– an expert coder should be specially selected for the study and given additional training before beginning to re-code coded clinical data previously submitted;
– the expert coder should randomly select a specified number of records, re-code them and complete a new abstract, then compare the original abstracted coded data and the re-abstracted data to determine if a discrepancy in coding exists;
– if there is **no** discrepancy, that is, if the coding matched exactly, the medical record is put aside;
– if there is a discrepancy in the coding, a reconciliation process is undertaken. In this step, the expert coder and a representative from the medical/health record service should review the record to determine whether the original coding or the re-abstracted coding is correct;

– if it is agreed that there is an error, the error is noted and the reason the incorrect code was assigned is identified, thus indicating why the discrepancy in coding occurred;

– feedback of results to coders is essential, as is retraining where required. A follow-up audit should be planned.

Samples of studies in clinical coding quality

Re-abstracting/coding audits conducted in Australia and the United States of America over the years using similar methods generally found that non-medical entries such as admission and discharge dates, institutional transfers and patients' sex and age produced good results. Results for diagnoses and procedures, however, were less accurate (Jones, 1993).

Types of clinical coding problems generally found by such studies

■ rules documented by WHO in ICD books were not actively followed;

■ problems appeared to occur in determining the appropriate specificity at which to code due to the level of documentation present in the medical record;

■ difficulties also appeared in the assignment of diagnosis type and in particular in the choice of the correct main condition;

■ some hospitals had coding policies contrary to established rules and guidelines, forcing coders to make an inappropriate choice of codes;

■ information in some medical records was inconsistent, making accurate coding difficult;

■ medical records were not reviewed in detail enough to capture all relevant information about diagnoses and procedures; and

■ guidelines were not sufficient for selecting the appropriate procedure code.

A survey of inpatient facilities contributing to the New South Wales Health Departments Inpatient Statistics Collection was undertaken in June 1991.

Although primarily undertaken to assess training needs for coding staff, the survey identified significant problems in the underlying delivery of coding services. The major findings included:

■ only 52% of inpatient facilities employed coding staff to abstract patient data and assign codes;

■ 52% of respondents had a computerized patient information system;

■ 54% of coding staff spent less than 10 hours a week on coding duties; 32% spent less than four hours per week coding; only13% were employed on a full-time basis;

■ 45% of coding staff had had formal education in coding;

■ of the 55% who did not have formal education, only a relatively low number had completed basic coding courses; and

■ the most frequent problems experienced were inadequate medical record documentation and difficulties with recruitment of staff.

The survey revealed that medical record documentation needed improvement through the development and design of strategies to enhance medical record formats.

A second study in 1993 produced similar results. This aimed to examine the quality of clinical documentation in medical records and the effect deficiencies may have had on the allocation of DRGs (Jones, 1993).

The study involved a retrospective survey of medical records of patients discharged from hospital in 1990/91. Doctors at the hospital who agreed to assist with the survey reviewed a sample of cases from high throughput DRGs. The diagnostic and procedural information was re-abstracted from the medical records and compared with previous data. The use of medical personnel in the study enabled the researcher to measure the quality of documentation at the documentation level as well as the diagnostic coding level (Jones, 1993). In the review:

- 67% of diagnoses were not correctly documented or were missing from the front sheet of the medical record; and
- 52% of surgical procedures were missing or incorrectly documented in the medical record (Jones, 1993).

Jones explained that these results indicated that there were substantial errors in the accuracy and completeness of medical documentation. Although the survey focused on DRG assignment, it was recognized that such errors could have a major effect on patient care, as clinicians rely heavily on accurate and complete documentation within the medical record for the continuing care and treatment of a patient. In addition, incomplete and ambiguous data, as well as the impact on patient care, have the capacity to jeopardize the quality of information produced by hospitals for planning, resource allocation, research and epidemiological studies (Jones, 1993).

Other studies conducted in Australia and the United States of America referred to by Jones also indicated that there were serious problems with the reliability of medical documentation.

The Jones study demonstrated that there is a strong link between documentation on the front sheet and in the body of the medical record and the process of clinical coding of diagnoses and surgical procedures. The study also demonstrated that the documentation was not of a sufficient standard for clinical coders to achieve high accuracy and consistency in coding.

It must be emphasized that the responsibility for recording accurate and timely diagnoses and procedures rests with the health care professional treating the patient. If documentation in the medical/ health record is inaccurate, the by-products of that information, clinical coded data and health statistical data will also be inaccurate.

Summary

The dilemma of clinical coding quality rests with ambiguity and inconsistency in medical record documentation, incomplete medical records, disinterested and sometimes uncooperative health care

providers, and inconsistent interpretation of established guidelines for clinical documentation and clinical coding.

