Chapter 7

Data Quality Management

Table Contents

Appendix

7.1 Introduction

The credibility of a birth defects surveillance program is built on a foundation of high-quality data. Information and results that are derived from surveillance data should be accurate, complete, and timely. Data quality influences the results of descriptive epidemiologic studies and, therefore, their interpretation. Data quality also affects the extent to which information can be utilized for planning, prevention, and intervention.

In this chapter, we will discuss some of the issues that affect the quality of data in surveillance systems and suggest methods for quality improvement. In Section 7.2 we present criteria designed to produce high-quality data. In Section 7.3 we introduce some relevant terminology. In Section 7.4 we discuss the relationship between data sources and quality, and in Section 7.5 we outline the distinctions between timeliness on the one hand versus thoroughness and completeness on the other. Sections 7.6 and 7.7 present various aspects of quality control and quality assurance, stressing the differences between the two. The importance of computer technology in support of quality improvement is particularly highlighted in Section 7.7. Nine specific quality improvements methods are discussed in detail in Section 7.8. References cited in this chapter may be found in Section 7.9.

This chapter contains a Data Sources Descriptive Assessment Tool that may help surveillance staff systematically evaluate the various data sources available to them (Appendix 7.1).

7.2 Criteria for High-Quality Data

High-quality data have a positive cascading affect on a surveillance program's outcome measurements – such as accuracy, completeness, and timeliness – which, in turn, can be monitored as a means to improve program performance. The term ***quality*** has many definitions and interpretations depending on use and intent. Philip Crosby, a total quality management expert, defines quality as "the conformance to agreed and fully understood requirements" (Dale and Bunney, 1999). In the surveillance field, this translates into the identification of a target, bench mark, or goal that defines the requirements against which results are measured.

Some experts in surveillance have suggested that the most important measurement indicators (or criteria) related to high-quality data are described by the mnemonic TACOMA (NAACCR, 2000). Data must be **T**imely, **A**ccurate, **C**omplete, **O**riented, **M**easurable, and **A**pplicable. The relative importance of these factors should be weighed and balanced, individually and in total, against the program's objectives and resources.

In the next section we define the terms on which the TACOMA mnemonic is based – timeliness, accuracy, completeness, oriented, measurability, and applicability – as well as several additional terms important for an understanding of data quality issues.

7.3 Terminology

Timeliness	The extent to which data are rapid, prompt, and responsive. For example, a birth defect case should be ascertained or reported to the program shortly after diagnosis. With rapid case identification, the program is able to provide timely prevention and intervention services, respond quickly to investigations, and monitor trends.
Accuracy	The extent to which data are exact, correct, and valid. For example, accurate diagnostic data affect a program's ability to provide reliable disease rates and to maintain data comparable to those from other programs. Diagnostic accuracy reflects the program's standard to conform to agreed-upon case definitions and requirements.
Completeness	The extent to which data are all-inclusive and comprehensive. For example, are all of the cases of birth defects that occur within the target population, within a specified time period, identified by the surveillance system?
Oriented	The extent to which data are focused, targeted, and intended. For example, programs should collect only those data that are appropriate to their goals and objectives. Programs should determine which data variables should be collected, how quickly they can be collected, and the resources available to be devoted to their collection. Having an oriented perspective parallels the 'minimum necessary' privacy standard of the Health Insurance Portability and Accountability Act or HIPAA (i.e., identify and use only what is necessary). (See Chapter 2 on Legislation for additional information on HIPAA and Chapter 4 on Data Variables recommended for consideration by birth defects surveillance programs.)
Measurability	The extent to which data are quantifiable, calculable, and objective. For example, the conformance to agreed-upon data definitions provides the foundation for quantitative evaluations.
Applicability	The extent to which information is relevant. Outcome measurements should be designed to promote data utilization. Information derived from the data should be beneficial to the target population or to public health interests.
Comparability	The extent to which the data in one data set conform with those in other data sets. For example, programs that agree to adhere to standard data definitions and case definitions produce data that can be evaluated and weighed against one another.
Thoroughness	The extent to which data collection activities are meticulous and exhaustive in completing a case abstract or case record. In other words, each data field on case abstracts and case records should be filled in.
Outcome measurements	Strategically planned results that may be quantitative or qualitative. Criteria, such as those described by TACOMA, or other defined factors, are specifically selected (and developed) to evaluate, track, and monitor a program target, goal, or benchmark. Desired outcome measurements are often developed in the planning stages of a surveillance program, for performance evaluations, and when adding new projects. Staff should identify the type or category of results to be measured in order to evaluate progress in achieving program goals and objectives (or study objectives, project targets, etc.).

