TECHNICAL DOCUMENT

Data quality monitoring and surveillance system evaluation

A handbook of methods and applications

www.ecdc.europa.eu

ECDC TECHNICAL DOCUMENT

Data quality monitoring and surveillance system evaluation

A handbook of methods and applications

This publication of the European Centre for Disease Prevention and Control (ECDC) was coordinated by Isabelle Devaux (senior expert, Epidemiological Methods, ECDC).

Contributing authors
John Brazil (Health Protection Surveillance Centre, Ireland; Section 2.4), Bruno Ciancio (ECDC; Chapter 1, Section 2.1), Isabelle Devaux (ECDC; Chapter 1, Sections 3.1 and 3.2), James Freed (Public Health England, United Kingdom; Sections 2.1 and 3.2), Magid Herida (Institut for Public Health Surveillance, France; Section 3.8), Jana Kerlik (Public Health Authority of the Slovak Republic; Section 2.1), Scott McNabb (Emory University, United States of America; Sections 2.1 and 3.8), Kassiani Mellou (Hellenic Centre for Disease Control and Prevention, Greece; Sections 2.2, 2.3, 3.3, 3.4 and 3.5), Gerardo Priotto (World Health Organization; Section 3.6), Simone van der Plas (National Institute of Public Health and the Environment, the Netherlands; Chapter 4), Bart van der Zanden (Public Health Agency of Sweden; Chapter 4), Edward Valesco (Robert Koch Institute, Germany; Sections 3.1 and 3.2).

Project working group members: Maria Avdicova (Public Health Authority of the Slovak Republic), Sandro Bonfigli (National Institute of Health, Italy), Mike Catchpole (Public Health England, United Kingdom), Agnes Csohan (National Centre for Epidemiology, Hungary), Yves Dupont (Scientific Institute of Public Health, Belgium), Irena Klavs (National Institute of Public Health, Slovenia), Anna Kurchatova (National Centre of Infectious and Parasitic Diseases, Bulgaria), Mathias Leroy (Scientific Institute of Public Health, Belgium), David Mercer (World Health Organization, Regional Office for Europe), Zsuzsanna Molnar (National Centre for Epidemiology, Hungary), Pierre Nabeth (World Health Organization), Elvira Rizzuto (National Institute of Health, Italy), Malgorzata Sadkowska (National Institute of Public Health, Poland), Gudrun Sigmundsdottir (Centre of Health Security and Communicable Disease Prevention, Island), Ewa Staszewska (National Institute of Public Health, Poland), Reinhild Strauss (Federal Ministry of Health, Austria).

Contributing ECDC experts: Cristian Avram, Arnold Bosman, Sergio Brusin, Bruno Ciancio, Denis Coulombier, Isabelle Devaux, Erika Duffell, Ana-Belen Escriva, Rodrigo Filipe, Graham Fraser, Keith Hodson, Frantiska Hruba, Charles Johnston, Marius Mag, Vladimir Prikazsky, Carmen Varela-Santos.

Acknowledgements
We would also like to thank the following individuals for their help, input and feedback: Elliot Churchill, Larisa Fedarushchanka, Lisa Ferland, Carmen Emily Hazim (all Public Health Practice, LLC; United States of America), Christopher Williams (Public Health England, United Kingdom), Michaela Dierke (Robert Koch Institute, Germany).

Suggested citation: European Centre for Disease Prevention and Control. Data quality monitoring and surveillance system evaluation – A handbook of methods and applications. Stockholm: ECDC; 2014.

Stockholm, September 2014

ISBN 978-92-9193-592-5

doi 10.2900/35329

Catalogue number TQ-04-14-829-EN-N

Contents

Abbreviations vii

1 Introduction 1
1.1 Background 1
1.1.1 Monitoring data quality 1
1.1.2 Evaluating surveillance systems 2
1.2 Surveillance attributes 2
1.2.1 Completeness and validity 2
1.2.2 Sensitivity, specificity, positive predictive value and negative predictive value 3
1.2.3 Timeliness 5
1.2.4 Usefulness 5
1.2.5 Representativeness 5
1.2.6 Other surveillance attributes 6
1.2.7 Development of data quality indicators 6
1.3 Description of a surveillance system 7
1.3.1 Surveillance objectives 8
1.3.2 List of diseases under surveillance and case definitions 9
1.3.3 Data sources and data flow 10
1.3.4 Surveillance networks 11
1.3.5 Populations under surveillance 12
1.3.6 Geographic coverage 12
1.3.7 Types of surveillance 12
1.3.8 Specification of the information to be reported 13
1.3.9 Reporting format 14
1.3.10 Data entry 14
1.3.11 System architecture 15
1.4 References 16

2 Monitoring data quality 17
2.1 Conceptual framework 17
2.1.1 Underreporting and under-ascertainment 17
2.1.2 Setting up a data quality monitoring system 19
2.2 Monitoring completeness 23
2.2.1 Describe the system 23
2.2.2 Plan the monitoring process 25
2.2.3 Measure internal completeness and summarise results 26
2.2.4 Interpret the results of the measurement 27
2.2.5 Make suggestions for remediation 28
2.3 Monitoring internal validity 32
2.3.1 Describe the system 32
2.3.2 Plan the monitoring process 33
2.3.3 Measure internal validity and summarise the results 34
2.3.4 Interpret the results of the measurement 36
2.3.5 Make suggestions for remediation 37
2.3.6 Questions and answers 39
2.4 Monitoring timeliness 40
2.4.1 Describe the system 40
2.4.2 Plan the monitoring process 41
2.4.3 Measure timeliness 42
2.5 References 44

3 Evaluating surveillance systems 45
3.1 A stepwise approach to evaluation 45
3.2 Plan the evaluation 45
3.2.1 Identifying triggers 45
3.2.2 Selecting the evaluation type and method 48
3.2.3 Engaging stakeholders 50
3.2.4 Setting up a team for surveillance evaluation 50
3.3 Evaluation of external completeness 51
3.3.1 Introduction 51

3.3.2 Describing the system....52
3.3.3 Planning the evaluation....52
3.3.4 Evaluating external completeness....54
3.3.5 Interpreting the results of the evaluation....58
3.3.6 Improving external completeness....59
3.4 Evaluation of sensitivity, specificity and positive predictive value....60
3.4.1 Factors which influence the sensitivity of a surveillance system....61
3.4.2 Methods....61
3.4.3 Results....63
3.4.4 Interpreting the results of an evaluation....63
3.4.5 Enhancing sensitivity, specificity and positive predictive value....64
3.4.6 Questions and answers....64
3.5 Evaluation of external validity....67
3.5.1 Identifying duplicate records....67
3.5.2 Calculating concordance through record linkage....67
3.5.3 Calculating concordance via follow-up of notified cases....69
3.5.4 Validity checks during statistical analysis....69
3.5.5 Validity of outbreak notifications....69
3.5.6 Indirect methods to evaluate validity....70
3.5.7 Ethical considerations....70
3.5.8 Interpretation and modifications....70
3.5.9 Questions and answers....72
3.6 Evaluation of usefulness....73
3.6.1 Introduction....73
3.6.2 Identifying expected uses of a disease surveillance system....73
3.6.2 Creating a user list....73
3.6.3 Documenting system usage....74
3.6.4 Remediation....74
3.7 Evaluating the costs of disease surveillance....75
3.7.1 Background....75
3.7.2 Classification of costs of public health surveillance....75
3.7.3 Cost analysis....76
3.7.4 Discussion....77
3.8 References....79

4 Intervention....84
4.1 Motivation of data providers for improving data quality....84
4.1.1 Why improve data quality?....84
4.1.2 How to motivate data providers to deliver good quality data....84
4.2 Improving surveillance system quality....85
4.2.1 How to motivate for surveillance system improvement....85
4.2.2 Reinforcing monitoring....86
4.2.3 Structuring output activities....87
4.3 References....89

Glossary....90

Figures

Figure 1.2a. Levels of case detection in the surveillance process for communicable diseases....5
Figure 1.3a. Key elements for the description of a surveillance system....8
Figure 1.3b. Simplified flowchart of a generic surveillance system (indicator-based)....11
Figure 2.1a. Possible steps to identify non-visible cases who do not seek healthcare....18
Figure 2.1b. General practitioners, case management and surveillance systems....20
Figure 2.2a. Steps designing and implementing a monitoring system for internal completeness....23
Figure 2.2b. Checklist: main considerations for the establishment of a monitoring system....25
Figure 2.2c. Methods at the data entry level for enhancing internal data completeness....30
Figure 2.3a. Monitoring internal validity....32
Figure 2.3b. Checklist: main considerations/decisions about internal validity when setting up a monitoring system33
Figure 2.3c. Surveillance systems-related factors which can negatively influence the validity of surveillance data..37
Figure 2.3d. Basic concepts for creating a reporting form....38
Figure 2.3e. Examples of validity checks....38
Figure 2.4a. Steps for monitoring timeliness....40
Figure 2.4b. Scope of a reporting system to measure timeliness....41
Figure 2.4c. Steps for monitoring timeliness....42
Figure 2.4d. Monitoring timeliness: main time delays....42
Figure 2.4e. Important timeliness elements in the Irish CIDR system....43
Figure 3.1a. Evaluating surveillance systems....45
Figure 3.2a. Selecting an appropriate evaluation type....48
Figure 3.3a. Key points for the evaluation of external completeness....51
Figure 3.3b. Venn diagram for a two-source capture–recapture study....56
Figure 3.3c. Venn diagram for a three-source capture–recapture study....57
Figure 3.4a. Key points for the evaluation of sensitivity, specificity and positive predictive value....61
Figure 3.5a. Key points for evaluating external validity....67
Figure 3.5b. Schematic presentation of the steps of a study of concordance between surveillance data and medical records (external data source)....68
Figure 3.5c. Approaches to increase data validity in a disease surveillance system....72
Figure 4.1a. Key points....84
Figure 4.2a. Key points....85
Figure 4.2b. Process to reinforce monitoring of data quality of surveillance system....87

Tables

Table 1.1a. Completeness of HIV data reported by countries of the WHO European Region (total: 53) in 2009 and 2010 2
Table 1.1b. Completeness of data on *Mycobacterium tuberculosis* infection, EU Member States, 2010 2
Table 1.2a. Two dimensions of completeness and validity 3
Table 1.2b. Sensitivity, specificity, positive predictive value, negative predictive value for a case reporting system/outbreak reporting system 3
Table 1.2c. Examples of data quality indicators 6
Table 1.3a. Data sources for indicator-based surveillance 10
Table 2.1b. Validity rules for data collected from the registered medical practitioner (RMPs) in the United Kingdom . 21
Table 2.2a. Checklist of background information needed when designing a monitoring system for internal completeness 24
Table 2.2b. Table summarising the monitoring plan for internal completeness 26
Table 2.2c. Table summarising the monitoring plan for internal completeness of common variables 26
Table 2.2d. Examples of fields that are not completed at initial notification 27
Table 2.2e. Completeness results by variable 27
Table 2.2f. Completeness results by record 27
Table 2.2g. Checklist of possible reasons for low completeness of the variable 'culture result' for tuberculosis notification 28
Table 2.2h. Examples of possible mandatory fields by disease 30
Table 2.3a. Checklist: necessary background information when designing a monitoring system for internal validity .. 33
Table 2.3b. Examples of validation rules and validation messages in TESSy 34
Table 2.3c. Number of notified cases and expected number of cases for a specific time period (e.g. week, month) and geographic area (e.g. whole country, region); based on background data 35
Table 2.3d. Proportion of notifications of specific diseases and expected proportion of notifications for a specific time period and geographic area; based on background data 35
Table 2.3e. Reported incidence and expected disease incidence in a specific population for a given time period; based on background data 35
Table 2.3f. Number of recognised false alarms/errors by health outcome for a specific time period 36
Table 2.3g. Validity results by health outcome and variable for a specific time period 36
Table 2.3h. Validity results by health outcome for a specific time period 36
Table 3.2a. Examples of triggers leading to an evaluation 46
Table 3.2b. Triggers for the evaluation of surveillance systems: suggested definitions and categories 47
Table 3.2b. Types of protocols 48
Table 3.2c. Examples of attributes to be evaluated and possible type of evaluation for specific triggers 49
Table 3.3a. Important information to be included in the evaluation protocol 52
Table 3.3b. Important information to be included in the evaluation protocol 54
Table 3.3c. Checklist of possible reasons for low external completeness of surveillance systems 58
Table 3.4a. Sensitivity (Se) and positive predictive value (PPV) of surveillance system X 63
Table 3.5a. Evaluation of validity of data in surveillance system X, comparison between surveillance data and data from a reference source 68
Table 3.5b. Absolute mean difference between dates in the notification dataset and in the medical record (or alternative data source) 69
Table 3.5c. Validity of reports of outbreaks of specific diseases 69
Table 3.5d. Summary of the validation process 71
Table 3.6a. Sample user list for data generated by the health surveillance system 73
Table 3.8a. Conceptual framework of public health surveillance 78

Abbreviations

CDC	United States Centers for Disease Control and Prevention
LIMS	Laboratory information management system
NOIDS	Notifications of Infectious Diseases
NPV	Negative predictive value
NSC	National surveillance coordinator
PPV	Positive predictive value
Se	Sensitivity
Sp	Specificity
TESSy	The European Surveillance System

1 Introduction

The European Surveillance System (TESSy) collects data on 52 infectious diseases and conditions collected through nominated disease-specific networks from 30 countries, whose disease surveillance systems differ substantially. This leads to problems with regard to data quality and comparability which can influence the quality of analysis of surveillance data.

Differences in notification rates are another factor that complicates data analysis. Differences in notification rates between countries can be explained by epidemiological factors, but also by differences in the national surveillance systems.

Data comparability at the national, sub-national and EU level can be improved by collecting appropriate information on the characteristics of surveillance systems, the quality of the data they generate, and their performance.

1.1 Background

In 2010, ECDC conducted a survey [1] among national surveillance coordinators (NSCs) which revealed substantial differences in practices with regard to monitoring data quality in EU Member States. The results of this survey were presented at the annual meeting of the NSCs in October 2010, where it was also decided to produce a handbook on monitoring the quality of disease surveillance data for public health professionals. Meeting participants also agreed that a working group composed of NSCs and surveillance experts working in national public health institutes should guide the development of this handbook. A proposal to improve data comparability at the EU level – in line with the objectives of the ECDC long-term surveillance strategy [2] – was presented to the ECDC Advisory Forum on 5 May 2011. Subsequently, a working group of 22 disease surveillance experts from 16 countries was established with a mandate to develop a handbook on data quality monitoring and surveillance system evaluation.

The overall objective of this project was to support processes for monitoring data quality and evaluating surveillance systems in EU/EEA Member States in order to provide accurate and timely information for decision-making. The handbook is aimed to support the daily work of public health professionals working with surveillance data on communicable diseases.

This handbook is based on input from the working group on surveillance system quality. In addition, a review of available literature was conducted.