The approach to improving data quality must be coordinated and systematic. Provision should be made for the allocation of sufficient resources to adequately monitor coded data quality. The ability to quantify changes in the level of data quality over time would provide a useful indicator to not only validate the accuracy of coded data, but also assess the effectiveness of strategies taken to improve the coding process.

References

Abdelhak M, Grostick S, Hankin MA, Jacobs E. *Health Information: Management of a Strategic Resource.* Philadelphia, WB Saunders Company, 1996.

Amatayakul M. Improving Coding in Light of New Demands for Data. *Proceedings of the 11th Health Records Congress,* 105–109. Vancouver, Canadian Health Records Association, 1992.

Hyde LA. Developing a Quality Management Program for Coded Data. *Journal of AHIMA*, 63/1: 50–52, 1992.

Jones L. *The Quality of Clinical Documentation and Subsequent Effect on DRG Assignment: A Report on the Findings of the DRG Documentation Study:* Case Mix Development Program. Commonwealth Department of Health, Housing and Community Services, Canberra, 1993.

Liu Aimin and Ma Jia-run. The Accuracy and Speed of ICD-9 Coding. *Proceedings of the 11th Health Records Congress,* 119–124. Vancouver, Canadian Health Records Association, 1992.

National Centre for Classification in Health. *Australian Coding Standards, Vol. 5, International Statistical Classification of Diseases and related Health Problems, 10th Revision, Australian Modification (ICD-10-AM).* Sydney, NCCH, 1998.

National Centre for Classification in Health. *Performance Indicators for Coding Quality (PICQ).* Quality and Education Division, Sydney, NCCH, 2001.

World Health Organization. *International Statistical Classification of Diseases and Related Health Problems, 10th Revision, Vol. 2.* Geneva, WHO, 1993.

5 Data quality of statistical reports

*S*tatistical reports based on health care data, whether generated manually or by computer, must be accurate and timely.

As with all health care data, the collection of meaningful statistics is important and an essential function of all health care facilities. Health statistics are used to:

- assist health care facility administrators and national, state or province health care authorities to make informed decisions in promoting the development of health care services; and
- actively co-ordinate statistical activities across government agencies and in discussions relating to health services with government authorities and the community.

> *The production of accurate, relevant and timely statistical information about the health status of the community is essential to identify risk factors associated with health and the use of, and need for, health services.*

Statistical reports based on health care data, whether generated manually or by computer, must be accurate and timely to meet the requirements of the health care provider, health care facility, local, state or province health authorities, and national governments in the production of quality health care information.

As for all health information, medical/health records are a primary source of this data. Whoever is responsible, however, whether medical record staff or other administrative staff, should be involved in the design and development of the protocol for collecting and processing the data.

If data quality is not maintained at the site of the original data collection, statistical information will not meet the requirements of the facility or government.

Health care statistics

Health care statistics are numerical data used to describe the population of a health care facility and also the services on offer. They provide information on:

- type of diseases treated by the facility;
- type of diseases occurring within a community and/or country;
- number of births and sex of newborns;
- number of deaths and age of death, including maternal, perinatal and infant deaths; and
- utilization of health care services within a community.

As statistics are only as accurate as the original documents from which they are produced, the person responsible for collecting and processing the data must ensure that the original source documents are accurate, complete and readily available.

Health care statistics are used for:

- comparison of present and past performance of health care facilities and appropriateness of the services offered;
- comparison of health status at the national and international level;
- planning future developments at health care facility and community and/or national level;
- appraising the work performed by all health care providers;
- evaluation of funding requirements of health care facilities and community services; and
- current and retrospective clinical and health services research.

All statistical reports should be monitored and evaluated to determine if they are accurate and complete. The first step would be to conform to uniform standards.

Standardization of statistical terms and formulae

Standards should be developed and used as guidelines for the collection of statistical data, as facilities and governments need to know that the information being collected and disseminated in a facility or in a community means the same when compared to other facilities or communities.

- Standard terms and definitions should be used if health authorities wish to be able to compare the health status of their population with other countries. It is important that they do not vary between health care facilities and services within the same country.

55

- Many countries already produce their own **Glossary of Healthcare Terms and Definitions** containing standardized terms, which they use as a basis for the definition of statistical terms and formulae.
- In the absence of a glossary and set of standard terms and formulae, a country should consider using the standard definitions and regulations regarding statistics for international comparison published in Vol. 2 of ICD-10 by WHO (1993). Alternatively, the WHO guidelines could be used as a starting point and a simple set of standard terms developed to meet the country's own requirements. Particular definitions such as *foetal death*, *maternal death* and *live birth* published by WHO should be used along with ICD-10 to enable meaningful statistics to be compared globally.

Determining the appropriateness of statistical reports

It is important that statistical reports are prepared in a timely and accurate manner.