7.4 Data Sources and Quality

Depending on the case ascertainment approach, birth defects cases are found-at or reported-from data sources. Therefore, the importance of the role data sources play in case ascertainment and surveillance should not be underestimated.

Quality issues surface because of variations among data sources. Some data sources may provide diagnostic information, but may lack important demographic information. Some may be service-focused, such that a precise diagnosis may not be important. Others may provide in-depth information on a specialty area, but may not identify other conditions that co-occur. Still others are administrative databases.

A single data source has the potential to affect multiple outcome measurements. For these reasons, programs should evaluate each data source in order to describe its basic characteristics, as well as to identify its potential strengths and weaknesses. A descriptive assessment tool should be designed to answer specific questions about each data source in relation to surveillance requirements. An example of such a tool is provided in Appendix 7.1 (Data Sources Descriptive Assessment Tool).

Quantitative evaluations should include outcome measurements for accuracy, completeness, and timeliness. Often data sources are evaluated in combination with other quality assessments. For example, diagnostic accuracy may be evaluated by staff reviewing a medical record to confirm a diagnosis that was identified-at or reported-from a data source. In this example, the data source is part of the evaluation because it is where the diagnosis case report originates; however, other aspects of the case ascertainment process may be evaluated as well.

Examples of quantitative evaluations are provided in Section 7.8 of this chapter on Quality Improvement Methods.

The program should:

- **Use** the data quality criteria in TACOMA as a guide when identifying outcome measurements and when evaluating data sources.

- **Identify** other factors that are important to consider, including those that relate to staff resources, such as ‘location of site’ and ‘volume of case reports’ (in relation to distance traveled).

7.5 Timeliness Versus Thoroughness and Completeness

Surveillance systems generally have limited resources to use in meeting program objectives. Additionally, staff face dilemmas in terms of prioritizing resources to achieve the outcome measurements of timeliness, thoroughness, and completeness. Should a program set a goal of timely data at the risk of potentially missing cases? Or risk losing timeliness by setting a goal of the most complete surveillance database? It is important for programs to achieve a balance that suits their needs, while also being responsive to external requirements, such as guidelines for submitting data to the National Birth Defects Prevention Network (NBDPN), as discussed in Chapter 10 on Data Collaboration and Dissemination.

Timeliness improves a system's ability to be responsive for investigations, up-to-date for monitoring trends, and current for referral to services. Thoroughness is a measure of finished versus unfinished case abstracts and case records. Clearly, data fields that are empty or inconclusive are not useful for most outcome measurements. Completeness is important because descriptive epidemiology – including the calculation of birth defects rates – is more comparable, accurate, and reliable when a surveillance program is confident that all cases have been ascertained.

When prioritizing resources to balance the quality indicators of timeliness, thoroughness, and completeness an important outcome measurement recommended by NBDPN is that the surveillance database be 95% complete by two years past the date of birth or fetal demise. Some programs may have a longer time period for reporting birth defects and, therefore, have a longer time period for case ascertainment. Still, it is important that surveillance systems be sufficiently responsive so that complete and timely data can be turned into useful information.

Programs should evaluate the factors that impact timeliness, thoroughness, and completeness. Often resources can be used more efficiently and effectively by streamlining or redeveloping procedures in individual areas, such as case finding, data collection, and data processing (see Section 7.8 on Quality Improvement Methods).

Timeliness, thoroughness, and completeness are often intertwined and affect other quality assessments. For example, the quality control methods that evaluate case finding and case abstracting may include outcome measurements for timeliness and thoroughness. Data source evaluations include a timeliness measurement.

The program should:

- **Develop** productivity guidelines and standards.
- **Use** TACOMA criteria, especially 'oriented' and 'applicable', to assess the factors that challenge timeliness and completeness.
- **Use** computer technology to improve timeliness. For example, consider using the Internet for case reporting. Internet and electronic reporting also ease the burden of case reporting at data sources.
- **Monitor** timeliness.

7.6 Quality Control and Quality Assurance

'Quality control' (QC) and 'quality assurance' (QA) can be defined as a set of methods, activities, and procedures designed to improve the results of specific outcomes. For birth defects surveillance programs, these outcomes are related directly to surveillance functions, such as case ascertainment and data collection. Although active and passive case ascertainment systems may use different methods and procedures for improving data quality, the goal is the same, namely high-quality data.