A first technical meeting of the working group on surveillance system quality took place in Budapest on 27 and 28 October 2011, where an outline of the handbook was agreed upon. Between October 2011 and February 2012, nine subchapters were drafted. A second technical meeting of the working group was held in Stockholm on 2 and 3 April 2012 to plan the next steps of the project.

This handbook:

- serves as guidance for surveillance system evaluation and data quality monitoring;
- supports self-assessments of surveillance systems at various levels from local to national and international;
- introduces a bottom-up approach to evaluation, starting at the level of the data provider;
- provides guidance in a practical and intuitive way; and
- includes a detailed glossary.

In addition, the approach developed in this handbook takes into account the following points:

- While monitoring is a continuous process and focuses on data quality, evaluation takes place at specific points in time and addresses a broader scope of issues than just data quality.
- Before deciding on an evaluation model suitable for the surveillance system in question, it is essential to establish drivers and triggers that would initiate the monitoring and evaluation process.

1.1.1 Monitoring data quality

It is important to monitor data quality and thus ensure that the collected data are meaningful so they meet the objectives of local, national and international surveillance systems. The quality of the initial data may determine the data quality at all stages of the reporting process. Monitoring data quality also helps to improve data analysis and interpretation in public health reports at all levels.

The two examples below show that results of HIV and TB surveillance data collected at the European level need to be interpreted in the context of completeness. Table 1.1a shows the results for the completeness of HIV variables collected in The European Surveillance System (TESSy), a database operated by ECDC [3]. Records for age and gender show a high level of completeness. However, completeness for stage (35%) and CD4 cell count (27%) is low.

Table 1.1a. Completeness of HIV data reported by countries of the WHO European Region (total: 53) in 2009 and 2010

	2009		2010	
Variables	Number of countries	Completeness (%)	Number of countries	Completeness (%)
Age	49	99	49	99
Gender	49	100	49	99
Transmission	48	84	48	87
Stage	43	35	43	33
CD4 cell count	23	27	25	31

Table 1.1b. Completeness of data on *Mycobacterium tuberculosis* infection, EU Member States, 2010

Variables	Number of countries	Number of cases with information	Completeness (%)
Age	29	73 221	99
Gender	29	73 812	100
Previous treatment	29	67 996	92
Culture result	26	59 637	81
Drug susceptibility testing (isoniazid and rifampicin)	26	26 158	73
HIV status	15	17 650	24

1.1.2 Evaluating surveillance systems

Most guidelines for the evaluation of surveillance systems in the literature use an approach that is attribute-oriented [4,5]. The following additional questions guided the approach developed in this handbook:

- What are the components of a surveillance system and how do they interact?
- What triggers the evaluation of a surveillance system?
- Which evaluation methods are appropriate?
- Which components should be evaluated?
- What do the results of the evaluation tell us?
- What are the possible interventions?

1.2 Surveillance attributes

This section provides definitions of surveillance systems attributes based on previous guidelines for the evaluation of surveillance systems [4-9], feedback from ECDC's national surveillance coordinators who participated in a survey in 2010 [10], and working group discussions on surveillance system quality.

In the survey 'Mapping current practices for monitoring and evaluating quality of surveillance data in EU/EEA Member States' [1], definitions focused primarily on data quality. Data quality is usually defined as the completeness and validity of the data recorded in a public health surveillance system [4,5]. However, other attributes, like sensitivity, representativeness and timeliness can also be used to evaluate the quality of a surveillance system.

In this handbook, definitions are not limited to completeness and validity and also include other surveillance attributes which can be relevant for the development of a common approach to monitoring data quality and evaluation of surveillance systems.

1.2.1 Completeness and validity

Completeness and validity both have an internal and an external dimension.

Completeness

Completeness can be considered as having two separate dimensions:

Internal completeness refers to whether there are missing and/or unknown data fields in a surveillance database and can be defined as 'the number of completed data fields out of the total number of data fields' (unknown and missing items should be included in the denominator).

External completeness relates to whether the data available to the surveillance system reflect the true number of cases affected by a given condition [4-6]. External completeness applies to the reporting process only and is equivalent to 'sensitivity of reporting' as described below. It can be a way to estimate underreporting of

surveillance data, but is does not measure under-ascertainment. One approach to evaluate external completeness consists of comparing at least two datasets from different sources of information that are supposed to provide surveillance information on the same disease (e.g. laboratory and notification data for case reporting of salmonellosis). A common method to measure external completeness is 'capture–recapture' [see Section 3.2]. However, other methods (for example the review of medical records) can also be used to compare datasets, depending on the disease under surveillance, the nature and accessibility of data sources, and other parameters.

Validity

In the context of surveillance, validity describes the ability to capture the 'true value' of the disease burden, such as incidence or prevalence, which is useful for the analysis of surveillance data. The 'true value' should be viewed in the context of a surveillance system. Validity may relate only to a limited number of cases, e.g. those diagnosed by health services covered by surveillance systems.

Validity has both an internal and an external dimension:

- **Internal validity** relates to the extent of errors within the system, for example, coding errors in translating from one level of the system to the next.
- **External validity** relates to whether the information recorded about the cases is correct. Evaluating external validity implies a surveillance indicator measured against a 'gold standard' value [6]. One possible way to conduct a validity study is to compare the recorded data to the original medical records. If data on the same patient/population are recorded at different points in time for the same information (disease/variable), differences can be due to a 'real change' over time, to chance, or to a bias in the measurement.

Table 1.2a. Two dimensions of completeness and validity

	Completeness	Validity
Internal	Proportion of complete data fields or values within a dataset.	Extent of errors within an information system and one data source.
External	Relates to whether the data available to the surveillance system reflect the true number of cases diagnosed with a condition in the population under surveillance. Can also be referred to as *sensitivity*.	Relates to whether the information recorded about cases is correct and exact when compared to an external database. This may require a comparison to a gold standard.

1.2.2 Sensitivity, specificity, positive predictive value and negative predictive value

Sensitivity (Se) is linked to specificity (Sp), positive predictive value (PPV) and negative predictive value (NPV). These four aspects are usually used in connection with the evaluation of a case definition. These values can be calculated for a case reporting system or for an outbreak reporting system (Table 1.2b).

Table 1.2b. Sensitivity, specificity, positive predictive value, negative predictive value for a case reporting system / outbreak reporting system

	Cases (condition present) / outbreak	Non case (condition not present) / no outbreak	
Reported	a True case True outbreak	b False case False outbreak	a+b
Not reported	c False non-case False 'no outbreak'	d True non-case True 'no outbreak'	c+d
	a+c Total cases Total outbreaks	b+d Total non-cases Total 'no outbreak'	N = a+b+c+d

Se = a/(a+c); PPV = a/(a+b); Sp = d/(b+d); NPV = d/(c+d)
N: population under surveillance

Sensitivity

Sensitivity in case-based disease surveillance refers to the proportion of cases in a population that are notified through the surveillance system [4]. Sensitivity of the surveillance case definition refers to the ability of the case definition to correctly classify cases to which it is applied. The sensitivity of the surveillance system is the number of cases reported by the surveillance system or 'true cases' (a), divided by the number of cases (a + c) in the community. Under certain circumstances, cases may not be reported to the surveillance system:

- An infected but asymptomatic person who does not seek medical attention will not get diagnosed and reported to the system.
- The public health reporting system can be responsible for the surveillance system receiving incomplete information: case definitions can be too specific, the healthcare system does not offer certain diagnostic tests, or cases were not classified (i.e. the case definition was not applied).

For event-based surveillance, sensitivity refers to the proportion of outbreaks occurring in a community that are picked up by the surveillance system.

Positive predictive value

Positive predictive value (PPV) in case-based surveillance is the proportion of real cases (*a*, 'true positive cases') reported through the surveillance system, divided by the total number of cases reported to the surveillance system:

$$PPV = a/(a+b)$$

The PPV reflects the probability that a case reported in the surveillance system is indeed a real case. For event-based surveillance, PPV reflects the probability that a detected outbreak is a bona fide outbreak.

Specificity

Specificity does not apply to case-based surveillance systems. Specificity, however, can be applied to event-based surveillance, where it refers to the number of weeks for which no outbreaks were correctly detected by the system (true negatives) in relation to the number of weeks without outbreaks (true negatives + false positives) as reported through other sources.

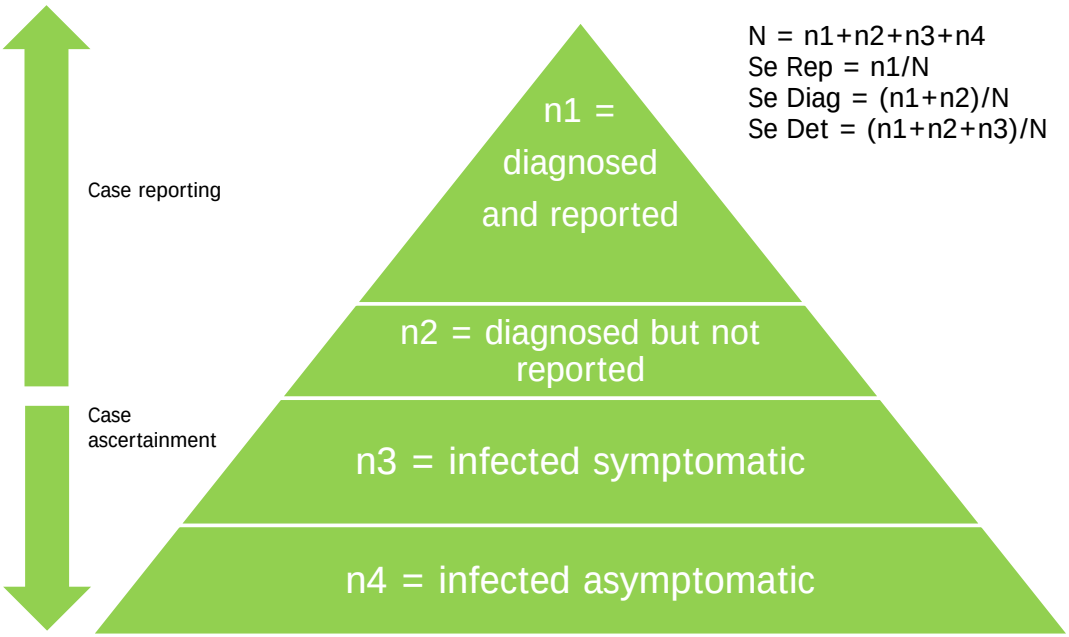
Negative predictive value

Negative predictive value (NPV) does not apply to case-based surveillance systems however, can be applied to event-based surveillance systems where it refers to the number of weeks in which the surveillance system correctly categorised outbreaks which did not occur (true negative) in relation to the number of weeks with no detected outbreaks (true + false negatives). The assumption is that if the system does not detect an outbreak, there is indeed no outbreak.

Case ascertainment and case reporting

Under-ascertainment of cases of infectious diseases can be defined as the number of infected individuals who are not diagnosed and hence not identified by the healthcare systems. Underreporting of notifiable health outcomes occurs when patients affected by the notifiable condition visit physicians or healthcare facilities, but once diagnosed, are not properly reported by the healthcare provider to the local public health unit. Only a percentage of infected persons seeking medical care are diagnosed and eventually reported to public health authorities: The more cases are included when moving up in the pyramid (Figure 1.2a), the more sensitive the surveillance system is. Methods to calculate external completeness or sensitivity are presented in Sections 3.2 and 3.4.

Figure 1.2a. Levels of case detection in the surveillance process for communicable diseases



In a report prepared by an expert group from the European Commission in 2001 [7], sensitivity of the EU-wide surveillance system was defined as the combined sensitivity of the national surveillance systems and the international system. Thus, sensitivity can be defined on three levels: (a) the proportion of cases diagnosed at the local level and reported to the national system that meet the standard case definition; (b) the proportion of cases detected by the local level that truly occur in the population, regardless of whether the cases sought medical care or a laboratory diagnosis was attempted (can usually only be determined by special studies); or (c) the proportion of cases reported to the national system that were also reported to ECDC. In practice, the sensitivity of the national surveillance systems will determine the sensitivity of the overall surveillance system. However, the sensitivity of national surveillance systems, i.e. (a+b), will vary widely from country to country for specific diseases. When knowledge on the differences in the sensitivity of national surveillance systems is of importance to the objectives of the EU-wide surveillance subnetwork, country-specific investigations with a clearly defined methodology need to be implemented in order to determine the sensitivity of the national surveillance systems and ensure the comparability of the country-specific data.

1.2.3 Timeliness

Timeliness reflects the speed between steps in a public health surveillance system [5]. The term 'reactivity' reflects the delay before public health actions are initiated.

1.2.4 Usefulness

Usefulness implies that surveillance results are used for public health action. Assessing usefulness consists in taking inventory of actions that have been taken in conjunction with the surveillance system. A public health surveillance system is useful if it contributes to the prevention and control of adverse health-related events, including an improved understanding of the public health implications of such events. A public health surveillance system can also be useful if it helps to determine that an adverse health-related event previously thought to be irrelevant is actually important. In addition, data from a surveillance system can be useful in contributing to performance measures, including health indicators that are used in needs assessments and accountability systems [5].

1.2.5 Representativeness

A public health surveillance system that is representative accurately describes the occurrence of a health-related event over time and its distribution in the population by place and person [5]. Knowledge on the representativeness of surveillance data on the national level is important for some of the proposed specific objectives of EU-wide surveillance. For example, if the coverage of a surveillance system is not national, the notification rate cannot be calculated using the country population as a denominator, which makes meaningful comparisons with other countries impossible.

Cases notified to a surveillance system may be derived unevenly from the population under surveillance and may therefore not be representative of the events in the population in general. Examples of situations that could affect

representativeness are uneven geographical coverage, prevalence of urban vs. rural reporting sites, minority populations not reached by the surveillance systems, and data sources not including all clinical settings (hospital vs. primary care settings).

Over time, the representativeness of surveillance data may change for a number of reasons, such as changes in legislation, surveillance infrastructure, clinical practice and reimbursement policies. A change in representativeness could lead to biased conclusions if trends in surveillance data are compared between countries or at the EU level. Stability in the representativeness of national disease surveillance needs to be monitored. Changes in the representativeness need to be assessed, and, if possible, quantitatively measured.

1.2.6 Other surveillance attributes

Surveillance attributes not covered above include simplicity, flexibility, acceptability, as well as the interrelated attributes stability, reliability and adequacy.

Simplicity

The simplicity of a public health surveillance system refers to both its structure and ease of operation. Surveillance systems should be as simple as possible while still meeting their objectives [5].

Flexibility

A flexible public health surveillance system can adapt to changing information needs or operating conditions with little additional time, personnel, or allocated funds. Flexible systems can accommodate, for example, new health-related events, changes in case definitions or technology, and variations in funding or reporting sources. In addition, systems that use standard data formats (e.g. in electronic data interchange) can be easily integrated with other systems and thus might be considered flexible [5].