The reports generated are important and are used as a tool of communication. Some reports only present the work accomplished during the reporting period and may not be particularly useful for problem identification or for decision making. A report that compares selected data and indicators over different time periods may prove more useful. For example, the data and indicators for a month and the year to date can be compared with data for the same month and year to date of the previous year.

As a general rule, all statistical reports should be carefully prepared and analysed to determine if totals and formulae are accurate, appropriate, consistent and reliable.

Questions that should be asked when determining the appropriateness of statistical reports include:

- Are the reports producing the type of information required by the health care facility and government authorities?
- Are the reports being used?
- Are the reports accurate, reliable, meaningful and timely?
- Are data displayed in an appropriate manner enabling them to be reviewed?
- How much of the report is not used?
- Could the reports be improved?

The comparison of selected data of two or more reports may reveal inconsistencies.

For example, a major difference in the death rate in January and February could identify an error in calculation or an error in the original data from the daily bed census.

In addition, a *face validity* check should be used to identify any possible quality problems. This procedure is the valuation of data to see if the information is logical and appears to be accurate. For example, at hospital level:

- the bed occupancy rate is shown as 110% when ideally it should be between 85% and 90%. Have incorrect figures been used to calculate the rate? Or has the facility added beds and the correct figure for the number of beds not been used?

- the maternal death rate is shown as 20%, could the decimal point have been omitted so that the figure should be 2%? Has there been an error in the calculation? Or has the wrong formula been used?

To ensure that the information is useful when preparing a statistical report, the purpose, objectives and scope of the statistical analysis must be defined. The person preparing the report needs to ask questions such as:

- What information does the user want?
- What information is available?
- Are the data routinely collected by the facility or will it require the collection of additional data?

Data presentation

How data are presented is important in ensuring the quality of statistics and statistical reports. Although most reports will be in tabular form, they would be easier to read if visual aids such as tables or graphs were used to illustrate clearly what the figures indicate.

> *The way data are displayed is important. Reports should be clear and concise, and leave no doubt as to what the figures represent. All presentations should be simple and readable with important facts highlighted.*

Statistical tables are used to summarize data and graphs present data in visual form. The decision as to how to present the data, that is, in the form of a table or graph, will depend on the purpose of the report and who will be using the information.

Tables

A table is the simplest means of summarizing a set of observations, and can be used for all types of data (Youmans, 2000). It is an organized arrangement of data, usually appearing in columns (reading down) and rows (reading across). The columns should be labelled, for example:

An example of table			
Name	**Age**	**Clinical Service**	**Length of Stay**
B. J. Singh	64	Cardiac	4.5 days

The essential components of a table should include (Youmans, 2000):

- ■ **title** – this must explain in simple words what is contained in the table;
- ■ **stub-headings** – the title or heading of the first column;
- ■ **column headings** – the headings or titles for the columns;
- ■ **stubs** – the categories (the left-hand column of the table);
- ■ **cells** – the information formed by intersecting columns and rows;
- ■ **source of footnote** – the source for any data should be identified in a footnote.

Advantages of using tables include (Youmans, 2000):

- ■ more information can be presented;
- ■ exact values can be read to retain precision;
- ■ supportive details can be provided;
- ■ less work and fewer costs are required in the preparation;
- ■ they are easy and inexpensive; and
- ■ flexibility is maintained without the distortion of data.

It is important to use good judgement when preparing a table. The table should be checked to be sure that it is logical and self-explanatory. One needs to ask questions such as:

- ■ Are headings specific and understandable for every column and row?
- ■ Do the totals add up?
- ■ Is it easy to read?

Graphs

A graph should be used as a pictorial representation of numerical data and in most cases proves the best medium for presenting data for quick review of relationships between various factors. They should be designed so that they convey the general patterns in a set of data at a single glance (Youmans, 2000). Graphs also supply a lesser degree of detail than tables but can be useful in displaying statistics.

Graphs should be easy to read, simple in content and correctly labelled. The presentation of data in the form of a graph is an excellent way to get the message across (Youmans, 2000). There are many computer software programs available to automatically convert data into graph form.

Whatever the type of graph used, it should (Johns, 2002):

- ■ display the data in such a way as to allow the user to think about the meaning of the data;
- ■ avoid distortion of the data;
- ■ encourage the user to make comparison with previous data; and
- ■ reveal data at several levels from a broad category to fine details.

The essential components of a graph include (Youmans, 2000):

- ■ **title** – the title must relate what the graph shows as simply as possible;
- ■ **legend or key** – when several variables are included on the same graph, i.e. males and females, it is necessary to identify each by using a legend or key;
- ■ **scale captions** – these are placed on both axes to identify the values clearly; and
- ■ **source footnote** – the exact reference to an outside source should be given.