Quality control is a retrospective and reactive approach to improvement that focuses on discovery and detection. Deficiencies and inaccuracies are found, resolved, and fixed so that final results or outcome measurements are accurate. As a result of QC procedures, high-quality data are created at the back end. In QC, the emphasis is on checking, investigating, containing, and adjusting (Dale and Bunney, 1999).

QC procedures may include re-case finding, re-abstracting, validity audits, timeliness monitoring, and data source evaluations. QC can also be used with data linkage, especially as this involves checking selected data fields, including birth weight, date of birth, name, etc. The results of QC procedures are used to evaluate, adjust, or correct the original data that were collected or the original circumstance that occurred.

Quality assurance is a proactive approach to improvement that focuses on prevention. Program functions are designed and activities are planned in advance to avoid inaccurate or deficient data. As a result of QA procedures, high-quality data are created at the front end or design stage. Often, the results of a QC method lead to QA activity. The QC method detects a deficiency, and the QA method redesigns the process to prevent its recurrence (Dale and Bunney, 1999).

QA procedures may include documentation (e.g., case finding, abstraction, medical records review, disease coding, data entry), the use of selective data sources, and the development and maintenance of the database infrastructure. Additionally, QA procedures can be implemented when specific outcome measurements require consistently high-quality data. Examples include (1) using an expert clinical reviewer to routinely evaluate case abstracts for data accuracy and thoroughness and (2) conducting medical records reviews to confirm a diagnosis prior to the data being used for projects like rapid case ascertainment, investigative inquiries, or statistical monitoring of trends. QA is cost efficient in the long run. Finding and solving problems can be time consuming and resource intensive, and unless the process is fixed, the same problems will continue to recur.

Maintaining high-quality data requires continual attention to improvement. Program performance is enhanced when quality improvement procedures are integrated into program operations and conducted in a consistent and systematic manner.

Refer to Section 7.8 (Quality Improvement Methods) in this chapter for specific examples of quality control and quality assurance applications.

The program should:

- **Maintain** documentation on program procedures, especially as these affect case ascertainment and data collection activities.
- **Record** and date decision items.

- **Identify** the sources of potential data quality issues and prioritize the impact of each on case ascertainment and surveillance. Some situations are provided in the 'quality issues' sections in other chapters in these guidelines.
- **Use** the TACOMA quality indicators to develop outcome measurements for evaluations. Of particular importance are quantitative evaluations of accuracy, completeness, and timeliness.
- **Design** meaningful evaluations, develop benchmarks, and track improvements. Quality assessments should be used to guide any decision to change or modify the program's practices and procedures.
- **Use** the results of quality control to design quality assurance procedures. Quality assurance is a self-propelling mechanism that ensures continual quality improvement.

7.7 Quality Control and Quality Assurance in the Surveillance Database

Computer technology provides many opportunities to implement quality control and quality assurance procedures. Computerization can promote standardization, perform queries on selected criteria, monitor timeliness, reduce duplication, and generate reports.

Quality assurance can be built into the design, development, maintenance, and expansion of the surveillance database. It is essential that the computer system address, at a minimum, the requirements of case ascertainment and data collection, data entry, information management, and statistical analysis. The system must also ensure security and privacy for the health information that is stored electronically (see Chapter 9 on Data Management and Security).

A database system should be documented thoroughly, with methods in place to track changes in procedures and processes and to identify security safeguards.

Standardization of data variables is an important quality assurance procedure. Data fields should have discrete definitions, and programs should standardize the information in a data field with unique codes or pre-formatted text. Drop-down windows can assist with this by providing choices and by placing limits on the options for the data field. Drop-down windows also prevent keying errors during data entry. Data fields can be programmed to perform logic checks for dates, time, age, gender-specific disease codes, and geographic information. Calculations can be programmed into data fields for measurements (such as weight, height, and head circumference) or can be programmed to complete a 'missing' measurement for a data field.

Software technology can also provide excellent resources for quality control. Procedures can be developed to monitor timeliness, productivity, and progress. Transaction logs can be used to monitor key activities and tasks. A posting-date field can be used to track staff entries as the case ascertainment process proceeds. Posting fields can also be used to monitor data source reporting trends, data collection activities, and data processing functions.