Acceptability

Acceptability reflects the willingness of persons and organisations to participate in the surveillance system [5]. Acceptability is influenced substantially by the time and efforts required to complete and submit reports or perform other surveillance tasks. Acceptability is linked to completeness of report forms and timeliness of data reporting.

Stability, reliability and adequacy

Stability refers to the reliability (i.e. the ability to collect, manage, and provide data properly without failure) and availability (the ability to be operational when needed) of the public health surveillance system [4]. Adequacy refers to the ability of the surveillance system to address its objectives.

1.2.7 Development of data quality indicators

In the context of monitoring and evaluation, an indicator is a quantitative metric that provides information to monitor performance, measure achievement and determine accountability [11]. While attributes of a surveillance system (like completeness and validity) are used to measure the performance of a surveillance system, indicators are quantitative measures associated with the attributes.

Some examples of defined data quality indicators are listed in Table 1.2c. An approach to develop data quality indicators should address the following issues:

- Specify attributes that are important for data quality monitoring and/or the evaluation of surveillance systems.
- List and define indicators as measures of the quality attributes; develop targets based on the literature, ECDC minimum levels, and the Member States' 'average'; modify as required by the local situation.
- Create a checklist of data quality issues to be monitored; this should help users to select indicators and set targets.
- Identify if there is a need for prioritisation. (Are there any quality measures that should be prioritised?).
- Consider using the data quality indicators as a part of the data quality monitoring report.

Table 1.2c. Examples of data quality indicators

Attribute	Proposed indicators
Completeness	• Completeness of data reported: percentage of cases recorded in a database with no missing required information (by disease, region, surveillance unit) • Percentage of missing information by required field (disease specific)
Validity	• Proportion of cases complying with case definition • Proportion of coding errors within a dataset • Proportion of values that comply with a gold standard or reference value (to be defined for specific diseases)

Attribute	Proposed indicators
Sensitivity, positive predictive value	• Sensitivity of outbreak detection: number of outbreaks detected divided by the number of occurring outbreaks. Is the surveillance system sufficiently sensitive to be useful in preventing or controlling the health problem? • If comparison with an external source: proportion of values that were not captured by the 'official' notification system
Timeliness and reactivity	• Proportion of surveillance units that submitted surveillance reports (weekly, monthly, annually) according to a predefined schedule • Proportion of outbreak notified within 48 hours of detection • Proportion of suspected outbreaks that were verified within 48 hours of notification • Average time interval between date of onset and date of notification by general practitioners/hospital (by disease, region, surveillance unit) • Average time interval between date of outbreak notification and date of first investigation (by disease, region, surveillance unit)
Representativeness	Population covered by the surveillance system (based on surveillance objectives), divided by the total population for a defined geographic area.
Usefulness (evaluation only)	Rating of the usefulness of the surveillance system. Indicators of usefulness can be described as actions taken as a result of surveillance outputs.
Simplicity (evaluation only)	• Rating of the simplicity of the surveillance system by implementers and users of the system • Amount and type of data necessary to establish that the health-related event occurred (a case of communicable disease, an outbreak) • Amount and type of other additional data collected on cases • Number of organisations involved in receiving case reports and outbreak reports from a surveillance unit • Amount of follow-up that is necessary to update data on the case • Method of managing the data, including time spent on transferring, entering, editing, storing, and backing up data • Time spent on system maintenance
Flexibility (evaluation only)	• Rating of the ability of the surveillance system to adapt to changing needs, as perceived by the national health managers and evaluators • One practical indicator of flexibility would a review of how the system coped in real-life situations that required its flexibility (e.g. capacity of the system to cope with the emergence of a new strain, change of disease severity, new legislation, availability of new technology, inclusion of additional diseases or outcomes, changes in case definitions, etc.)
Acceptability (evaluation only)	Acceptability of the surveillance system as rated by the data providers and implementers. Quantitative measures of acceptability can include: • subject or agency participation rate (if high, how quickly it was achieved?); • interview completion rates and question refusal rates (if the system involves interviews); and • physician, laboratory, or hospital/facility reporting rate.

1.3 Description of a surveillance system

The first step when monitoring data quality or evaluating surveillance systems is the description of the surveillance system.

During an evaluation, the people tasked with the description of the system usually:

- collect, read and summarise all available documentation;
- validate their summary with systems operators and collect additional information through interviews and by recording daily activities;
- identify the components of the systems that need to be evaluated;
- have an idea of how the various attributes can be derived, based on the characteristics of the system; and
- pay additional attention to the description of components which are critical for system performance.

The description of the surveillance system is an essential step of the evaluation process because a number of performance issues can already be detected at this stage.

Some key elements are listed in Figure 1.3a. Below we also provide definitions of key terms, offer practical guidance on how to describe these elements, and give examples from TESSy. Although the description of elements has a clear European perspective, it can also be applied to national and/or sub-national levels.

The below list is by no means exhaustive and should be revised at a later point in time in order to incorporate the findings of other ECDC-coordinated projects related to the development of descriptors and standards for surveillance. In a first step, ECDC wants to list and categorise the descriptors of surveillance systems and group them in different categories, e.g. legal, structural, and procedural (including IT system), which will clarify the different functions in a surveillance system. As a second step, standards for surveillance will be established, which

may include methods for the analysis of surveillance data, a compilation of indicators, targets associated with indicators, and thresholds for action.

Figure 1.3a. Key elements for the description of a surveillance system

- Surveillance objectives
- List of diseases under surveillance and case definitions
- Data sources, data flow
- Surveillance networks
- Population under surveillance
- Geographic coverage (regional, national)
- Type of surveillance:
  - Passive vs. active
  - Compulsory vs. voluntary
  - Comprehensive vs. sentinel
- Specification of the information to be reported
  - Case-based and aggregated
  - Variable specification
  - Frequency of data collection
- Reporting format
  - Paper-based
  - Electronic
- Data entry
  - Web-based
  - Interface-mediated data entry
  - Open-source software
- Database architecture
  - Centralised
  - Server selection

1.3.1 Surveillance objectives

As specified in the CDC guidelines, 'the purpose of the system indicates why the system exists, whereas its objectives relate to how the data are used for public health action [...] The purpose and objectives of the system, including the planned uses of its data, establish a frame of reference for evaluating specific components' [5].

Surveillance objectives of a surveillance system should be stated in a dedicated document related to the development of the surveillance system. In some countries, these objectives can be part of a country's legislative and regulatory provisions on communicable disease surveillance, prevention and control.

Examples of EU surveillance objectives

ECDC and its competent bodies for surveillance have developed a set of objectives for surveillance at the EU level. Objectives were formulated with respect to general and disease-specific surveillance and have been agreed upon by all Member States. These objectives guide the discussion on disease-specific surveillance networks and their further development in the EU.

As described in the ECDC long-term surveillance strategy 2008–2013 [2], the main objectives for EU surveillance are to:

- monitor trends in communicable diseases over time in order to assess the present situation and compare communicable disease trends across Member States (including healthcare-associated infections and antimicrobial resistance) in order to respond to disease outbreaks above certain warning thresholds and to facilitate appropriate evidence-based action;
- detect and monitor any multinational infectious disease outbreaks with respect to source, time, population and place in order to provide a rationale for public health action;
- contribute to the evaluation and monitoring of prevention and control programmes targeted at infectious disease surveillance in order to provide evidence for recommendations to strengthen and improve these programmes at the national and European level;
- identify population groups at risk and in need for targeted prevention measures;
- contribute to the assessment of the burden of communicable diseases on the population, using such data as disease prevalence, complications, hospitalisation, and mortality; and
- generate hypotheses on (new) disease sources, modes of transmission and groups most at risk, and identify needs for research and development and for pilot projects.

1.3.2 List of diseases under surveillance and case definitions

Most administrative areas (subnational, national, international) maintain – through their respective public health institutions – a list of diseases under surveillance. These lists represent the foundation for the reporting of communicable diseases. In most EU countries, the list of diseases under surveillance is endorsed by a regulatory process and part of public health legislation.

When preparing for the evaluation of a public health surveillance system, all diseases covered by the system should be listed. In most EU countries, surveillance of communicable diseases is not centralised in one system, but handled by several surveillance systems. Consequently, it is important to determine whether the diseases included in the surveillance system are part of an 'official' list established by an authorised public health institution and have to be reported in accordance with legal or regulatory processes.

Example of the list of diseases to be reported at the EU level

The list of communicable diseases which have to be reported at the EU level includes 52 diseases and conditions grouped by categories [13,14]:

- Respiratory tract infections: influenza (seasonal and avian), legionellosis, tuberculosis.
- STI, including HIV infection and blood-borne infections: chlamydia infection, gonorrhoea infection, hepatitis B, hepatitis C, HIV/AIDS and syphilis.
- Food- and waterborne diseases and zoonoses: anthrax, botulism, brucellosis, campylobacteriosis, cholera, cryptosporidiosis, echinococcosis, VTEC infection, giardiasis, hepatitis A, leptospirosis, listeriosis, salmonellosis, shigellosis, toxoplasmosis, trichinellosis, tularaemia, typhoid/paratyphoid fever, variant Creutzfeldt-Jakob disease and yersiniosis.
- Emerging and vector-borne diseases: malaria, plague, Q-fever, tick-borne encephalitis, SARS, smallpox, viral haemorrhagic fevers (including CCHF and chikungunya), West-Nile and yellow fevers.
- Vaccine-preventable diseases: diphtheria, invasive *Haemophilus influenzae* disease, invasive pneumococcal disease, measles, invasive meningococcal disease, mumps, pertussis, poliomyelitis, rabies, rubella and tetanus.
- Antimicrobial resistance and healthcare-associated infections

Case definitions

The criteria to be considered for the development of case definitions for surveillance purposes are clinical criteria, laboratory criteria, and epidemiological link. For each disease, the status of a case should also be defined as follows:

- **Possible:** classification of cases is based on clinical criteria. Diseases for which cases can be reported as possible are vCJD, VTEC, influenza, polio, SARS, smallpox, tuberculosis.
- **Probable:** classification of cases is based on clinical criteria, epidemiological link, and presumptive laboratory criteria. Diseases for which cases can be reported as probable include anthrax, botulism, brucellosis, hepatitis A, Legionnaires' disease, listeriosis, measles, meningococcal infection, mumps, pertussis, plague, rabies, rubella, syphilis, tetanus, VHF, WNV, yellow fever.
- **Confirmed:** classification of cases is based on laboratory confirmation. All diseases should be reported as confirmed after completion of diagnosis.

In the context of an outbreak, time and place should be mentioned. For outbreak investigation, a typical case definition would include:

- clinical criteria and laboratory findings to characterise the disease;
- a clear time period within which we count cases;
- a precise identification (personal characteristics) of the population from which we count cases and its location.

Case definitions for surveillance purposes may not necessarily apply in the context of outbreak investigation (where place and time of occurrence pays a role). Only case definitions for surveillance purposes will be included in the scope of this handbook.

Standard case definitions [15] will make it possible to compare disease occurrence between populations, in different locations, and at different times. Using case definitions removes potential bias related to different criteria being used across the system (e.g. exposure criteria for inclusion/exclusion of a case).

The last major revision of case definitions for diseases relevant to ECDC was in Commission Decision of 8 August 2012 amending Decision 2002/253/EC [16].

1.3.3 Data sources and data flow

A data source for a disease surveillance system can be defined as a place where the initial information on the disease to be reported is collected and from which data are sent to public health authorities. For example, healthcare professionals can act as 'data providers' and gather and report information on communicable diseases to the designated public health authorities.

Data sources

The data sources described in this handbook refer to indicator-based surveillance processes, which include notifications and case reporting [2]. Laboratories, general practitioners, and hospitals are the most common sources of information for case-based communicable disease reporting systems (Table 1.3a). Other sources (healthcare based or not) can also be included in a public health reporting system.

Further sources and types of data are described in ECDC's FEM Wiki [17].

Table 1.3a. Data sources for indicator-based surveillance

Data source	How to describe
Laboratory	A microbiological diagnosis is part of the confirmation process for most of the communicable diseases reported in surveillance systems. Therefore, laboratories for microbiological analysis represent a major source of data. Laboratory diagnosis can be reported to public health authorities in two different ways: • Direct reporting to public health authorities. Laboratories participating in the surveillance system might be linked through a network dedicated to surveillance activities. • Indirect reporting through general practitioners. The laboratory diagnosis is attached to the report provided by general practitioners. Several types of laboratories contribute to public health and clinical microbiology functions. Therefore, laboratory data sources can be classified by: • area covered: local, subnational, national or international; • administrative status: public or private; • population covered: human, animal, environmental; and • point-of-care tests (also known as bedside or near-patient tests).
General practitioners	General practitioners can participate in a surveillance system in different ways: • Notification of cases of infectious diseases through mandatory notification • Reporting of cases of infectious diseases on a voluntary basis General practitioners participating in a case-reporting system may be based in a private/public clinic/hospital or a community health service.
Hospitals	Hospitals are an obvious data source for healthcare-associated infections and/or antimicrobial resistance. Other infections, which are not necessarily hospital acquired, but rather community-based, can also be reported through hospital networks, for example pneumococcal infections and severe respiratory infections such as influenza.
Other data sources	Healthcare based: • Syndromic surveillance data • Electronic health record Not healthcare based: • Mortality monitoring • Behavioural surveillance • Veterinary • Environmental • Food safety • Drug post-licencing monitoring

Note: Data sources for public health information systems of countries participating in European disease surveillance are listed by disease in the ECDC Annual Epidemiological Report [13].