Presentation can be made more effective by the selection of the most appropriate graph to accompany the data. An overabundance of graphs, however, should be avoided.

Advantages of using graphs include (Youmans, 2000):

- they are attention getters;
- they are more easily understood than tables;
- they bring out hidden facts;
- they display trends or comparisons more vividly; and
- a picture is worth a thousand words.

Whatever format is selected to present statistical data, it must be simple, easy to read and carefully prepared. The quality of the original data must be of a standard to support the presentation and enable those viewing the report to be confident in the details presented.

Computer generated statistics

Health care statistics and a great deal of health care documentation in many countries are produced electronically via sophisticated computerized hospital information systems.

This often results in the collection of more sophisticated data that is readily retrievable. This tendency, to produce more information than required, often makes the report more complicated and less meaningful.

All statistical collections need to be monitored to ensure that the information being produced is useful and relevant. The administrators must find out what is needed, and when, why and how it is to be used. There is a misconception that because a computer produced the reports, they must be correct. All computer systems must have in-built edits and validation checks to help ensure the production of accurate information.

A computer can calculate statistics only from the data entered. As the final report is dependent on the original source data, whether produced by a computer or manually, if data are NOT correct, the information generated will be incorrect.

Data quality of vital statistics

A special branch of statistical data collection is vital statistics. Vital statistics in health care include births and deaths.

Births

Data on births are reliant on the completion of birth certificates collected and signed by the health professional delivering the baby, usually a doctor or midwife. The data need to be accurate and complete and must be processed quickly to enable the appropriate statistics to be compiled.

Birth certificates are not always issued from a health care facility. In many countries, babies are born at home, or in special "birthing centres" and the person delivering the baby is responsible for seeing that the birth certificate is completed accurately and forwarded to the appropriate government office. He or she is also responsible for recording the information on a data collection form for processing either by the facility within which the birth occurred or the local government area office if the baby is born at home or in a remote centre. This information must at some time reach the government official responsible for producing statistics about births.

Deaths

The primary source of mortality statistics is the death certificate, which is also forwarded to the appropriate Registrar of Births, Deaths, and Marriages. In the case of death, again not all people die in a hospital. Many die suddenly at home or elect to die at home if terminally ill. In addition, if the patient is in a remote area where access to a hospital is not available, the local doctor usually looks after him or her. If the patient dies, the local doctor is responsible for the completion of the death certificate. All death certificates must be forwarded to the appropriate local government office, in many countries via the funeral director.

As with birth certificates, the accuracy of the death certificate is important. There are certain requirements for completing death certificates, which must be followed, to enable the production of accurate and reliable information on the major causes of death in a community and country. When only one cause of death is reported, this should be the underlying cause of death, which is used for tabulation (WHO, 1993). When more than one cause of death is recorded, the *underlying cause of death* must be used for tabulation. A government officer receives the information from the various sources and uses ICD (whatever edition is being used in the country) to code the underlying cause of death. The same data quality principles discussed in clinical coding apply to the coding of death certificates.

An overview of causes of death is released annually in most countries, giving detailed information on deaths, and causes of death, which is comparable across all states, provinces and countries.

Causes and sources of poor quality statistical data

There are numerous causes and sources of poor quality statistical data, including:

- **incomplete and inaccurate original source data** – in hospitals for example, this can arise from inaccurate ward census reports, incomplete medical records and inaccurate clinical codes. Inaccurate or delayed transcription of data into either manual or computer systems may also be the cause of poor quality statistical data;
- **insufficient data** documented in health records – at primary health care level, problems arise due to insufficient data recorded by the attending health professional. This in turn

leads to inaccurate coding and problems in the collection of accurate and meaningful statistics;

■ **poorly designed data collection forms** and poor database design in the computer;

■ **inaccurate transfer of data onto statistical collection forms** or errors in the transfer of data into a computer at the point of service;

■ **poor data transfer from one document to another** – for example, errors in transferring clinical codes from the front sheet of the medical/health record to the data collection/abstract form;

■ **lack of approved procedures relating to data collection** – poorly defined procedures for collecting and processing statistical data are a major problem in obtaining quality data;

■ **unreliable storage media and backup systems** in computerized collections;

■ **inaccurate coded clinical data,** as discussed in Chapter 4;

■ **lack of training of data collection staff**, particularly in relation to requirements for accurate and timely statistical data collection;

■ **problems with data entry devices**, such as equipment failure and lack of sufficient devices at appropriate treatment areas within the facility;

■ **lack of staff training** in the proper use of data entry devices both for manual systems and computer-based systems;

■ **paper records not available for data collection** of statistical data when required; and

■ **lack of standard terms and formulae** or incorrect application of standard terms and formulae.