Any number of outcome measurements can be developed to track quality indicators, including measuring accuracy and completeness. Additionally, computer technology is uniquely suited to detect duplicate cases in the surveillance system. Information can be cross-linked on many different data fields, including name, date of birth, hospital of birth, mother's maiden name, etc.

There are almost limitless ways that computer technology can be used in quality control. The database integrates and supports surveillance activities. As such, the inputs and outputs of the database play a role in each TACOMA quality indicator. A well-designed database improves program efficiencies, outcome measurements, and data utilization (see Section 7.8 on Quality Improvement Methods).

The program should:

- **Identify** situations in case ascertainment and data collection where computer technology can be used to detect or prevent problems and to track measurements.

7.8 Quality Improvement Methods

Methods to measure and ensure high-quality data may vary depending on the approach to case ascertainment.

In active case ascertainment, field staff engage in the process of case identification, including gathering information and confirming a diagnosis for the case abstract. Quality control is directed at improving the way staff ascertain cases. In passive case ascertainment, the surveillance system receives case reports from data sources. Staff are not engaged in collecting the information on a case report. Additionally, a diagnosis reported on a case report is not usually confirmed prior to entry into the database. Therefore, in passive case ascertainment, quality control is directed at improving the results of the data collection process.

Although the ascertainment approaches are different, quality control and quality assurance methods can be used to achieve comparable levels of data quality across surveillance programs regardless of the ascertainment approach used.

While the list is not all inclusive, some of the methods used most frequently by birth defects surveillance programs for quality control are described below. Some are useful regardless of the case ascertainment approach and can be modified to suit the specific programmatic needs.

On the following pages we describe the following quality improvement methods in detail:

- Re-case finding
- Re-abstracting
- Validity audits and medical records reviews
- Clinical review
- Reliability and inter-rater agreement checks
- Timeliness measurements
- Data source evaluation
- Comparison/verification between multiple data sources
- Computer technology

Improving Quality through Re-Case Finding

Purpose

To evaluate the accuracy and comprehensiveness of the case-finding process.

Background

The case-finding process, used primarily in active case ascertainment, involves staff identifying potential birth defects cases at data sources.

Method

For re-case finding, perform the same steps and functions as for case finding. Develop procedures to evaluate results from the different pathways and steps in the process. Re-case finding should be conducted on a sample of information sources. The sample should consist of an appropriate number of entries, either from a single log or from multiple logs.

Outcome Measurements

- *Evaluation of results between the original case-finding activity and the quality control process.* This includes calculating the false positive and false negative rates at different steps in the case-finding process. In other words, this QC procedure evaluates the decision making that results in identifying a case versus a non-case.
  - Compare the QC list and the original staff review list of potential cases found at a data source during initial case finding. This is the list that identifies which cases go on to a medical records review and which do not.
  - Compare the results of re-reviewing the medical records. This involves QC re-reviewing medical records that were selected for review during the original case-finding activity and reviewing (for the first time) some medical records that were not on the original staff review list.
  - Determine the timeliness of the case-finding process.
- *Evaluation of compliance with case-finding procedures, including assessing decision-making skills.*

Frequency

It is important to develop a benchmark for re-case finding and to monitor outcome measurements periodically. The frequency with which re-case finding is conducted should be based on the demonstrated expertise and proficiency of the staff.

Quality Assurance

- Update case-finding procedures.
- Streamline the process to improve timeliness.

Tips

The case-finding process is a critical step in case identification. Not only is it important to evaluate staff effectiveness in identifying cases (and not missing any), it is also recommended that programs evaluate program efficiencies in case finding. For example, programs should evaluate the types of conditions that are considered potential cases. An evaluation might consist of determining how many confirmed diagnoses resulted from using a 'potential condition' in the initial steps of case finding. Some programs include ICD codes (i.e., searching through a hospital's disease index) as 'potential conditions'. An evaluation might consist of evaluating the effectiveness of searching using disease codes to identify a potential case in relation to whether specific codes were predictive in identifying a true birth defects case (i.e., an abstract is created).

Improving Quality through Re-abstracting

Purpose

To evaluate the accuracy and comprehensiveness of information that is entered on a case abstract form (hard copy or computer screen).

Background

Abstracting, used in active and passive case ascertainment, is the process of gathering and recording specific information from logs, medical records, or other information sources onto standard case abstract forms or computer screens.