Data flow

A flowchart for a generic case reporting system of infectious diseases is described in Figure 1.3b. The following elements should be considered:

- Data providers or data sources as described in the previous section
- Processes for clinical diagnosis, case confirmation, and gathering of additional information

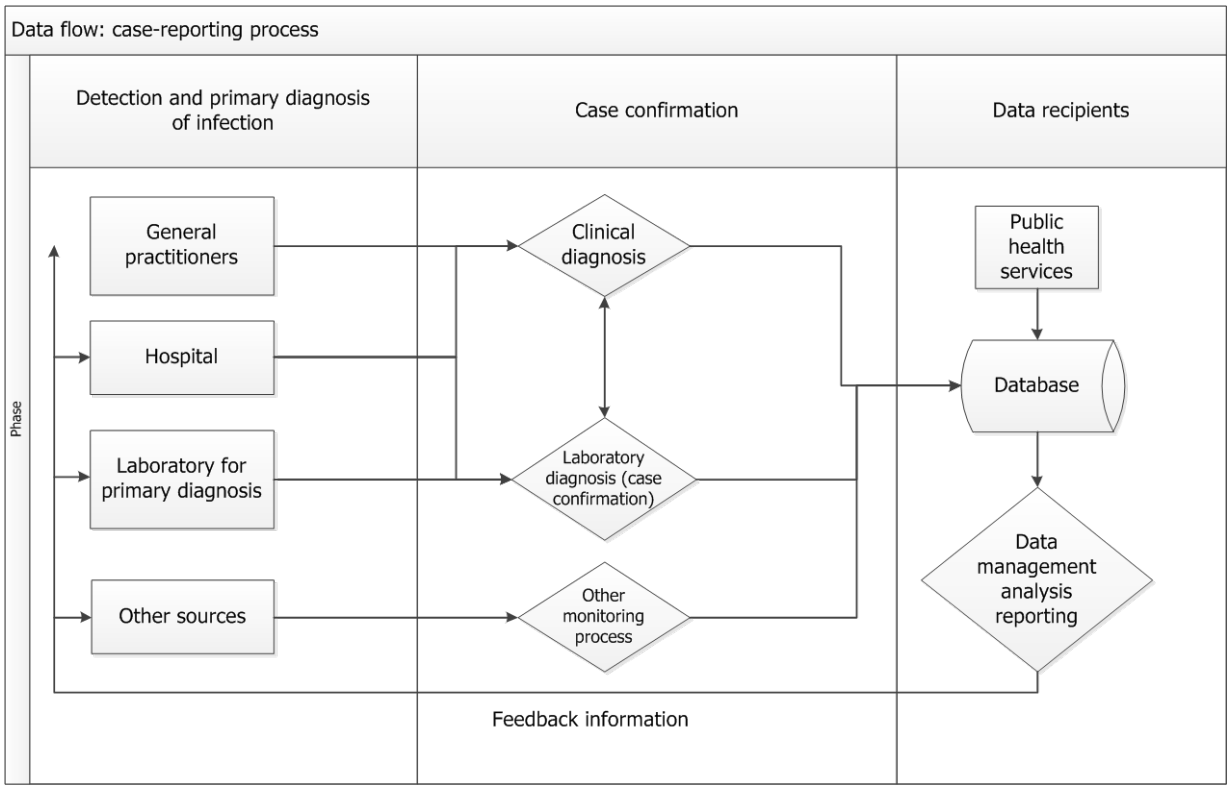
- Public health institutions (data recipients) that provide feedback information to participants of the case reporting process, public health professionals, and the general public

It is important to describe the data flow to further characterise channels of communication between the various stakeholders that participate in the surveillance of a disease or a group of diseases. The data flow of a reporting system will be essential for evaluating the complexity of a system and for identifying steps where information quality could be affected. The roles and responsibilities of the data providers and the various stakeholders in the flow of surveillance data through the reporting system should be identified.

As recommended in the CDC guidelines, the following steps should be taken [5]:

- Cite any legal authority for the data collection.
- Describe where in the organisation the system resides, including the context (e.g. the political, administrative, geographic, or social climate) in which the system evaluation will be done.
- Describe the level of integration with other systems, if applicable.
- Draw a flowchart of the system.

Figure 1.3b. Simplified flowchart of a generic surveillance system (indicator-based)



1.3.4 Surveillance networks

Networks for the surveillance of infectious diseases are formed by all partners participating in the reporting process of the diseases under surveillance. The establishment of a network is one of the first steps in the development of a surveillance system. In most cases, public health networks for the surveillance of infectious diseases are established by public health authorities to address public health objectives for disease surveillance. The role and scope of a public health network for the surveillance of infectious diseases can be extended to other public health functions like prevention, monitoring and evaluation, preparedness, early warning and response, programme management, and evidence-based guidance. Surveillance of communicable diseases requires concerted efforts and collaboration between stakeholders and partners in and between countries.

A network participating in the surveillance of infectious diseases can be described within the framework of data sources and data flow (see 2.2.3). Surveillance network functions can be described by established communication channels, feedback loops, training activities, support actions to data providers, and (financial) incentives to report data to the system.

ECDC surveillance networks

In the EU, countries collaborate on surveillance and response activities. European surveillance is conducted through disease-specific networks, where nominated contact persons in the Member States report data to ECDC within an agreed time frame. Reports are sent back to Member States for comments before publication. See also: http://www.ecdc.europa.eu/en/activities/surveillance/european_surveillance_networks/Pages/european_surveillance_networks.aspx.

Collaboration between international public health agencies is necessary to maintain early warning and response functions. The current network for disease surveillance was originally established by Decision No 2119/98/EC of the European Parliament and of the Council of 24 September 1998 setting up a network for the epidemiological surveillance and control of communicable diseases in the Community. On 22 October 2013, this Decision was replaced by Decision 1082 of the European Parliament and of the Council on serious cross-borders threats to health [12].

1.3.5 Populations under surveillance

The population under surveillance can be defined as the general population or targeted groups. The targets can be based on risk groups (e.g. injecting drug users for HIV surveillance), specific age categories (e.g. children under five years of age for *Haemophilus influenzae* surveillance), or other determinants. The population under surveillance should be determined according to the objectives of the surveillance system and should take into account feasibility.

1.3.6 Geographic coverage

The geographic coverage represents the geographic unit (municipality, region, country or any other pre-defined geographic area) where disease surveillance is conducted. Some surveillance systems are run at the local level for targeted exposed populations (for example surveillance of gastroenteritis in the vicinity of a sewage treatment plant) while other systems are operated at a larger scale, for example the mandatory notification of infectious diseases to public health authorities. The geographical coverage of disease surveillance is linked to the concept of representativeness of the surveillance system.

1.3.7 Types of surveillance

The following types of disease surveillance are currently used in the EU.

Passive and active surveillance

According to the TESSy metadataset, passive surveillance 'relies on the physicians, laboratory or hospital staff or other relevant sources to take the initiative to report data to the health department'. A key data quality concern of passive surveillance systems is external completeness.

An active surveillance system, according to TESSy, is a system that is 'based on the public health officials' initiative to contact the physicians, laboratory or hospital staff or other relevant sources to report data'. Active surveillance typically generates high-quality data (high levels of completeness, validity and timeliness) through a direct link with data providers. Active surveillance relies on a prompt response from public health professionals, who will usually monitor replies and can respond to low response rates in order to improve external completeness. Completeness for key variables, and thus better representativeness, will be achieved if distribution and composition of data providers are carefully chosen. Data comparability will be improved if data are collected through a common surveillance protocol. The main disadvantages of active surveillance are the substantial resources needed and the subsequent need to focus on high-priority diseases and questions that justify the additional expense.

Compulsory versus voluntary

Surveillance systems rely on processes and individuals. Member States use different approaches to elicit surveillance data from their sources. Some systems make data submission mandatory or quasi mandatory (either by law, professional edict, policy or guidance), whereas others rely on a voluntary approach. A legal or professional obligation to report may lead to greater external completeness and may increase timeliness but is unlikely to have an impact on representativeness or validity.

Comprehensive versus sentinel

Comprehensive surveillance systems include reports of cases of infectious diseases that occur within the whole population of the geographical area covered by the surveillance system (national, regional). Sentinel systems rely on notifications from a selected group of physicians, hospitals, laboratories, and other institutions. In the case of sentinel systems, representativeness is often established as part of the system-design process. The choice between sentinel and comprehensive systems is based on surveillance objectives, characteristics of the disease, and the availability of resources.

If surveillance operates through sentinel systems, the characteristics of the sentinels should be described:

- Number of physicians, laboratories, and hospitals participating in the system
- Sampling methods for the selection of the sentinels
- Denominators for the population under surveillance

1.3.8 Specification of the information to be reported

Datasets for surveillance systems are usually in a case-based or aggregate format. The inclusion of variables is ideally based on clearly defined surveillance objectives.

Case-based and aggregated reporting

A case-based dataset format includes information on individual records related to a disease/health event. An aggregate format for a dataset includes information related to a group of people classified under the same category for a disease or health-related event, for example the number of people reported by age and/or gender. Aggregated reporting systems are selected to reduce costs and minimise risks associated with the unnecessary exposure of personal data. However, from a data quality perspective, it can be difficult to assess the quality of aggregated data. Validation rules cannot be applied by the system to the same extent as when evaluating aggregated data, and data quality monitoring is limited to measures that apply to batches of aggregated data. Because of the limitations on data quality monitoring in systems that use only aggregated data, system owners should consider implementing whole-system evaluations on a regular basis.

Specification of variables

Variables and related coding systems for a disease/health-related event should be specified for all case-reporting disease surveillance systems. For each disease, the specification of variables should be based on surveillance objectives.

TESSy variables

For example, the metadataset of TESSy includes a common set of variables which is reported for all diseases; TESSy also offers an enhanced set of variables for selected diseases. The common set of variables includes age, gender, outcome, dates (onset, diagnosis, notification, used for statistics), reporting country, classification (clinical criteria, epi-linked, lab result). The enhanced set of variables may include the following categories of variables:

- Dates (hospitalisation, treatment, death)
- Geographic/origin: place of residence/notification, country of birth/nationality
- Travel-related: imported/probable country of infection.
- Transmission (human immunodeficiency virus, sexual transmitted infections, food- and waterborne disease)
- Laboratory: pathogen, test method, serogroup, AMR
- Diagnosis: clinical presentation, previous diagnosis
- Treatment/vaccination
- Underlying conditions, complications

Frequency of data collection and data reporting

The frequency of data collection and data reporting should be related to the objectives of the surveillance system and the public health importance of the disease in question. For example, diseases which have to be reported in accordance with the legally binding International Health Regulations, e.g. SARS, should be notified immediately to all surveillance levels. Diseases requiring outbreak investigation and rapid public health intervention (e.g. salmonellosis) should be notified immediately at the local level, and within 48 hours at the national level.

The frequency of data collection and the frequency of data reporting should be described for each disease.

TESSy data collection

At ECDC, for example, data are collected annually for most of the 52 diseases under surveillance. Selected zoonoses are reported quarterly. Measles and rubella data are reported monthly. Influenza is reported on a weekly basis, and some other diseases (e.g. travel-associated Legionnaires' disease, diphtheria) are reported daily. Some factors can influence the frequency and way of reporting at the EU level:

- International/cross-border outbreaks have to be notified immediately, using event-based surveillance tools (no reporting through TESSy).
- Diseases for which the main objective is the monitoring of trends over time and the description of risk factors can be reported to TESSy annually.
- Diseases (e.g. influenza) that spread rapidly and need to be characterised should be reported more frequently to inform preparedness, prevention and control measures.
- The frequency of reporting to the EU level cannot be higher than the frequency of reporting to the national level.

1.3.9 Reporting format

Reporting format, as treated in this document, refers to paper-based and electronic record keeping.

Paper-based systems

Unless data are entered into an information system at some point, it is very unlikely that any rigorous data quality control system can be established. Quality monitoring will prove difficult as it will require a lot more effort (compared to an electronic system) if based on paper records alone.

The key advantage of paper is that it requires no specialist technical knowledge and it entirely circumvents the difficulties of interfacing between information systems. On the other hand, a paper form cannot ensure data quality at the point of data recording, e.g. by limiting data entry to a pre-populated list of terms.

Data that are reported on paper will have to undergo a transcription step. Although there are computer-based technologies that can support this now (e.g. form scanning and optical character recognition software), transcription always adds the risk of introducing additional errors.

Collections that rely on paper-based solutions should consider implementing completeness and validity checks across as much of the dataset as possible. It should also be kept in mind that paper systems cannot enforce completion of mandatory fields or other validation checks.

Electronic systems

An electronic case reporting system is a computerised system hosting medical records on diseases/health events. Additional manual steps can be included for human-moderated quality control. In practice, this advantage is often lost, for example when manual quality control takes place outside the direct line management of the system owner. To maximise benefits while minimising risks, surveillance systems should aim to become wholly automated end-to-end. For each automated step, measures which monitor data quality should be introduced to ensure that the automation process does not introduce errors.

1.3.10 Data entry

There is a wide range of data entry methods available; as a rule, all data entry systems should validate user input before processing it.

Web-based data entry

This frequently used method requires manual online data input: information on a case or series of cases is entered into a data entry mask on a web page. Advantages are that this technology is common, cheap, and the introduction of a manual step reduces costs compared to a system that relies on integration with other systems. However, web-based data entry also introduces the risk that data are not regularly entered.

Web tools can, and should, perform data-validation checks.

Interface mediated data entry

If data are initially collected or recorded in an external system, e.g. a laboratory information management system (LIMS), it may be possible to wholly automate data export by developing an interface. Using international standards such as Health Level Seven International (HL7) and SNOMED Clinical Term will support increased interoperability in future. Currently, however, many systems that support this type of data transfer rely on custom interfaces. The data quality risks inherent in automated data transfer can be identified by a whole-system evaluation. For instance, problems with the identification of cases are caused by different interpretations of the case definition and the communication between the reporting system and the external system

Usually, an evaluation will identify and remedy problems caused by the user interface. Additional reports on data quality should be produced to identify potential problems, for example if a sending system was modified.

Finally, validation rules on data entry can already be introduced at the level of the sending system (e.g. at the LIMS or when a record is entered online), thus ruling out data entry errors and ensuring that the data are compatible with all applications, including surveillance.

Open-source software

The great advantage of open-source software is that it is free. The downside is that the software is not supported in an enforceable way, i.e. by a legally binding contract. From a business continuity perspective many organisations feel that open-source solutions are not sufficiently robust.

1.3.11 System architecture

Determining which system architecture to use and how to design the system is an essential question when setting up a disease surveillance system.

Centralised versus distributed systems

The architecture of a system can have repercussions on data quality. In theory, neither a centralised approach nor a decentralised/distributed approach to database design need to have a negative effect on data quality. Centralised systems may prove to be cumbersome, so a database designer must ensure that systems are able to run all necessary checks for data quality monitoring. Conversely, the key problem experienced by owners of distributed data systems is one of data validity – if data are changed in one part of a distributed database, rules must be in place to propagate and audit that change elsewhere in the database.

Server selection

The choice of server hardware and database solution can have a marked effect on data quality. Server hardware varies in levels of stability, maturity and speed, and the choice of server may be predicated on the requirement for index-driven as opposed to direct-disk-access querying. The choice of database software can affect data quality because of variations in data security, data resilience, data auditing and data indexing. To mitigate risks caused by the choice in soft- and hardware, the validity of data needs to be thoroughly monitored.