Methods to improve data quality of statistical reports

Methods to improve data quality start with the improvement of data collection, data transfer, clinical coding and subsequently the production of statistical reports. Measures in place to ensure data quality in all documentation should include:

■ **monitoring original data**, such as ward census reports, discharge lists, medical/health records, lists of outpatient and clinic attendances, to ensure they are correct and received within an established time period;

■ **follow-up procedures** – a procedure should be in place to enable staff to immediately follow-up on missing reports and records;

■ **checking ward census reports** – these reports should be checked daily to ensure that:
  – every patient listed as a transfer in or transfer out of a ward appears as a transfer on the census report of the appropriate ward receiving or transferring the patient;
  – the number of patients remaining at the end of the day agrees with the number obtained by adding to the patients remaining the previous day to the number of entries to the

ward (admissions and transfers in) and from this total subtracting the number of discharges, deaths and transfers out from the ward;

– all indicators should be checked for accuracy and to be sure that they make sense. For example, the number of patients remaining at the end of the time period should be lower than the number of beds available for use, unless of course more than one patient occupies a bed on the same day.

Other methods to improve the quality of statistical reports

■ Ensure that data collected are meaningful to the persons collecting them by conducting training programs.

■ Ensure that all processes relating to the accuracy, validity, reliability, completeness, legibility, timeliness, usefulness and accessibility of data are in place.

■ Develop appropriate procedures for collecting statistical data such as:
  – verification of data before entry;
  – determining that the terms and definitions used comply with standards previously set;
  – prepare reports in a logical, useful and meaningful manner. That is, present figures in the most appropriate format to enhance the usefulness of the information; and
  – check reports for face validity and consistency.

■ In computer systems, ensure data quality assurance such as edit and validity checks.

■ Proofread all reports before distribution to be sure that they are error-free.

■ Design a plan for conducting quality assurance audits. This involves routine checks of data submitted to the processing centres and management. Steps would be to:
  – assign responsibility for the quality review programme;
  – formulate a program to check for completeness of coverage, accuracy and consistency of data against pre-determined standards;
  – determine what is to be checked;
  – conduct the audit on all statistical reports;
  – review findings and follow-up problems identified and provide appropriate feedback to all staff involved. This is essential at all times but particularly for staff in remote areas working in isolation;
  – retrain staff if required; and
  – repeat the process at a specified time period.

■ The person responsible for the presentation of statistical reports should examine the final product carefully before distribution to verify that the information presented is correct. Questions they should ask include:

–   do the totals match?
–   are the figures presented in the best way?
–   are they easy to understand?

These questions should be addressed and steps taken to correct the report if discrepancies are noticed. For example, when examining the discharge analysis report, you notice that the death rate is 01%, which is an odd percentage. This should alert you to the fact that it may not be correct. You would need to recalculate the rate and carry it to two decimal points. For example, the corrected figure could be 1.38%. By showing it as 1.38% or 1.4% may be more meaningful to the administration or government authorities.

Summary

The ultimate responsibility for data quality in the collection, presentation and distribution of statistical data rests with the administrator of the health care facility and the person in charge of information services at national level.

All persons collecting and processing data must accept responsibility for the accuracy and reliability of the data they collect and process.

The standardization of terminology, data elements and formulae used are important and must be adhered to if the health care facility wants to compare data from previous years, with other facilities and with other countries.

References

Johns M. *Health Information Management Technology*: Chicago, American Health Information Management Association, 2002.

World Health Organization. International Statistical Classification of Diseases and Related Health Problems, 10th Revision. Vol. 1, Rules and Guidelines for Mortality and Morbidity Coding. Geneva, WHO, 30–123, 1993.

World Health Organization. International Statistical Classification of Diseases and Related Health Problems, 10th Revision. Vol. 2. Statistical Presentation. Geneva, WHO, 124–138, 1993.

Youmans K. Basic Statistics for Health Information Professionals. Revised Edition. Chicago, AHIMA, 2000.

6 Data quality in public health records and reports

If information generated in public health reports is inaccurate or not available when needed, the consequences could be devastating to the community.

In many countries, public health provides both direct clinical services and community-based services, which require good data collection and recording systems to be able to provide accurate and understandable information for government policy-makers, community leaders, health planners and health care providers (Abdelhak et al., 1996).

As for all health information, public health records and reports must be accurate and valid, reliable and effective.

Because public health authorities are responsible for determining disease patterns, preventing disease and injuries, and educating consumers to take care of themselves, it is essential that the health information used is accurate and reliable.

Public health prevention measures are used to protect whole communities or populations from communicable diseases, epidemics and environmental contaminants (Abdelhak et al., 1996). If information generated in public health reports is inaccurate or not available when needed, the consequences could be devastating to the community. Clean water is essential to ensure the well being of whole communities and detailed, accurate and timely information on chemical poisoning from fertilizers or industrial waste, for instance, is crucial.