Method

For re-abstracting, gather and abstract information from the same information source and record the data using the same abstract format (e.g., hard copy or computer screen). Re-abstracting should be conducted on a sample of information sources and a range of diagnosis categories

Outcome Measurements

- Comparison of the results of the quality control method to the results from the original case abstract and evaluation of the differences. Evaluation of the percentage and type of false positive cases.
- Identification of types and categories of errors or deficiencies. This may include disease coding, incomplete or missing information, and data entry errors. Includes the types of data variables that are problematic.
- Evaluation of compliance with abstracting procedures and guidelines.
- Determination of the timeliness of the abstracting process.

Frequency

It is important to develop a benchmark for re-abstracting and to monitor outcome measurements periodically. The frequency with which re-abstracting is conducted should be based on the expertise and proficiency demonstrated by staff.

Quality Assurance

- Update case abstracting guidelines.
- Provide training in disease coding, as applicable.
- Incorporate additional standardization into the data entry process. For example, provide drop-down windows to select and limit choices and to prevent key stroke errors.

Tips

Conduct an ***abstraction form review*** to identify differences and errors on completed abstraction forms. The abstraction forms should be checked for completeness, logic, and correct coding. Additionally, it is useful to categorize the types of data variables that are problematic to abstractors. For a given time period, QC should document, for each field staff member, the total number of abstraction forms reviewed and the number that have errors, such as incomplete or illogical data and incorrect coding.

Improving Quality through Validity Audits and Medical Records Reviews

Purpose

To evaluate the accuracy and comprehensiveness of a diagnosis that is reported by a data source or represented in a listing (e.g., hospital disease index) at a data source.

Background

In programs using passive case ascertainment, birth defect cases reported by data sources are accepted without confirmation. Active case ascertainment systems may use a listing of diseases provided by data sources, in disease-coded format, as part of case finding.

Method

The medical record, or other medical information report, is reviewed at the site or data source that reported the diagnosis or provided the diagnosis in a listing. This method is also used in the data sources audit.

Outcome Measurements

- *Predictive validity.* This is the degree to which an original measurement (e.g., reported diagnosis) successfully predicts a valid or confirmed outcome of interest. In other words, it represents agreement between the case report from the data source and the medical records review performed by surveillance staff.
- *Evaluation of missed diagnoses.* In other words, how many more diagnoses were identified by the medical records review process.
- *Identification of disease-coding issues,* especially as this pertains to data sources that report birth defects in a coded format (e.g., administrative databases such as the hospital discharge data set).
- *Incorporation of an evaluation of the data source* with the validity audit.
- *Timeliness* of the review process.

Frequency

Passive case ascertainment systems rarely have the resources to confirm all reported cases through medical records review. Therefore, the frequency of validity audits depends on program resources, requirements, and priorities. However, it is important to develop and maintain a certain level of validity audits. Programs should develop benchmarks, set goals, monitor results, and adjust program procedures.

Quality Assurance

- Identify and use data sources that report a confirmed diagnosis.
- Select diagnoses for consistent, concurrent, and timely validity audits. It is recommended that the diagnoses be from the set of birth defects that are reported to NBDPN. This QA procedure is primarily for passive case ascertainment systems.
- Identify disease codes that are problematic for describing birth defects precisely. Prioritize which ones should have a consistent validity audit. This QA procedure is applicable for passive case ascertainment systems that use the ICD-9-CM coding system, and can be adapted to accommodate active case ascertainment programs that use the hospital disease index during case finding.

Improving Quality through Validity Audits and Medical Records Reviews (continued)

Tips

Validity checks are a quality control tool. Although used primarily by passive case ascertainment systems, the tool is relevant for active ascertainment programs as well (e.g., active ascertainment key data entry systems or on-line abstracting). Validity checks in birth defects surveillance provide a way of evaluating the accuracy of what was reported (or represented) compared with what was 'validated' or confirmed after an investigation or medical records review.

Improving Quality through Clinical Review

Purpose

To review the diagnoses listed on the case abstract or in the case record for accuracy and plausibility.

Background

Information on birth defects cases is gathered and compiled by staff in active case ascertainment. Information is reported and collected from data sources in passive case ascertainment.

Method

Case abstracts or case records are examined by a designated clinical expert.

Outcome Measurements

- Evaluation of the information recorded on the case abstract.
- Identification of abstracting or coding problems.
- Percentage and types of agreement or disagreement with clinical review result.

Frequency

Programs should develop a benchmark for volume and types of case abstracts that should be reviewed and monitor the rate of agreement. In other words, the program should determine whether all case abstracts should be reviewed, or merely a percentage.