1.4 References

1. European Centre for Disease Prevention and Control. Survey report: mapping current practices for monitoring and evaluating quality of surveillance data in EU/EEA Member States. Stockholm: ECDC; 2010.
2. European Centre for Disease Prevention and Control. Surveillance of communicable diseases in the European Union. A long-term strategy: 2008–2013. Stockholm: ECDC; 2008. Available from: http://www.ecdc.europa.eu/en/aboutus/Key%20Documents/08-13_KD_Surveillance_of_CD.pdf
3. European Centre for Disease Prevention and Control, WHO Regional Office for Europe. HIV/AIDS surveillance in Europe 2010. Stockholm: ECDC; 2011. Available from: http://www.ecdc.europa.eu/en/publications/Publications/111129_SUR_Annual_HIV_Report.pdf
4. World Health Organization. Communicable disease surveillance and response systems. A guide to monitoring and evaluation. Geneva: WHO; 2006.
5. Centre for Disease Control and Prevention. Updated guidelines for evaluating public health surveillance systems. Recommendations from the Guidelines Working Group. MMWR July 27, 2001/50 (RR13);1-35
6. Doyle TJ, Glynn MK, Groseclose SL. Completeness of notifiable infectious disease reporting in the United States: an analytical literature review. Am J Epidemiol. 2002 May 1;155(9):866-74.
7. Ruutu P, Breuer T, Desenclos JC, Fisher I, Giesecke J, Gill N, et al. Framework and recommendations for evaluating surveillance systems within the community network for communicable diseases. Brussels: Directorate General Health & Consumer Protection; 2001.
8. Rajaraman P, Samet JM. Quality control and good epidemiological practice. In: Ahrens W, Pigeot I, editors. Handbook of epidemiology. 2nd ed. Heidelberg: Springer; 2014.
9. Buehler JW, Hopkins RS, Overhage JM, Sosin DM, Tong V; CDC Working Group. Framework for evaluating public health surveillance systems for early detection of outbreaks: recommendations from the CDC Working Group. MMWR Recomm Rep. 2004 May 7;53(RR-5):1-11.
10. Devaux I. Mapping of the current practices for monitoring and evaluating data quality in surveillance system for communicable diseases in EU/EEA countries. Paper presented at: ESCAIDE, Lisbon 2010.
11. UNAIDS. An introduction to indicators. Monitoring and evaluation fundamentals. Geneva: UNAIDS; 2010.
12. Commission Decision of 22 October 2013 on serious cross-border threats to health and repealing Decision No 2119/98/EC of the European Parliament and of the Council. Available from: http://ec.europa.eu/health/preparedness_response/docs/decision_serious_crossborder_threats_22102013_en.pdf
13. Commission Decision of 28 April 2008 amending Decision 2002/253/EC laying down case definitions for reporting communicable diseases to the Community network under Decision No 2119/98/EC of the European Parliament and of the Council. Available from: http://ec.europa.eu/health/ph_threats/com/docs/1589_2008_en.pdf
14. Commission Decision of 22 December 1999 on the communicable diseases to be progressively covered by the Community network under Decision No 2119/98/EC of the European Parliament and of the Council. Available from: http://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri = CELEX:02000D0096-20120905&from = EN
15. FEM Wiki, Field Epidemiology Manual. Defining a case. [internet]. 2010 [cited 2014 Aug 25]. Available from: https://wiki.ecdc.europa.eu/fem/w/wiki/defining-a-case.aspx
16. Commission Decision of 8 August 2012 amending Decision 2002/253/EC laying down case definitions for reporting communicable diseases to the Community network under Decision No 2119/98/EC of the European Parliament and of the Council. Available from: http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri = OJ:L:2012:262:0001:0057:EN:PDFht
17. FEM Wiki, Field Epidemiology Manual. Source and types of data. [internet]. 2010 [cited 2014 Aug 25]. Available from: https://wiki.ecdc.europa.eu/fem/w/wiki/sources-and-types-of-data.aspx

2 Monitoring data quality

This handbook only covers data quality issues related to the case reporting process (and associated underreporting).

2.1 Conceptual framework

This section covers several aspects of underreporting and data quality monitoring.

2.1.1 Underreporting and under-ascertainment

Underreporting and under-ascertainment can have a significant impact on data quality. Various degrees of underreporting and under-ascertainment can be observed during the data quality monitoring process. Underreporting can be detected at the healthcare and public health system levels, while assessment of under-ascertainment would require additional studies at the population level.

Possible reasons for underreporting in the healthcare and public health system

There are many situations that may lead to underreporting which are important to consider when trying to improve data quality. Underreporting, for example, can occur if a patient is not visibly ill when visiting the doctor (e.g. patients with chronic but notifiable conditions). In addition, many patients will not be reported at their first visit but when laboratory results are available or an epidemiological link is established in the context of a possible outbreak.

Underreporting may also occur if a patient has more than one reportable condition and only one is reported. Another example could be a patient presenting in a serious clinical condition and the physician focuses on the patient, simply forgetting to notify the case.

Cases that are not reported can be classified as follows:

- **Non-visible cases**: the person is not in the public health surveillance database. This may be caused by failure to register the case with the healthcare provider, or an attempt to register was made but the communication or the software system failed. In addition, some cases may be non-visible to the surveillance system because:
  - the clinician, hospital epidemiologist, or laboratorian (healthcare provider or reporter) failed to report the notifiable health outcome; or
  - the surveillance system (e.g. communication system or software) did not fully support the actions initiated by the healthcare provider (or reporter) so he or she could not file a formal report.
- **Visible cases**: the registered case is present and registered in the public health surveillance database, but has been misdiagnosed, misclassified, or miscoded.

As mentioned in Chapter 1, validity issues need to be categorised:

- Misdiagnosis: health outcomes with unspecified symptoms in the absence of laboratory confirmation (e.g. chlamydia)
- Miscoding: health outcomes which were not reported because the data provider forgot to do so.
- Misclassification: health outcomes reported with inappropriate case definition category (case reported as confirmed if in fact probable). If, for example, tick-borne encephalitis is reported in an area where the vector is not present, it is highly unlikely that the disease was actually reported there.

When assessing underreporting, quality control measures should focus on data completeness and validity.

Possible reasons for under-ascertainment of cases

Under-ascertainment of notifiable health outcomes occurs when patients do not visit healthcare services. However, failure to seek healthcare assistance is only one of the causes of under-ascertainment. Other reasons can be unequal geographical distribution of healthcare services; unavailability of diagnostic tools in certain areas; different primary care ascertainment attitude due to different experience, size, location, and socio-economic characteristics of the served population.

These patients are not visible to the surveillance system. The following reasons for under-ascertainment are common:

- Routine surveillance does not capture marginalised high-risk groups (e.g. commercial sex workers, intravenous drug users, men having sex with men, the Roma population).
- Asymptomatic carrier states, especially for diseases that have a wide clinical spectrum, may be missed more easily than severe cases.
- Unawareness of the disease among physicians and the general population.

- Ill people may not seek medical care for various reasons, e.g. they cannot take time off work, cannot afford medical bills/medication, cannot access the hospital, or feel their illness does not warrant a trip to the doctor.

When anticipating or recognising under-ascertainment, quality control measures should focus on improving the sensitivity of the system.

Assessment of underreported cases in the context of public health surveillance

As described in the WHO guidelines [1], the core functions that characterise a case reporting system are detection, reporting, case registration, and confirmation. Other activities should also be considered in detail, including laboratory testing (if required), feedback, and data analysis and interpretation.

By considering in detail each type of activity attached to these functions, one can determine the factors that may create gaps in, or impediments to, data reporting. For example, regarding case detection and case reporting, some hard-to-reach populations may be at a greater risk of acquiring infectious diseases. Therefore, several demographic and behavioural characteristics might impact under-ascertainment and underreporting.

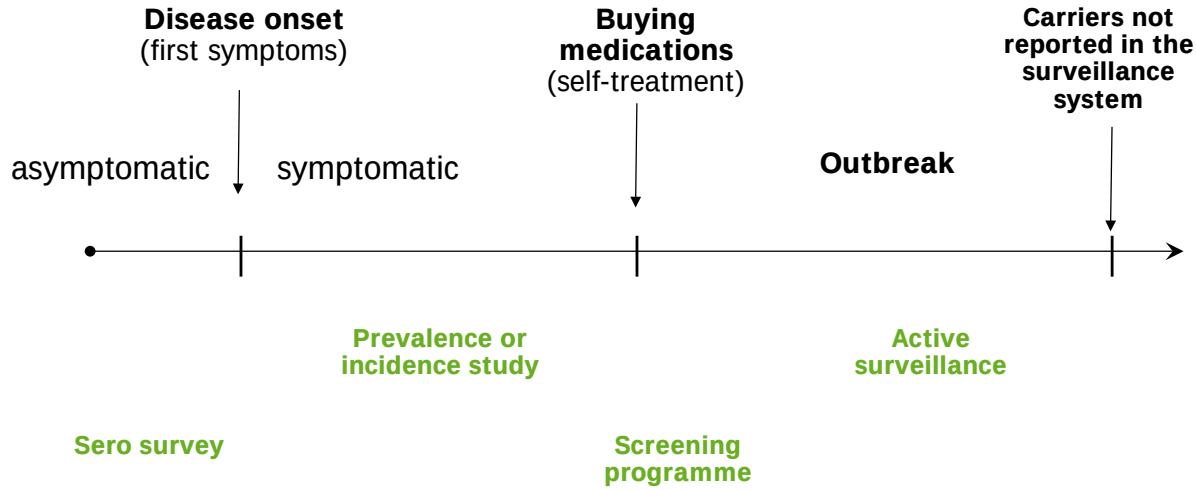
For example, the risk of infection for vaccine-preventable diseases may be higher in some populations because of low childhood vaccination coverage. Commercial sex workers are a population at risk for sexually transmitted diseases. Intravenous drug users are at risk for blood-borne diseases as some share and reuse needles. Reaching populations who rarely visit healthcare services is difficult because of the fear of stigmatisation and discrimination, economic reasons (unpaid sick leave), and a lack of education on healthcare.

Non-visible (underreported or under-ascertained) cases in public health surveillance are difficult to identify. However, it is possible to try to determine where the 'case loss' might happen by analysing the surveillance data flow (Figure 2.1a.). Unreported symptomatic cases may be visible to healthcare providers not covered by the surveillance system, cases may buy over-the-counter medications, or are only identified in the context of an outbreak investigation. However, specialised studies are able to estimate the true disease prevalence or incidence if necessary. Possible options include:

- Monitoring sales of specific drugs over time through electronic systems
- Active surveillance
- Identification of non-visible cases during an outbreak (through contact investigation of visible cases)
- Screening of risk groups or samples of the general population (e.g. pregnant women, refugees, prisoners, blood donors, screening before surgery).

The only way to find asymptomatic cases would be through specific sero-surveys (e.g. serological surveys). This could be relevant for diseases such as hepatitis B and C and HIV.

Figure 2.1a. Possible steps to identify non-visible cases who do not seek healthcare



Priority diseases to be assessed for underreporting

Some diseases tend to be consistently underestimated. It is important to also pay attention to diseases that present only a potential threat to public health, without neglecting those that pose a real threat. Criteria to identify priority diseases for which under-ascertainment can affect prevention and control should be developed.

Examples of underreported diseases include emerging infectious diseases, defined as diseases whose incidence rates have increased in the past 20 years and threaten to increase in the near future.

Emerging infectious diseases include:

- those caused by newly identified microorganisms or newly identified strains of known microorganisms (e.g. SARS, AIDS);
- new infections resulting from change or an evolution of an existing organism (e.g. influenza, *Clostridium difficile*);
- known infections that spread to a new geographic area or population (e.g. West Nile virus, malaria, viral meningitis)
- newly recognised infections in an area undergoing ecologic transformation (e.g. Lyme disease, tick-borne encephalitis)
- pre-existing and recognised infections re-emerging due to drug resistance or to a breakdown in public health (e.g. tuberculosis, NDM-1-producing Enterobacteriaceae)

2.1.2 Setting up a data quality monitoring system

According to Andres G. Lescano et al., it is 'important to assess data quality at each state of implementation [of a surveillance system] using a diverse mix of data sources and analytical methods. Careful, close monitoring of selected indicators is needed to evaluate whether systems are reaching their proposed goals at each stage' [2].

Introduction

Information systems can collect, collate, analyse, interpret and disseminate surveillance data in many ways. Although every system will be set up and developed in a unique set of cultural, political and economic conditions, there are learning points that can be captured and shared to help improve data quality in these systems. In addition, although all data quality attributes are important (the main data quality attributes are defined as completeness, validity and timeliness), depending on the purpose of the system, some attributes are likely to be more important than others. This chapter is intended to aid surveillance system owners and coordinators in identifying areas where systems may suffer in terms of data quality.

Developing a strategy for establishing a monitoring system

Public health services in Europe operate surveillance systems for reasons as diverse as supporting local public health action, early warning and emergency response, national and local policy, national and local commissioning, performance management, research, and reporting to international organisations such as ECDC.

A pragmatic monitoring system, i.e. the regular assessment of data internal to a system to ensure data quality is fit for purpose, should be in place for all surveillance systems. It is recommended that a pragmatic approach to selecting data quality measures is taken because, although it is very simple to run large, comprehensive, automated reports, it is easier to see important changes when viewing a concise report.

What is 'fitness for purpose'?

An information system is fit for purpose if it enables users to make an accurate assessment or prediction of the real-world situation it is trying to describe [3].

For instance, a laboratory surveillance system may be set up to model the number of cases of *Salmonella* spp. in a given country in order to identify outbreaks and put in place an action plan to help reduce the number of cases. An evaluation of that system will identify whether the used case definitions are useful and appropriate. It may also identify where assumptions have been made that influence elements of data quality such as completeness, e.g. the assumption that all cases of *Salmonella* are laboratory confirmed, or all laboratory-confirmed cases are reported. Monitoring of data quality is purely internal to the system and is intended to occur frequently or continuously in order to identify problems or track improvements in data quality as they happen.

What is 'the system'?

An information system comprises a series of processes or protocols involving the collection, processing and presentation of information which may or may not contain elements of technology [4]. In defining where data quality monitoring should take place, one must assess the system as a whole and determine what is necessary and practical to record and measure.

The point at which the value of the data quality exceeds the cost of data collection and data analysis associated with a data quality measure can be considered as an appropriate focus when monitoring of data quality. The cost of data quality monitoring decreases as automation increases. An ideal situation would be to seamlessly collect all data items associated with data quality, present them electronically in a surveillance system and automate data-quality output.

General steps for setting up a data quality monitoring system

The process below summarises a strategic approach for setting up a data quality monitoring report with a concise list of data quality measures. It corresponds with the planning phase described for the common approach to monitoring data quality in Chapter 1.

Step 1: Identify objectives for the system itself: document the system's fitness for purpose.

Step 2: Identify objectives for other systems: document fitness-for-purpose measures for other users of the data (e.g. ECDC).

Step 3: Identify the process modelled by the system, and choose measures to monitor the timeliness of data.

Step 4: Identify collected dataset and used system rules, and choose measures to monitor the completeness and validity of data.

Step 5: Use evaluation data to set up representativeness measures and inform the selection of additional monitoring measures.

Step 6: Review the chosen measures as part of every system evaluation.

To illustrate this process, we have modelled the British statutory notification system, a system which only collects information on clinically diagnosed (not laboratory confirmed) and suspected diseases.

Step 1: Identify and document the system's fitness for purpose

The system was set up to notify the English publish health system of cases of suspected or diagnosed diseases of public health significance. Notifications would enable the rapid implementation of public health actions to limit the spread of infections and limit further damage. The system is fit for purpose if reports are received by the public health organisation within three days of a patient being first clinically suspected or diagnosed as having a disease of public health significance.