Data collection and dissemination

Governments are reliant on the accuracy of information contained in records and reports:

- for the production of high-quality, user-orientated and dynamic statistical information on the health status of the community;
- on risk factors associated with disease and the use of, and need for, health services and for informing the community about these matters;
- on providing statistics to the government to assist in formulating, monitoring and evaluating health politics and progress made in the provision of health care services; and
- when conducting national health surveys to obtain information on a range of health related issues.

Surveys comprise a core data set, which will be repeated in successive surveys to monitor trends in the health of a community over a period of time, and a supplementary data set, which can be varied from survey to survey to address key issues of the day (Abdelhak et al., 1996).

Surveys are designed to collect information within a community:

- about recent and long-term illness or injury;
- about the use of health services;
- on medications ordered and used in response to illness or injury;
- on smoking patterns;
- on alcohol consumption;
- on physical exercise;
- on dietary changes; and
- on women's health issues.

Data recorded and information produced need to be accurate and reliable if they are to be useful.

Public health services

Departments of public health play an important role in the collection and dissemination of national health statistics, which must be accurate, clear, concise and understandable. As a government department they are responsible for:

- the collection of a vast range of historical and current data used for continuity of present and future health care services;
- maintaining extensive links with governments and international agencies; and
- developing and maintaining statistical standards and classifications to undertake complex large-scale statistical activities and provide a range of services.

The core activities of public health include:

- collection, analysis and dissemination of quality statistical information on the health status of the population, the availability of health services, and community-based immunization and health screening programmes;

- development of public health policy at state, province and national level;
- development of quality assurance programmes to monitor the collection of accurate and appropriate data and policy decisions;
- compilation and publication of statistics on the health of particular population groups, such as children, including infants and babies, women, indigenous people and ethnic groups; and
- analysis of trends in mortality within the community and country.

Causes and sources of poor data quality

- Reporting requirements are often poorly defined and dysfunctional.
- Data collection is often incomplete and inaccurate.
- Statistics come from a wide range of sources without the use of standardized collection methods, making consistency difficult.
- Lack of standards for reporting statistical data.

Methods to improve data quality in public health data collection

Abdelhak et al. (1996) highlights some ways public health authorities can implement an effective integrated information system. They include:

- redesigning programme specific data systems into integrated systems;
- developing an integrated, centrally managed electronic network that provides access to national, state and local information systems;
- using a data system that helps to provide services to the public;
- implementing a data management system that meets local needs in systematic collection, analysis and monitoring of standardized baseline data;
- introducing a link between local and state-wide databases in the private and public sectors;
- determining data use and dissemination standards;
- employing technical assistance to ensure a high standard of data analysis, dissemination and communication;
- developing and maintaining a fully integrated and secure computer network;
- evaluating and disseminating information on new health information technologies; and
- using a system of tracking clinical and environmental laboratory information.

Summary

The population focus of public health helps to accomplish changes and improve the health status of the community. Accurate and timely information is essential when governments are working towards improving sanitation, ensuring clean water, maintaining immunization programmes and eradicating infectious diseases such as malaria, diptheria and tuberculosis.

Public health authorities also need to have current and accurate data on diseases such as HIV/AIDS, sexually transmitted diseases, diabetes, heart disease, all forms of cancer and many more. In addition, they need to monitor child abuse and motor vehicle accidents.

As for all health care data, public health authorities need data collection, analysis and dissemination systems to provide accurate, valid, reliable, consistent and complete information to evaluate existing resources, plan for the future and introduce measures to anticipate problems before they arise.

References

Abdelhak M, Grostick S, Hankin MA, Jacobs E. *Health Information: Management of a Strategic Resource.* Philadelphia, WB Saunders Company, 1996.

Glossary

Aid post	A small clinic or room in an outlying area where health care and health screening are offered.
Allied health professional	Physiotherapy (physical therapy), occupational therapy, speech therapy, social worker, etc.
Audit trail	A programme that records the access and/or action that occurs in a computer record by logging the user identification and the file identification, recording date and time of access and action carried out.
Case mix	The type of cases treated within a health care facility based on characteristics such as resource consumption, diagnosis or procedure, or reason for visit.
Clinical coding	The assignment of a specific code to a narrative statement of diagnoses and procedures.
Clinical data	Data about a patient's care and treatment, whether admitted to a health care facility or treated in a primary health care environment.
Clinical staff	Doctors, nurses, health extension officers, nurse practitioners, midwives and allied health professions.
Coding	A procedure that assigns a numeric code to diagnostic and procedural data based on a clinical classification system.
Coding quality indicator	A measure of a particular aspect of coding practice.
Daily admission list	A daily list of all patients admitted to the hospital.
Daily discharge list	A list of inpatients discharged on a specific day.
Data dictionary	A set of core uniform definitions and data items relating to the full range of health services and a range of population parameters.