Quality Assurance

- Update and standardize abstracting and disease coding procedures.
- Train staff in the deficiencies cited and evaluate compliance concurrently.
- Increase the volume of clinical reviews, as required.

Tips

A clinical reviewer should be proficient at disease coding since the literal text of the diagnosis needs to be translated into the most accurate disease code.

In passive case ascertainment, the medical records from all data sources that reported a diagnosis for a respective birth defect case should be available to the clinical reviewer. Document the policies and procedures for the clinical review to ensure standardization. Include instructions for assigning the disease code.

Improving Quality through Reliability and Inter-Rater Agreement Checks

Purpose

To evaluate rate of agreement between two or more persons for the outcomes of interest.

Background

Results of case ascertainment and data collection should be consistent, especially when staff are required to make abstracting decisions.

Method

- *Dual-entry coding system (double-checking of assigned code).* At least two coders assign codes from the same list of diagnoses.
- *Dual-entry data entry*. At least two staff key information from the same case abstract into the surveillance database.
- *Dual clinical review*. At least two clinical reviewers examine the same abstracts and provide results.
- *Dual medical records reviews*. At least two staff review the same medical records and abstract information per program procedures. This may include evaluating disease code assignments. Some passive case ascertainment programs may benefit by including a clinical expert in this inter-rater reliability evaluation.

Outcome Measurements

- Rate of agreement
- Type of deficiencies
- Compliance with abstracting and other program procedures

Frequency

Programs should develop benchmarks and periodically evaluate for continued consistency.

Quality Assurance

- Write precise procedures.
- Develop decision-making flow charts.
- Train staff with respect to addressing any deficiencies noted.
- Develop standardized data definitions for each data element. When applicable, develop a list of acceptable responses for a data element. Use drop-down windows to facilitate selecting from a list.
- Use technology to increase the accuracy of abstracting and data entry.

Tips

Keep a log of decision-making items and make sure it can be referred to easily. This is important for abstracting and coding procedures. Update procedure manuals, date-stamp all changes. When disease reporting rules or procedures change, make the changes effective as of the beginning of a calendar year.

Improving Quality through Timeliness Measurements

Purpose

To evaluate rapidity and readiness.

Background

All areas of case ascertainment and data collection affect how responsive the program is in meeting goals and objectives with respect to timeliness.

Method

The time interval between two or more points of interest is measured. Often the measurement is from one task to the next or from start to finish.

Outcome Measurements

Timeliness measurements can be used to evaluate and improve many areas within a surveillance program including productivity and program performance. Examples include:

- *Reporting time lags*. A measurement of the time it takes for a case report to be received-in or identified-to the birth defects program.
- *Case-finding process*. An evaluation of the time it takes to identify a case, review the medical record(s), abstract information, and complete the abstract.
- *Data processing time lags*. A measurement of how quickly information is processed for use.

Frequency

Timeliness can be evaluated readily. Tracking measurements can be monitored using software technology and developing date-posting fields. Queries, internal logs, and reports can facilitate this quality improvement method.

Quality Assurance

- Implement changes to case ascertainment procedures or processes to improve timeliness.
- Use laptops to reduce redundant steps.
- Work with data sources to improve consistency in reporting, including using electronic case reporting and Internet reporting.
- Develop computer transaction logs.

Tips

Evaluate the program's desired outcome measurements in relation to how long it takes to achieve them. Use the criteria in TACOMA, especially as they relate to improving timeliness. For example, the criteria 'oriented' and 'applicability' focus on selecting data variables that are important to the program. Include an evaluation of the reasons for unfinished case abstracts or case records. Data variables that consume a great deal of resources to collect should be re-evaluated for intent and usefulness.

Improving Quality through Data Source Evaluation

Purpose

To ensure that birth defect case reporting is complete, accurate, appropriate, and within the guidelines for timely reporting.

Background

Birth defects are found-at or reported-from data sources. Data sources vary in purpose, organizational structure, and scope.

Method

The source of the diagnostic information is evaluated for accuracy, completeness, and timeliness. This method may combine the methodology of other procedures, such as validity audits and timeliness measurements, and may also include re-case finding.

- *Accuracy.* The medical record, or other medical information report, is reviewed at the site or data source that reported the diagnosis or provided the diagnosis in a listing (see Validity Audits).
- *Accuracy.* For a large data source, such as hospital discharge data (an administrative data base), the audit may be designed to focus on a suspected hospital or unusual patterns of disease-code use.
- *Timeliness.* Time lags for reporting are evaluated per data source.
- *Completeness.* Passive case ascertainment utilizes the steps taken in active case-finding to identify all of the potential and confirmed cases of birth defects at the data source. This procedure is more difficult for passive case ascertainment to implement because of the staff resources needed to conduct comprehensive case-finding (see Chapter 6 on Case Ascertainment Methods).