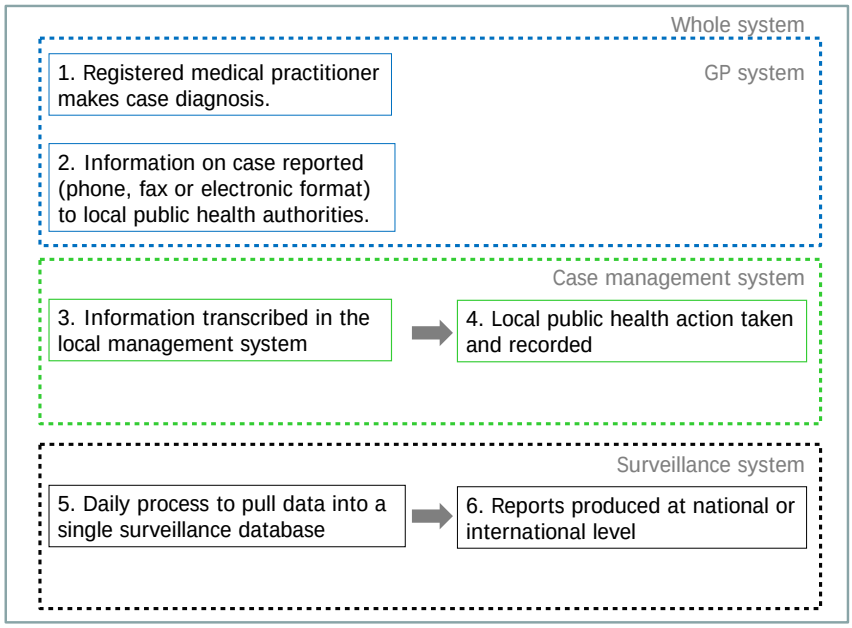
Step 2: Document fitness-for-purpose measures for other users of the data

The system is also used as a general surveillance system, and data are used to identify broad exceedance and monitor trends. The system is fit for purpose if complete, valid and representative data are available in a reasonable timeframe. ECDC requires that data reported by the system are complete, valid, representative and timely.

Step 3: Identify the process modelled by the system, and choose measures to monitor the timeliness of data

Identifying data quality issues around timeliness requires an overall understanding of the process the surveillance system is modelling. As part of a system evaluation, system owners can draw a process map to describe this process. A simplified example of the notification of clinically diagnosed/suspected disease is shown below (Figure 2.1b.).

Figure 2.1b. General practitioners, case management and surveillance systems



Each box represents an event in the process that is modelled by the system. The public health authorities are able to routinely monitor information collected in the surveillance system component (see bottom of figure).

To identify an appropriate timeliness measure for data quality monitoring one must establish whether the purpose of the system requires it. In this example, a system could be used for a number of reasons, and each reason might utilise a different measure as explored below.

In order to meet timeliness requirements, i.e. receiving a notification within three days of the first consultation, the monitoring system must measure the time elapsed between boxes 1 and 3. It may also be interesting to see how long it takes to implement public health action from the point the public health system is aware of the case (the time elapsed between boxes 3 and 4) or the point the health service becomes aware of a disease/outbreak (the time elapsed between boxes 1 and 4).

Typically, surveillance requirements focus on timeliness of reporting, and system time delay (i.e. the time elapsed between box 1 and box 6) should be measured.

It is acceptable to not take timeliness measures when regularly monitoring data, as long as timeliness is not an indicator of whether the system meets its objectives.

Step 4: Identify collected dataset and used system rules, and choose measures to monitor the completeness and validity of data

Data completeness is an assessment of how often a given field contains a value. Validity is a measure of whether the value in a given field is correct. Completeness and validity can be measured separately as a percentage to give an overall indication of how well a given field is populated within a database. More commonly, completeness and validity measures are combined to give an indication of the valid completeness of a given field.

The first step to selecting an appropriate completeness/validity measure is to document the complete dataset for the system, including derived fields. Once documented, the list of fields can be reviewed together with the users of the system to identify where a high level of completeness is most important. Reasons for measuring include the following:

- The system's purpose relies on a pre-defined completeness rate in order to meet its purpose (e.g. diagnosis code).
- The completeness of a field has a direct effect on the workload of the system users (e.g. a unique person identifier).
- The completeness of a field is likely to fluctuate over time.
- The field is required to support data analysis.

Validity can only be monitored if the system contains a set of rules which determine the accepted and valid values for all fields. It is possible to manually monitor data validity (this is a common approach to ensure validity in systems with a low number of line listings of cases), but larger systems which process a high number of datasets rely on automated processes to ensure data validity.

In order to support the combination of field validity and completeness measures (e.g. pregnancy status), validity measures are often used to monitor system quality. If automated extraction and loading techniques are used, files can be rejected if data validity criteria are not met (e.g. if a record is not associated with a case identifier). Measuring how many files were rejected can give an indication of the operations that need to be integrated into the data management process. The level of validity (and whether validity is adequately assessed) can only be determined after an evaluation of the rules used to assess field or file validity.

Below (Table 2.1b.) is a sample dataset with validity rules from the United Kingdom.

Table 2.1b. Validity rules for data collected from the registered medical practitioner (RMPs) in the United Kingdom

Health protection (notification) regulations 2010: notification to the proper officer of the local authority	
Data item name	**Validity rules**
Registered medical practitioner (RMPs) reporting the disease	
Name	From a list of RMPs
Address	
Post code	Must be a valid post code
Contact number	Valid format for a UK land line or mobile number
Date of notification	Must be a valid date format in the past
Notifiable disease	
Disease, infection or contamination	From a valid list of SNOMED CT disease codes

Health protection (notification) regulations 2010: notification to the proper officer of the local authority	
Data item name	**Validity rules**
Date of onset of symptoms	Must be a valid date format in the past
Date of diagnosis	Must be a valid date format in the past
Date of death (if patient died)	Must be a valid date format in the past
Index case details	
First name	
Surname	
Gender (M/F)	M/F/unknown
Date of birth	Must be a valid date format in the past
Ethnicity	From a valid list of ethnicity codes
NHS number	A valid NHS number with correct check digit
Home address	
Post code	Must be a valid post code if populated
Current residence if not home address	
Post code	Must be a valid post code if populated
Contact number	Valid format for a UK landline or mobile number
Occupation (if relevant)	
Work/education address (if relevant)	
Post code	Must be a valid post code if populated
Contact number	Valid format for a UK landline or mobile number
Overseas travel, if relevant (destinations and dates)	
Derived data fields	
Date case reported	Must be a valid date format in the past
Date case entered into case management system	Must be a valid date format in the past, must be after date case reported
Date case reported to surveillance system	Must be a valid date format in the past, must be after date case reported

The minimum data requirements to ensure meaningful public health action must cover the patient (name, date of birth, gender and NHS number), identify the disease this patient might have (the diagnosis code), and provide correct contact information on the patient and the diagnosing clinician (address and phone number).

To ensure that the system is also fit for the purpose of surveillance it should also have pre-populated values in a drop-down list for the diagnosis, dates of onset, reporting, and a patient post code to support geospatial mapping.

As these data fields are key fields, it is necessary to monitor the validity of these fields against the validity rules presented above. Additional rules might be needed. If data validity presents an actual problem, solutions need to be found to improve data validity.

Step 5: Use evaluation data to set up representativeness measures and inform the selection of additional monitoring measures

'A public health surveillance system that is representative accurately describes the occurrence of a health-related event over time and its distribution in the population by place and person' [22].

Monitoring data representativeness should be considered if systems rely on processes that vary over time or if the coverage of a population under surveillance fluctuates. In practice, any system that collects health data from more than one site and/or at more than one point in time should strongly consider adopting representativeness measures.

Representativeness over time requires the establishment of a set of longitudinal data quality measures, possibly by a series of reports. The monitoring report can help to identify hypotheses as to why a system is not representative and is often a starting point for a more rigorous analysis or evaluation. For instance, a laboratory surveillance system may see a gradual year-on-year increase in the numbers of reported tuberculosis (TB) cases, which may be due to a greater reliance on automated data capture systems, an increase in the number of reporting laboratories, or an actual increase in the instance of the disease. It should be noted that a standard monitoring report can only identify a steady rise and not the underlying cause.

Representativeness across the population can be on the basis of geography or other factors (socio-economic group, ethnic group). Lack of representativeness may result in a bias in the surveillance data. It is difficult to monitor

representativeness directly and routinely. However, proxy measures can be used that generate alarms when there is suspicion of representativeness issues. A direct assessment of representativeness requires a thorough revision of the surveillance system, which is usually carried out in conjunction with a large-scale system evaluation.

Indirect representativeness measures that can be monitored over time will usually include a count of cases or reports over time (or other dimensions/categories). Discussions with system users may illustrate possible reasons for observed variations in these indirect measures, and some of these could be associated to changes in system representativeness.

In the British statutory notification system (Notifications of Infectious Diseases or NOIDS), a system evaluation was performed which found (through a survey) that general practitioners varied in their understanding of the law regarding the notification of infectious diseases. In response to this, monitoring measures could be set up to determine the representativeness of the system over time.

To measure whether general practitioners are effectively reporting, information would be required on who is reporting (name of general practitioner and possibly post code of the practice) and how often (date reported). These data could be used to identify variations in the reporting from general practitioner practices across the country.

Step 6: Review the chosen measures during every system evaluation

A system evaluation gives the system owner the opportunity to review the entire system, including elements which normally cannot be assessed during the continuous monitoring process.

In the example from the British statutory notification system, an evaluation might have resulted in a proposal to automate the extraction of data from systems used by general practitioners to improve the external completeness of the system by identifying cases which otherwise would have not been reported.

System modifications may also inform the development of data quality monitoring reports that look at the internal completeness of the entire dataset to demonstrate that automatic reporting can improve the completion of fields in a case report. A system review report may also look at the number of cases reported by general practitioner practices that pilot-tested the modified system and compare them with the numbers reported in previous years.

2.2 Monitoring completeness

The basic steps of the monitoring process for the internal completeness of a surveillance system are presented in Figure 2.2a below. The approach is sequential and applies to each of the following attributes: completeness, validity and timeliness.

Figure 2.2a. Steps designing and implementing a monitoring system for internal completeness

Step 1. Describe the system

Step 2. Plan the monitoring process

Step 3. Measure internal completeness and summarise results

Step 4. Interpret the results of the measurement

Step 5. Make suggestions for remediation

2.2.1 Describe the system

All major steps needed for a comprehensive description of a surveillance system are outlined in Section 1.3 above. Summaries of this description are helpful when designing processes to monitor the internal completeness of a surveillance system as this information will determine the characteristics of the monitoring system (Table 2.2a.).

Table 2.2a. Checklist of background information needed when designing a monitoring system for internal completeness

Descriptor	Specification for internal completeness
Priorities/objectives of surveillance system	Describe the main purpose of the surveillance system in order to define variables for which high levels of completeness are required. For example, describing the temporal and/or geographical distribution of cases and identification of probable risk factors. In order to estimate the geographic distribution of cases, a system needs to have a high level of internal completeness for geographic variables: 'place of residence' or 'place of exposure'. The variable 'travel' is particularly important when the surveillance system is used to estimate the percentage of travel-related cases of a disease.
Type of surveillance system	Mandatory systems and active surveillance systems usually require a degree of completeness in order to comply with the regulations or objectives of the system.
Type of reporting	Specify if the system is: • electronic; • paper-based; or uses • a mixed approach
Type of data	Establishing a monitoring system for the completeness of case-based data is more complicated than for aggregated data.
Levels of reporting system	A detailed description of the data flow through the various levels (district, regional, national) should be carried out to identify levels where cases/information could be lost: • Reports are sent from data providers to local public health departments and from local public health departments to the central/national level. • Reports are sent from data providers directly to the central level. • Reports are sent from data providers to the regional level and then to the national level.
Diseases being monitored	Identify which diseases in the surveillance system should be monitored for data quality.
Variables/fields by disease	Description of all collected variables, their formats and validation checks. This may result in ideas on how to simplify reporting and thus lead to improved completeness. Disease-specific variables should be customised to meet surveillance objectives.
Providers of data for the surveillance system	• Description of all data providers: one data provider versus many • Data entry (or completion of notification forms for paper-based systems): specify name/title/position and qualifications of the data entry operator (doctor, nurse, administrator)
Healthcare system constrains	Description of resources used by the healthcare system to run disease surveillance activities in order to identify possible limitations that may hamper data completeness. Even though these restrictions are external to the surveillance system, they should be anticipated and addressed, and the way they affect the monitoring system should be predicted. Such restrictions may include any of the following: • Lack of computers or other communication equipment, e.g. fax machines • Lack of personnel • No or slow internet access • Lack of technical support (no information technology professional available) Healthcare system factors which can influence surveillance quality include private vs. public insurance policy, policy for laboratory testing, organisation of primary care assistance (as opposed to secondary and tertiary assistance), regional vs. national governance, etc.
Variables added to the system at a later stage	Description of the process and timeline of data entry from first notification to case confirmation and final clinical outcome. For example, the time interval between the first report and the completion of laboratory testing is determined by the availability of tests, applicable techniques, laboratory routines, and similar factors. The total duration of this delay should be known when calculating and interpreting the completeness of variables for laboratory investigation.
Approaches already in place for managing missing data	Description of analysis and plan for handling missing data. This will improve the understanding of current data quality issues and suggest solutions on how to manage missing values.

2.2.2 Plan the monitoring process

In order to set up a data completeness monitoring system, a number of preliminary decisions need to be made.

Figure 2.2b. Checklist: main considerations for the establishment of a monitoring system

No.	Question
1	Which diseases will be monitored?
2	Which variables (by disease) will be included in the system?
3	Which measures will be used?
4	What is the acceptable level of internal completeness (by variable)?
5	Which level represents the monitoring standard?
6	How often will system data be used for follow-ups, analyses, and interpretation of results?
7	Which approach will be used for managing missing data?

Decisions 1 and 2

These decisions relate to the diseases and the variables by disease that will be monitored. The number of diseases and variables depends on the available resources and the priority objectives of the system.

One should also keep in mind that the surveillance priorities of systems may change over time. For example, during an influenza pandemic some additional surveillance components may be required for the early assessment and investigation of cases, which is usually not necessary for seasonal influenza surveillance. The same applies to measles or other diseases for which a global eradication goal has been set. Here, a high level of internal completeness of the system is essential, especially regarding information on importation, travel within the incubation period, and vaccination history.

When deciding on the scope of the monitoring system, high data quality requirements should be limited to essential variables. A large number of monitored diseases/variables will lead to an increased workload while producing results of limited use.

Decision 3

This decision involves how completeness will be measured and which types of completeness checks will be performed at different levels of the reporting process and within the information system (information technology infrastructure). Different checks can be suitable at different levels but data completeness should be monitored at every step of the data flow. Completeness can be measured by disease or by percentage of completed data fields (e.g. variable, geographical unit, and reporting year). If the number of variables included in the surveillance system is large, completeness can be measured for selected variables that are essential to address surveillance objectives. Completeness can also be measured over time by performing a time series analysis of the percentage of completed data fields for specific variables; it should also be checked if any public health intervention could have impacted the time series.