Data element	A unit of data, usually defined by a set of attributes specifying its definition, its identification, the way it is represented and its permitted values.
Day only	Day only patients are admitted in the morning and discharged in the afternoon. Day only patients are NOT day only patients if they stay in hospital overnight.
Diagnosis related group	A patient classification scheme that provides a means of relating the type of inpatients a hospital treats to the costs incurred by the hospital.
Discharge summary	A summary of a patient's stay in hospital written by the attending doctor.
Disease index	Lists diseases, conditions and injuries by the specific code number for each disease, condition or injury, based on a clinical classification system to allow for retrieval of medical records for research by each specific code.
DRG	See Diagnosis Related Group
Emergency patient	Attends a hospital or health care facility needing immediate attention for a disease or injury.
Face validity	Validation of data to see if the information is logical and appears accurate.
Front sheet	The first form in the medical record. Also called identification and summary sheet
General outpatient	In developing countries, a general outpatient is a patient attending the outpatients department of the health care facility without an appointment. These patients do not include accident and emergency patients.
Health care data	Items of knowledge about an individual patient or group of patients.
Health care facility	Hospital, community health centre, outlying clinic, or aid post, etc.
Health care standard	Best practice principles and guidelines for collection and storage of health care data.
Health care statistics	Numerical figures used to describe the type of services and number and type of patients treated by a health care facility over a specific period.
Health information	Health care data that have been organized into a meaningful format.
Health record	A single record of all data on an individual's health status, including birth records, immunizations, reports of all physical examinations as well as all illnesses and treatments given in any health care setting. Often used interchangeably with "medical record," but is a broader concept.

HIM	Health information manager – the person who manages the health information service.
HIS	Hospital information system – a collection of data relating to patients and their care.
Hospital number	See Medical Record Number.
ICD-9	International Statistical Classification of Diseases (9th revision) published by WHO.
ICD-10	International Statistical Classification of Diseases and Related Health Problems (10th revision) published by WHO.
ICPM	International Classification of Procedures in Medicine, published by WHO.
Identification number	See Medical Record Number.
Inpatient	A patient who has been admitted to the health care facility. Inpatients usually occupy a bed in a health care facility.
Main condition	The condition, diagnosed at the end of the episode of health care, primarily responsible for the patient's need for treatment or investigation. If there is more than one such condition, the one held most responsible for the greatest use of resources should be selected. If no diagnosis was made, the main symptom, abnormal finding or problem should be selected as the main condition (WHO, 1993).
Master patient index	Contains identification information of all patients admitted to a health care facility and is the key to locating a patient's medical record.
Medical record	A collection of facts about a patient's health history, including past and present illness(es) and treatment(s) written up by the health care professional treating the patient.
Medical record audit	A retrospective review of selected medical/health records or data documents used to evaluate the quality of care or services provided compared with pre-determined standards.
Medical record number	The number used to identify the patient's medical record and file the medical record. Also referred to as hospital number, identification number or unit record number.
Medical record room	Usually a small medical record department in a developing country.
MPI	Master patient index.
Operation index	Lists operations and procedures by a specific code number based on an operation or procedural classification system. The index enables the retrieval of medical records of all patients who have undergone a specific operation or procedure while in hospital.

Outpatient	A patient who attends an outpatient department, is not admitted to a health care facility, and does not occupy a bed for any length of time.
Patient held health record	A record kept by the patient, or parent if a child, which covers the life of a patient from birth to death. All health professionals caring for the patient record their findings and treatment in the record. Also referred to as a longitudinal record.
Patients' master index	See Master Patient Index.
Patients' name index	See Master Patient Index.
Primary data	Data obtained from the original data source, e.g. documentation in a patient's medical/health record and data collected in the daily bed census.
Principal diagnosis	The condition established after study to be chiefly responsible for occasioning the admission of the patient to hospital for care (USA definition). The diagnosis established after study to be chiefly responsible for occasioning the patient's episode of care in hospital or attendance at the health care facility (Australian Definition).
Procedure index	See Operation Index.
Secondary data	Data sets derived from primary data.
Specialist outpatient	An outpatient who attends a specialist clinician in the outpatients department. A specialist outpatient is usually a patient with a chronic problem (hypertension, diabetes, etc.), a paediatric patient or a recent inpatient.
Standard	A specimen or specification by which something may be tested or measured.
Tracer	A card, usually the same size or slightly larger than the medical record, which replaces the medical record in the file when the record is removed for use elsewhere in the hospital.
Underlying cause of death	The disease or injury that initiated the train of morbid events leading directly to death or the circumstances of the accident or violence that produced the fatal injury.
Unit record number	See Medical Record Number
Unique patient characteristic	Something about a patient that does not change, such as his or her mother's maiden name or a national identification number, or social security number.