Outcome Measurements

- *Completion of the descriptive assessment* of the data source.
- Refer to *validity audits* and *timeliness audits*.
- *False positive rate*. What is the level of diagnostic quality from a data source?
- *Completeness rate*. What is the rate of missed individuals with birth defects? These are individual cases that were not reported-to or identified-at the data source.
- *Evaluation of data collection methods*. Is the format used for reporting cases contributing to missed case reports?

Frequency

Each data source should be evaluated at least once to assess a level of quality.

Quality Assurance

- Use multiple data sources. One data source rarely provides comprehensive information.
- Use data sources that report a confirmed diagnosis.
- Involve the data source in discussions related to quality indicators. Develop mutually agreed-upon strategies for resolving issues.
- Encourage data sources to report cases in an electronic format, including using the Internet. This may improve timeliness and completeness. Confidentiality and privacy can be assured via encryption and other safeguards.

Improving Quality through Data Source Evaluation
(continued)

Tips

Staff from programs using passive case ascertainment often review medical records in medical records departments, and some review autopsies at pathology departments. However, these staff usually do not engage in case-finding (i.e., combing through information sources to find potential cases of birth defects). Passive ascertainment staff should engage the data source in discussions prior to a case-finding audit. It is important to involve staff at the data source in planned activities to answer their questions. A contact person at the data source should be identified to ensure minimal disruption of normal work flow once the case-finding process begins.

Improving Quality through Comparison/Verification Between Multiple Data Sources

Purpose

To compare diagnosis, and other information, that is reported-from or identified-at different data sources.

Background

Programs are encouraged to use multiple data sources for case ascertainment. A single data source is rarely able to provide comprehensive or accurate information.

Method

- Compare information that is collected from multiple data sources in order to determine what information is accurate and complete. Examples include:
  - *Confirm or invalidate a diagnosis* based on a higher level of diagnostic expertise or clinical specialty. For passive case ascertainment this could mean that a diagnosis that is reported from a high-quality data source is considered to be confirmed or valid.
  - *Clarify an incomplete or imprecise diagnosis*. Conduct follow-up to gather better information.
  - *Identify incomplete data fields* on the case abstract or case record. Some data sources may not have complete information on a birth defect case, which results in an incomplete or deficient case report.
  - *Update the case abstract or case record* with more timely information. This includes address, names, and contact information.
- Develop procedures to identify duplicate case abstracts or case records in the database. Common situations that result in duplicate case abstracts or case records are mistakes with date of birth, use of multiple or incomplete names, and adoptions.

Outcome Measurements

- Rate the data sources. Assign 'quality' grades for specific criteria (e.g., diagnosis quality, complete address).
- Evaluate the value-added benefit that a data source provides. For example, if two data sources identify the same cases but one source provides a higher total volume of cases, evaluate the rationale for using both data sources.

Frequency

The use of multiple data sources is strongly encouraged. However, a surveillance program needs to understand the potential differences in quality among data sources and adjust procedures accordingly. In active case ascertainment, the comparison and verification of information can be done in an ongoing manner. In passive case ascertainment, where each case report may not be read by staff upon receipt, a benchmark should be established and key factors evaluated. At a minimum, comparison and verification should be done annually; otherwise the volume of inconsistencies or differences may turn into a resource-intensive effort to reconcile them. Computer technology greatly enhances a program's ability to systematically conduct comparison and verification procedures.

Improving Quality through Comparison/Verification Between Multiple Data Sources (continued)

Quality Assurance

- Combine or merge data that are collected or abstracted into a central case abstract or case record upon receipt. This minimizes the possibility of creating a duplicate abstract or record and reduces redundant staff work.
- Develop data linkage procedures for the large administrative, computerized data sets, such as vital records, hospital discharge data, hospital disease index, and Medicaid. Data linkage can also be developed to accommodate smaller clinic-based information systems, such as cytogenetics laboratories, genetic services, and specialty clinics. A key factor in data linkage is using standardized data variables (see Chapter 4 on Data Collection Variables).
- Develop decision-making and hierarchy models for use in comparison and verification of data elements. Programs should determine which data sources are considered a high-quality information source for specific data variables.