Decisions 4 and 5

Determining an acceptable level of completeness for a variable involves a decision which depends on several factors. This level may vary substantially between countries, based on different national priorities and objectives for the surveillance system. For example, if an eradication programme for brucellosis is in place, place of exposure is an important variable, and a high level of completeness (for example 80%) is required. Since cases of the disease under eradication are rare, the probable place of exposure should always be reported so that appropriate measures can be taken. On the other hand, when brucellosis is endemic and mainly concerns occupational exposure (e.g. farmers), the respective acceptable level of completeness can be lower.

The level of data completeness considered acceptable is also based on international targets and standards. For example, the World Health Organization requires countries that strive to eradicate measles to report a comprehensive set of case data (e.g. place of exposure, measures taken, and contacts).

The level of completeness expected from a disease surveillance system has to be decided in accordance with a country's resources and the capacity of its healthcare system. It is essential that the acceptable level of completeness is feasible and can be achieved within the restrictions of the existing system.

The acceptable level of completeness (reporting proportion) may change over time, depending on possible changes of the factors mentioned above.

When monitoring surveillance systems, the intervals for follow-ups, analyses and the interpretation of results need to be decided prior to the actual monitoring. In addition, the team responsible for monitoring the system should be nominated well ahead of time. The analysis and interpretation of system checks should take place as soon as possible after the actual check – and within the capacity of the system.

Decision 6

This decision relates to the frequency of follow-ups, analyses and interpretations in connection with the results of a system check; it also addresses the criteria that should be applied to nominate the staff members responsible for these functions. The analysis and interpretation of system checks should be as close in time to the actual system checks as possible. In addition to the routine checks and follow-ups, a periodic evaluation of the monitoring system should be performed; this evaluation should include a review of its scope and requirements.

Decision 7

This decision relates to the approach used to manage missing data at an early stage and the results of such measures. Often, an effective monitoring process is available, but the outcomes are not used to improve the system.

The following blank table summarises the process for setting up a monitoring system for completeness (Table 2.2b.).

Table 2.2b. Table summarising the monitoring plan for internal completeness

Variables to be monitored, by disease	Acceptable level of completeness	Frequency of follow-up	Possible approaches for management of missing data
Disease 1			
Variable a			
Variable b			
Variable c			
Variable d			
Variable e			
...			
Disease 2			
Variable a			
Variable b			
Variable c			
Variable d			
Variable e			
...			

Variables that are common for all diseases under surveillance (e.g. age, gender, and case classification) can be arranged in a table format (Table 2.2c.).

Table 2.2c. Table summarising the monitoring plan for internal completeness of common variables

Variables to be monitored for all diseases	Acceptable level of completeness	Frequency of follow-up	Possible approaches for management of missing data
Variable a			
Variable b			
Variable c			
Variable d			
Variable e			
....			

2.2.3 Measure internal completeness and summarise results

An easy and straightforward way to measure the internal completeness of surveillance data is to calculate the percentage of missing values ('unknown' or 'blank') within a dataset. More specifically, the percentage of missing records by variable and also the percentage of missing values by record can be calculated.

Time should be taken into account when measuring completeness by indicating a) the percentage of records that are complete on initial report and b) the percentage of records that are complete within an appropriate time interval after the first submission. The latter usually applies to specific variables such as laboratory results because laboratory results may become available after the case notification. Table 2.2d presents some examples of variables that are usually not completed at the time of initial notification.

Table 2.2d. Examples of fields that are not completed at initial notification

Variable	Estimated time
Laboratory result (tuberculosis)	Sputum culture result may not be available for as long as two months.
Laboratory result (brucellosis)	Culture result may not be available for four weeks.
Laboratory result (yersiniosis)	Culture result may not be available for four weeks.
Serotype or phage type results (especially for reporting of clusters)	Unspecified: depends on the time needed for referral of isolates from primary laboratories to reference laboratories and the time needed for serotyping/phage typing at the reference laboratory.
Clinical outcome	Time cannot be specified.

When calculating the percentage of missing values, the denominator should reflect the number of reported cases to which the field/variable is actually applicable. The completeness for some variables will depend on the value of another variable. For example, a value for *Salmonella* serotype is expected only when stool culture for *Salmonella* is actually positive. Cases with a negative stool culture, or cases for which stool culture was not performed, should be omitted when you calculate the percentage of missing values for the field 'serotype'.

Internal completeness can be summarised using simple tables which contain the percentage of missing records for all variables; further tables can be created with the mean number of missing values for each record and disease (Tables 2.2e and 2.2f).

Table 2.2e. Completeness results by variable

Variable	% of records with completed data field	Target level of completeness	Action required
Disease A (n* = ?)			
Variable a			Yes/no
Variable b			Yes/no
Variable c			Yes/no
Variable d			Yes/no
Variable e			Yes/no
...			

** n = number of records (cases reported)*

Table 2.2f. Completeness results by record

	Mean number of missing values, by record	Target level of completeness	Action required
Disease 1			Yes/no
Disease 2			Yes/no
Disease 3			Yes/no
Disease 4			Yes/no
Disease 5			Yes/no
...			

2.2.4 Interpret the results of the measurement

In order to facilitate the interpretation of completeness measures, results should be accompanied by an explanation of possible factors which contribute to low completeness. These factors can either refer to the surveillance system itself or to external circumstances or constrains (economic, political). These factors can be investigated through a specific evaluation of the surveillance systems. A checklist for summarising possible reasons for low internal completeness is presented in Table 2.2g.

A list summarising the relevant completeness measures for the diseases under surveillance – together with desired completeness targets – can be compiled. Variables which do not meet the desired targets of completeness can be flagged automatically by the system.

It is also suggested to check whether internal completeness has a temporal distribution, which is useful in identifying seasonal reasons for low internal completeness of specific fields. This can be easily accomplished by drawing a line graph with the monthly (or weekly) percentages for the year in question. If completeness is lower only during some weeks of the year (e.g. in summer), it is important to point this out and research the reasons.

Table 2.2g. Checklist of possible reasons for low completeness of the variable 'culture result' for tuberculosis notification

Disease and variable	Possible reasons for low completeness	Result
Tuberculosis, culture result	The format of the reporting form is not appropriate and information is missing.	Yes/no
	The electronic data entry system is not user friendly (e.g. not convenient to go back and complete missing data for a previously entered record).	Yes/no
	Variable specifications are not well documented (e.g. no clear definition of the variables to be collected).	Yes/no
	Lack of a surveillance protocol; poor or missing instructions for the data providers	Yes/no
	Poorly designed coding system which does not allow for an appropriate level of completeness	Yes/no
	Missing values indicate that culture has not been performed	Yes/no
	Data lost during transfer of data from one database to another	Yes/no
	Completeness measured too early for cases with recent onset	Yes/no
	Other reasons. If yes, describe:	Yes/no

If the completeness rate is subject to seasonal variation, possible explanations should be mentioned. For example, the low completeness of the variable 'culture result' for tuberculosis in August can perhaps be explained by a change in laboratory routines in the same month. In other instances, a drop in completeness rates may reflect summer holidays taken by the personnel whose absence leads to delays in updating the database.

Data providers should be contacted to provide feedback on the reasons for the low completeness of a variable, for example by asking them to complete the above checklist and research the actual reasons behind the low completeness levels.

2.2.5 Make suggestions for remediation

Database operators can use a number of approaches for the management of missing data.

Managing/retrieving missing data

Main approaches to recover missing data:

- Personally following up with data providers to request missing data
- Using an automated and systematic approach for follow-up
- Managing missing data during statistical analysis

Following up with data providers

This is the most widely used method. The main advantages of this approach are simplicity and the fact that it leads to direct and timely results. This approach does not necessarily require the support of a digital automated system and is also applicable to paper-based reporting systems: epidemiologists/public health professionals who work at the local level can carry out manual checks, on a case-by-case basis, upon reception of paper-based notifications.

Another advantage of direct contacts between public health professionals and data providers is that data providers are usually clinical doctors who can provide additional information needed for public health action because they would know, for example, whether the family of a patient with meningococcal meningitis received prophylaxis, or whether screening tests were performed on contacts of a patient with tuberculosis.

The main drawback of this approach is that it is time-consuming (and therefore not cost-effective) and labour-intensive. Moreover, this approach has no training component, so data providers will not be trained to collect more complete data. If contact with the data provider is delayed, it becomes increasingly difficult to locate data on specific patients, especially in hospitals with no electronic systems.

Automated, systematic approach

This approach refers to avoiding missing data at the data entry stage: all electronic forms/data entry masks for local or centralised public health authorities have a built-in mechanism which automatically performs a completeness check when data are entered.

The system automatically generates reports, which are then sent to data providers to inform them of, and request, missing data. Such automated reports can be generated at the local (preferable) or central level. For example, a central level report with missing geographic data can be sent to the corresponding local public health authorities to request missing information; alternatively, it can be sent directly to the data providers, depending on the characteristics and structure of the system. If a report on missing data was generated by a system at the local

level, it can be sent directly to the data provider, e.g. hospitals, laboratories, or private doctors. This approach only puts moderate demands on data providers and the monitoring team.

The main drawback of this approach is the time and effort that has to be invested when setting up the system and the fact that information technology experts are needed. In addition, this approach lacks direct communication channels with data providers to resolve misunderstandings or problems with the database or reporting forms.

Choosing an approach depends on the characteristics of the surveillance system, on the processes already in place, and the capacity of the public health authorities at the local and central level. A mixed approach is also an option, and in some cases is the most efficient one.

Examples of automated requests for missing data include the following:

- Automated requests can be used for revising pending laboratory results. Completing follow-ups and summarising the missing information should be as user friendly as possible.
- An automated mechanism to notify local public health authorities or doctors that they need to complete the information on the outcome of the disease (e.g. for listeriosis, for which it is important to know if and how many people died).
- An automated message for tuberculosis can be sent to clinicians asking for follow-up (e.g. regarding compliance to treatment).

If the same case is notified by two different sources, the two records can be compared for missing variables. For example, when a salmonellosis case is reported through a mandatory notification system and the national reference laboratory for salmonellosis, and the age of the case is only given in one of the two databases, missing information can be easily transferred. Matching and comparing data from different sources can improve data completeness for both systems. It is important, however, that the internal validity of both systems is confirmed.

Modifications during data analysis

Apart from the built-in automated mechanisms for increasing completeness, there are some specific methods that can be used to treat missing values during data analysis. These methods depend on the nature of the missing values (missing at random/not at random) [9].

The term 'missing at random' can be used when the probability that a value is missing is unrelated to the value itself or to the value of any other variable [9]. For example, when information for the variable 'travel history of case' is missing independently of the true value of the nationality of the case, the time of year, or any other characteristic, missing data are considered 'at random'.

This type of missing data does not affect the validity of the gathered data and mainly impairs the power of the dataset, therefore the impact on the data quality is smaller than the impact of data missing not at random.

Example: When the culture results for a tuberculosis case are missing, and their absence is not related to the result of the culture (positive/negative) or other characteristics of the person (e.g. age, nationality), the culture result can be considered a value 'missing at random'. This may happen by mistake during data entry, or because someone forgot to enter the data.

On the other hand, when the probability that a value is missing is related to the value itself or the value of other variables of the dataset, the missing value is characterised as 'not at random' [9].

If the assumption of values 'missing at random' is plausible, missing values can be estimated using various statistical technics (e.g. imputation) [11].

Examples:

- Place of residence or travel history is missing more frequently for cases that involve immigrants than for those that involve native-born citizens.
- Information on sexual behaviour is reported more frequently for homosexuals than for heterosexuals.

These missing values impair the validity of the data. In order to obtain an unbiased estimate of parameters one has to model 'missingness,' meaning that a model has to be created that accounts for the missing data. More complex models can be used for estimating missing values.

This approach is not always feasible when dealing with surveillance data, and experts on modelling should be involved.

Suggestions for improving completeness

Apart from retrieving missing data, structural changes in the system and its components may be needed in order to increase internal completeness. Gap analysis (also called need-gap analysis) is a broadly used technique for determining the steps that are needed for moving from a current state (completeness based on the results of monitoring) to a desired future one (acceptable completeness).

The monitoring system is actually the tool that is used to document the difference/gap between the expectations regarding completeness and the actual internal completeness of the system. The next step after that is to document areas that need improvement. A list of factors/gaps that may have contributed to the low internal completeness of the system can be drawn. This list is a simple method for identifying problems in a notification workflow (see Section 3.2.4).

The factors which will have to be addressed in order to achieve completeness objectives are shown in Figure 2.2c. Sometimes the observed level of completeness differs substantially from acceptable completeness, so the first objective usually is to close this observed gap.

The choice of the methods/changes that will be used to close the above gap depends on: a) the results of the measurement, b) the factors chosen to be addressed, c) the available resources, d) the characteristics of the system, and e) the methods are already in place to increase internal completeness.

Figure 2.2c. Methods at the data entry level for enhancing internal data completeness

1. Use of mandatory fields
2. Use of repeatable fields
3. Logical order of variables
4. Coding of possible values
5. Appropriateness of data values
6. Automated help texts/notes during entry
7. Additional automated notes/comments on the relevance of specific information
8. Checks preventing losses during data transfer or flagging up discrepancies between different databases

Use of mandatory fields

Mandatory fields are those that require 100% completion. This means that in an electronic system, data cannot be saved and sent unless those fields are filled in. In some cases, the system lets the user save incomplete datasets. However, incomplete items will eventually be flagged up by the system for further examination (usually an automated process).

In systems which transmit notifications by fax or mail, data validation cannot be automated. The completeness of a mandatory field can only be verified when data are manually entered into an electronic system (either at the local or central level), and epidemiologists need to ask their data providers for any missing information.

Fields that can be defined as mandatory for all diseases/ for a subset of diseases

Mandatory fields include the following:

- Fields necessary for matching and comparing data from different sources
- Fields containing essential descriptive data (time, place, characteristics of cases)
- Fields needed for case classification
- Fields containing important information for public health action
- Measures of generally unknown factors (newly emerging diseases)
- Variables that contain important information on current system priorities

Examples of fields that could be considered mandatory are included in Table 2.2h.

Table 2.2h. Examples of possible mandatory fields, by disease

Data entry field	Diseases covered
Matching and comparing data from different sources	
Unique personal identifier	All diseases
Disease identifier	All diseases
Notification date	All diseases
Descriptive data	
Age	All diseases
Gender	All diseases
Place of residence	All diseases
Date of onset of symptoms	All diseases
Necessary information for case classification	

Data entry field	Diseases covered
Laboratory results	(Almost) all diseases
Important for public health action	
Occupation/school attendance	Meningococcal meningitis
Food handler	Typhoid/paratyphoid fever
Emerging diseases	
Travel history	Cholera, dengue, malaria, poliomyelitis, rabies, yellow fever, etc.
Priorities of the system	
Brucellosis (in the process of being eradicated)	Place of exposure

One should not define too many mandatory fields in a surveillance database as this would work as an anti-incentive to notification/reporting. Therefore, fields should be only defined as mandatory if the necessary information for completion can be found and is essential to the achievement of the surveillance objectives.