Appendices

Medical Record Review Form

Medical Record Number: _____   Attending Doctor: _____
Admission Date: _____   Discharge Date: _____
Reviewer: _____   Date: _____

CRITERIA	Y	N	N/A
Informed consent for treatment signed by patient	☐	☐	☐
Past medical history and history of present illness	☐	☐	☐
Documented within 24 hours of admission and prior to surgery:	☐	☐	☐
Past history, etc. contains:			
Chief complaint	☐	☐	☐
Medical history	☐	☐	☐
Family history	☐	☐	☐
Social status	☐	☐	☐
Review of symptoms	☐	☐	☐
Physical examination	☐	☐	☐
Treatment plan	☐	☐	☐
Initial nursing plan documented within 24 hours of admission	☐	☐	☐
Progress notes documented daily, signed and dated	☐	☐	☐
Pre-anaesthesia assessment documented, signed and dated	☐	☐	☐
Surgery/procedure performed documented, signed and dated	☐	☐	☐
Post-operative monitoring of patient in Recovery room	☐	☐	☐
Operation report documented immediately following surgery:	☐	☐	☐
Operation report includes:			
Procedure	☐	☐	☐
Findings	☐	☐	☐
Specimen(s) removed	☐	☐	☐
Postoperative diagnosis	☐	☐	☐
Discharge note recorded in progress notes on day of discharge:	☐	☐	☐
Discharge note includes:			
Condition on discharge	☐	☐	☐
Final diagnosis	☐	☐	☐
Prognosis	☐	☐	☐
Follow-up details	☐	☐	☐
Discharge summary completed by attending doctor:	☐	☐	☐
Discharge summary includes:			
Summary of hospitalization	☐	☐	☐
Treatment and medications	☐	☐	☐
Final diagnosis (principal)	☐	☐	☐
Associated diagnoses	☐	☐	☐
Procedures	☐	☐	☐
Prognosis and follow-up	☐	☐	☐

APPENDIX 2

Simple Data Collection Form for a Manual Medical/Health Record Tracking System
Accessibility of Records Criteria

CRITERIA	YES	NO	COMMENT
1. Does it take staff more than 5 minutes to locate a medical record in the permanent file?	☐	☐	_____
2. Does it take staff more than 5 minutes to locate an incomplete medical record?	☐	☐	_____
3. If a medical record is not on file, is a tracer or outguide in its place?	☐	☐	_____
4. Are names and numbers correctly and legibly written on folders?	☐	☐	_____
5. Are stacks of medical records found lying on desks and in no order?	☐	☐	_____
6. Does the medical record department/office have a procedure for medical record location?	☐	☐	_____

Data Collection Form: Content of Medical Records

CRITERIA	Y	N	COMMENT
Does each medical record on file contain:			
1. A discharge summary is included?	☐	☐	_____
With significant findings?	☐	☐	_____
Procedures performed?	☐	☐	_____
Treatment given?	☐	☐	_____
Patient's condition on discharge?	☐	☐	_____
Final diagnosis — main condition and other conditions?	☐	☐	_____
Instructions for follow-up?	☐	☐	_____
2. Patient history and physical examination present?	☐	☐	_____
Is the chief complaint recorded?	☐	☐	_____
Details of present illness?	☐	☐	_____
Relevant past and family history?	☐	☐	_____
Body system review?	☐	☐	_____
Report on physical examination?	☐	☐	_____
Provisional diagnosis?	☐	☐	_____
3. An operation report (if a surgical patient) present?	☐	☐	_____
Includes a preoperative diagnosis?	☐	☐	_____
Includes a postoperative diagnosis?	☐	☐	_____
Specify findings?	☐	☐	_____
Contain the surgeon's signature?	☐	☐	_____
Contains the date of procedure?	☐	☐	_____
4. Progress notes from admission to discharge?	☐	☐	_____
Written at least daily?	☐	☐	_____
Signed and dated?	☐	☐	_____
Reflective of the patient's admission status?	☐	☐	_____
A chronological report of the patient's progress during hospitalization?	☐	☐	_____
Contains a final discharge note?	☐	☐	_____
5. Pathology, laboratory and X-ray reports if ordered included?	☐	☐	_____
6. Do all entries in the medical record contain signatures?	☐	☐	_____
Are all entries dated, including day, month and year?	☐	☐	_____
Are times of treatment noted?	☐	☐	_____
Contains a signed consent for treatment?	☐	☐	_____
7. Patient identification recorded on all pages?	☐	☐	_____
Do all pages contain correct patient identification?	☐	☐	_____
Are all entries legible?	☐	☐	_____
8. All nursing notes are signed and dated?	☐	☐	_____