Tips

This QC method is enhanced by using computer technology and developing a systematic approach.

Improving Quality through Computer Technology

Purpose ***To use technology in quality improvement efforts.***

Background Surveillance systems are information management systems whose operations are enhanced by computer technology (see Chapter 9 on Data Management and Security).

Method

- *Quality assurance.* Use software to prevent problems and enhance standardization.
  - *Build in range checks* to prevent inaccurate abstracting and data entry. These checks can be created for any data variable with a defined parameter of acceptable or measurable values. Date range checks can be used for age, date of birth, date of fetal demise, date of death, LMP (date of last menstrual period). These become the dates that other dates (e.g., date of case report) are compared to for rationale. Other types of range checks are Apgar scores, gestational age, and birth weight.
  - *Develop automated calculations and conversions* for specific data fields. Examples include birth weight, time and LMP .
  - *Promote the use of coded data.* Develop codes for text information. This method can be applied to any data variable definition that has multiple acceptable responses. Examples include disease, geographic, race, and ethnicity codes. Programs can develop code sets for data sources, specific sites, types of procedures, family history, physicians, etc
  - *Use drop-down windows for data fields.* This approach is useful with long text entries and for text that has been converted to a code.
  - *Use standard data collection variables (and data definitions)*, to accommodate record linkage and electronic transfers (see Chapter 4 on Data Collection Variables).
- *Quality control*. Develop procedures to detect, measure, and enhance effectiveness.
  - *Perform logic edits*. Review existing program documentation and syntax to ensure that the computer application is performing as intended. For example, when computer applications are used to convert or calculate data field values, make sure the results using the formula(e) are accurate.
  - *Create date-posting fields* to monitor timeliness.
  - *Develop transaction logs*. This is a method that tracks and dates additions, deletions, and other changes to the database.
  - *Create queries and reports* to track desired outcome measurements.
  - *Develop methods, using key data variables, to find duplicate cases* in the database.
  - *Develop queries to identify problem situations*. Examples include:
    – Some birth defects should not be counted due to prematurity or low birth weight.
    – Some ICD codes are problematic for birth defects.
  - *Develop information management systems* to improve the efficiency of program operations, including case ascertainment.
  - *Develop methods to improve timeliness* of case reporting. This includes using Internet reporting and other electronic methods, with appropriate security measures to protect confidentiality and privacy.

Improving Quality through Computer Technology
(continued)

Outcome Measurements

Track measurements from the QC and QA methods that are developed.

Frequency

Once developed, computerized quality procedures can be run on a consistent and systematic timeframe. Systems and software also facilitate flexibility for ad hoc queries and reports. Information management systems are ongoing system enhancements.

Quality Assurance

Design, maintain, and update to:

- Prevent problems at the source
- Promote standardization
- Improve program efficiencies, including timeliness
- Facilitate data retrieval and analysis
- Assist in tracking measurements

Special considerations for passive case ascertainment programs

A dilemma that primarily affects passive case ascertainment programs is how to retain the integrity of the database, while also resolving data quality problems. In other words, how do programs identify and use accurate information, especially since the majority of diagnoses in the data base are accepted as reported (i.e., not confirmed by staff)? For example, if a diagnosis is reported from a data source and is determined to be inaccurate, or incomplete, it should not be counted in statistical analysis. However, for epidemiological and evaluation purposes, this diagnostic information (and the associated information that accompanies the case report) should not be deleted from the database. A method to resolve this issue could be to develop a mechanism (perhaps a data field) that identifies or flags a diagnosis that is not accurate (valid) or should not be counted. Programs are encourage to develop methods to resolve these kinds of issues in a way that best suits the program's needs.

Tips

- Before building a computerized data collection system evaluate the current manual data collection instrument to determine what works and what doesn't work.
- Avoid programming an on-line data collection system based on your hard copy instrument. Once programming is completed, it is often difficult to undo.
- Prior to developing and expanding the data base, evaluate the program's needs, i.e., how the data will be used, how the data will be accessible, data transfer, etc.

7.9 References

The North American Association of Central Cancer Registries, Inc. (NAACCR). *Standards for Cancer Registries, Volume III: Standards for Completeness, Quality, Analysis, and Management of Data.* Springfield, IL: Registry Operations Committee of NAACCR; September 2000.

Dale B, Bunney H. *Total Quality Management Blueprint.* Malden, MA: Blackwell Publishers Inc.; 1999.