Mandatory fields should be compatible with the requirements of ECDC's The European Surveillance System (TESSy). However, additional mandatory fields may be necessary to meet national surveillance objectives and address the local epidemiologic situation.

Using repeated fields during data entry

The use of repeated fields can make data entry easier and quicker. The benefit of this method should always be balanced against the possibility that errors will be introduced. It is not a suitable method for variables that have many different possible values.

Example:

- The field 'source of isolation' for *Salmonella* cultures can be repeated (stool samples for the vast majority of reported cases).

Logical order for variables

A misplaced variable can lead to low completeness, so structuring the reporting form is important. For example, all data fields connected to the clinical history of a case should be grouped together, as this information will be provided by the treating physician. The same applies to laboratory data, usually reported by a microbiologist.

Example:

If the field 'outcome' in an electronic reporting form is included under the section on laboratory results, it will often be left blank as this field is usually completed by clinicians and not microbiologists: clinicians are likely to skip over this section of the form because they think it is only relevant to microbiologists, not noticing that the variable on the outcome of the disease requires their input.

Predefined values

Appropriate coding of categorical variables is important because it makes data entry easier for data providers and leads to increased completeness. Providing lists with predefined values (e.g. pre-populated drop-down field options) for as many fields as possible can be very helpful, especially for frequently diagnosed infectious diseases, such as salmonellosis.

Examples:

- Place of residence (predefined list of regions, prefectures)
- Country of origin (predefined list of countries)
- Ethnicity (predefined list)
- *Salmonella* serotype (list of predefined serotypes instead of a free-text field)

Predefined data values should be appropriate. In general, values should be specific, understandable and mutually exclusive. Predefined options (such as pre-populated drop-down lists) in an electronic system should be frequently updated and always include the option 'other' and a free-text field. Additionally, the option 'unknown' should be included in the list so that missing and unknown values can be recorded.

Example:

- Low completeness for the field 'serotype' for salmonellosis may be caused by new serotypes of *Salmonella* which were not included in the list of predefined choices.

Help texts/notes during entry

Certain fields may require automated help texts, a definition, or examples. Help texts can increase completeness because data providers may skip fields or questions they do not understand.

Examples:

- The field 'high-risk group' may need further clarification as it can mean different things to different doctors or in different settings.
- The field 'imported case' also needs a clarification as it depends on the incubation period of the disease, the mode of transmission, and the travel destination (for diseases not endemic in a country). For diseases with a long incubation period, characterisation of a case as imported may be quite complicated.

Context-specific messages

When a field is skipped, automated pop-up messages can prompt for the missing data. Prompts should be brief and state in simple words why the missing information is important and how it will be used by public health authorities.

One of the main reasons for low internal (and also external) completeness is that clinicians consider notification as a bureaucratic procedure for the collection of statistical data. The goal for database operators should be to make it clear that surveillance data are used for public health action.

Examples:

- When the field 'occupation' is not completed under typhoid/paratyphoid fever or tuberculosis, a pop-up message could emphasise that certain public health measures are based on the occupation.
- A pop-up message for the field 'travel history' for tuberculosis could explain that this information could be helpful to locate contacts during transit, as tuberculosis is not usually a travel-related disease.

Avoiding data loss during transfer

Loss of data is sometimes a problem when transferring data to a database that does not support the language of the originating database in which the data were created; information contained in the variables can become unreadable, or get lost, if a local language is not supported.

2.3 Monitoring internal validity

Internal and external validity is defined in Section 2.1 of this handbook. Information should be sufficiently free of error to allow the objectives of the surveillance system to be met. Invalid data can impair the understanding of the true epidemiology of the communicable diseases under surveillance and can lead to unnecessary or even harmful public health interventions. Regardless of how errors occur, they can have substantial influence on the effectiveness of the surveillance system and the subsequent public health measures. However, improving internal data validity will not correct for biases introduced by external validity issues.

The steps for monitoring internal validity of a surveillance system are listed in Figure 2.3a.

Figure 2.3a. Monitoring internal validity

Step 1. Describe the system

Step 2. Plan the monitoring process

Step 3. Measure internal validity and summarise the results

Step 4. Interpret the results of the measurement

Step 5. Make suggestions for remediation

2.3.1 Describe the system

The key elements for the description of a surveillance system are listed in Section 1.3 and in Figure 1.3a. More specifically, before preparing a monitoring system for internal validity, system operators need to develop a keen insight into the main processes of the surveillance system. When designing monitoring processes, one should be familiar with the main characteristics of the respective surveillance system (Table 2.3a).

Table 2.3a. Checklist: necessary background information when designing a monitoring system for internal validity

Characteristics	Specifications for internal validity
Priorities/objectives of a surveillance system	For example, in order to identify a gastroenteritis outbreak, a high level of internal validity is needed for the variables 'number of gastroenteritis cases', 'epidemiologic link with other cases', or 'incriminated food item/water'.
Levels of reporting system	The level (district, regional, national) at which the monitoring system will be established.
Health outcomes under surveillance	Numbers of diseases; types of diseases
Numbers of variables/fields by disease	Increasing the number of fields usually increases the possibility of errors and warrants a more complex monitoring system.
Types of variables	Different types of variables (e.g. categorical, numeric, text) require different approaches during the monitoring process.
Providers of data and/or source of data	Depending on the source (one versus many, laboratory versus clinical records) and the provider (doctor, nurse, administrator), different validity issues can be expected.
Adapting approaches which are already in place in order to manage invalid data and identify discrepancies/possible errors	Proposals for a) approaches for correcting invalid data, and b) additional changes/validity checks – based on what is already in place.

2.3.2 Plan the monitoring process

Some basic considerations about internal validity to be taken into account when preparing the monitoring system are described in Figure 2.3b. The decisions that will be made during preparation mainly depend on the characteristics of the surveillance system and the available resources of the system.

Figure 2.3b. Checklist: main considerations/decisions about internal validity when setting up a monitoring system

- Which diseases will be monitored?
- Which disease variables will be included in the monitoring system?
- Which measures will be used to quantify validity issues?
- How does one prevent invalid data entry?
- Which level of internal validity, by variable, is acceptable?
- Which levels of the system will be checked by the monitoring system?
- How often will follow-ups, analyses and interpretation of results be carried out?
- What kind of approaches will be used for managing invalid data?

Different types of errors can occur as data are processed by the surveillance system: random errors and systematic errors.

Random errors

The most common ones are the random errors during completion of the reporting form or during data entry. These errors may cause false alarms for public healthcare services. Often, these errors can be identified by implementing simple validation rules. Below are some examples of this type of error:

- Reporting 40 cases of gastroenteritis instead of four
- Reporting a case of poliomyelitis in a European country instead of a case of pertussis
- Choosing serotype typhi instead of typhimurium for some salmonellosis cases
- Choosing the value 0 (no travel history) instead of 1 (travel history) when reporting a malaria case
- Typing 01/05/2012 as the date of symptoms onset instead of 01/01/2012
- Choosing congenital syphilis instead of syphilis
- Choosing 1 (female) instead of 0 (male) when reporting a case of sexually transmitted disease among MSM (men who have sex with men)

Systematic errors

Some types of errors can occur systematically during the reporting of surveillance data and thus are difficult to detect. It has been documented, for example, that data entry operators show a systematic digit preference when filling in numeric fields for age (ages ending in 0 and 5 are overrepresented) or date (01, 10, 15 and 20 are overrepresented) [9,10]. This digit preference may be also combined with the tendency to avoid certain 'unpleasant' numbers, such as 13.

Errors can also occur because of insufficiencies of the system. For example, new diseases or parameters that the system has not taken into account cannot be reported. When *S.* enterica serovar 4,[5],12:i:- was first recognised, some systems reported related cases either as unknown or as *S.* typhimurium because this new monophasic *Salmonella* serovar was not included in the pre-populated list of values.

Unclear case definitions or case definitions that are not widely known also lead to reports with invalid epidemiologic data. For example, paediatricians sometimes report diagnosed cases of childhood diseases such as measles as verified, even though the EU case definition requires laboratory confirmation.

Since perfect validity is not a realistic goal, and errors are expected to occur during reporting, an acceptable level of validity should be set for each single variable before the system is activated. The acceptable level should be based on the objectives of the surveillance system, its characteristics, and the resources available to the system.

For some variables, especially those that are important for public health action, a validity rate of 100% may be needed, while for other variables a lower level of validity may be acceptable. Practical restrictions should always be taken into account when designing a monitoring system for validity so that the established standards of the system can be met. The same criteria should be used when selecting the health outcomes (e.g. cured) and associated variables that will be monitored. Ideally this would include all possible health outcomes, all variables in the monitoring system, and all set goals regarding validity.

2.3.3 Measure internal validity and summarise the results

Validity checks identified during the planning of the monitoring process can be conducted at different stages of the reporting process: during data submission, during validation or during analysis.

During data submission/manual data entry or as part of a file upload process, validation rules can be applied to each variable. For example, during the automatic validation process TESSy distinguishes three error levels:

- Errors. A TESSy *error* message indicates that the structure of the file should be corrected. The file is rejected, should be reformatted and then uploaded again.
- Warning: A TESSy *warning* message is always combined with a course of action to improve data quality.
- Remark: A TESSy *remark* informs the user that steps should be taken to improve data quality.

Table 2.3b. Examples of validation rules and validation messages in TESSy

Variables in rule	Severity	Validation rule	Validation message
'DateOfNotification', 'DateOfOnset'	Warning	If DateOfOnset is after DateOfNotification	DateOfOnset cannot be later than DateOfNotification; DateOfOnset will be set to [Default value].
'Gender', 'Transmission'	Error	If Gender is 'F' and Transmission is 'MSM'	If reported transmission category is men who have sex with men (Transmission = MSM), then Gender should not be female (F).
'DateOfDeath', 'Outcome'	Warning	If DateOfDeath is 'UNK' and Outcome is 'D'	If it is known that a person died, it is usually expected that the DateOfDeath is reported.
'Transmission', 'TransmissionHetero'	Warning	If Transmission is 'HETERO' and TransmissionHetero is not reported	Transmission subcategory TransmissionHetero should be reported if transmission category is 'HETERO' (Transmission = 'HETERO').
'DateOfAIDSDiagnosis', 'Stage'	Warning	If DateOfAIDSDiagnosis is not reported and Stage is 'AIDS'	If stage of disease is reported as AIDS (Stage = AIDS), then DateOfAIDSDiagnosis should be reported.
'CD4Cells'	Error	If CD4Cell count is less than 0 or greater than 6000 per cubic millimetre.	Usually, CD4 cell counts range from 0 to 1 500. Extremely high values are rare, and the upper limit has been set to 6 000 per cubic millimetre.

During data exploration the system checks the distributions of variables, performs cross-tabulations to identify outliers, and inconsistent values, etc.

A data analysis can identify inconsistencies in the data and show irregularities and deviations from standard patterns. The analysis of surveillance data is mainly descriptive, and validity assessments should be defined when analysing data on people, time, and place.

When certain signals are detected, for example a single autochthonous case of poliomyelitis in a European country or an outbreak of gastroenteritis, it is important to verify that the surveillance data are correct before applying public health measures.

The validation steps to be performed during a descriptive analysis can be summarised as follows:

- Comparing absolute numbers. The number of notified cases should be compared with the expected number of cases for this disease (for a given time period and geographic area); deviations should be checked for possible errors. This approach is more suitable for rare diseases (e.g. anthrax, haemorrhagic fevers).

- Proportion of notifications of a disease. The proportion of notified cases of a specific disease in relation to the total number of cases for all diseases is compared with the expected proportion of cases of the disease. This proportional morbidity is a useful measurement in the absence of denominators (e.g. 30% of all reported cases this year are tuberculosis, 10% of all reported cases are measles).
- Compare incidence. The rate of notified cases using the population as denominator is compared with the expected rate in the same population based on previous periods. Observed and expected incidences are compared, which provides a practical way to identify errors for common diseases (e.g. the incidence of *Campylobacter* infection was 12 cases/100 000 population in one of the Member States in 2010.)

The results of the descriptive analysis will be compared with the expected values which were calculated based on the available background surveillance data. These background data may refer to a whole country or to smaller geographic regions and populations, or even to a specific hospital or healthcare unit. Furthermore, they can refer to different time periods (e.g. year, semester, month). Selection of time and place will be decided according to the capacity of the monitoring system. In general, more calculations (of increasing specificity) may reveal more discrepancies in surveillance data.

The question that normally arises at this point is whether a difference between an observed and an expected value is associated to epidemiological factors or due to an error (for example invalid data), which will have to be investigated further.

There are various mathematical models which help to detect statistically significant deviations from expected values [11,12]. These models take into consideration the day-of-week effect, seasonality, and several other factors. When performing this type of analysis, it is important to check whether an observed value (which may or may not be statistically significant) is an artefact or real.

When the expected number of cases (the expected proportion or the expected incidence) is significantly higher than the observed one, one of the possible explanations could be increased morbidity. Before initiating public health measures, authorities should confirm than this is not a false alarm due to enhanced notifications, demographic changes, or notification errors.

The blank tables below can be used for the presentation of the results of a descriptive analysis in the context of identifying errors/false alarms (Table 2.3c, Table 2.3d and Table 2.3.e).

Table 2.3c. Number of notified cases and expected number of cases for a specific time period (e.g. week, month) and geographic area (e.g. whole country, region); based on background data

Disease	Number of notifications	Expected number of notifications
Typhoid fever		
Death		
Hypertension		
...		

Table 2.3d. Proportion of notifications of specific diseases and expected proportion of notifications for a specific time period and geographic area; based on background data.

Disease	Proportion of notifications*	Expected proportion of notifications
Typhoid fever		
Death		
Hypertension		
...		

** Number of notified cases of the disease/total number of notified cases*

Table 2.3e. Reported incidence and expected disease incidence in a specific population for a given time period; based on background data

Disease	Number of notifications per 100 000 population*	Expected number of notifications per 100 000 population
Typhoid fever		
Death		
Hypertension		
...		

** Or any other reference unit, e.g. 1 000 000*

Another approach to identifying possible errors/inconsistencies is to use calculations that pertain to specific variables.

Additional errors can also be detected by analysing disease distribution and frequency by age and gender, the type of hospital, and other factors (appropriately selected for each outcome). For example, the distribution of congenital syphilis or toxoplasmosis cases by age can pinpoint mistakes made during entry of the patient's age. Large discrepancies among hospitals can help in identifying systematic errors. When, for example, a hospital reports the vast majority of salmonellosis cases without epidemiologic link or suspected food vehicle, while other hospitals report much lower percentages for the same disease, one has to suspect a lack of validity. An equal lack of validity can be assumed if a hospital shows significant differences in the number of notifications from one year to another (provided that was no substantial change in morbidity or in the pathogen). A review of outliers can be used to validate the notifications.

If two or more sources of information for the same cases are available (e.g. a notification dataset and a laboratory dataset), the two records can be matched and compared for consistency; this can be done either manually (this is usually done by epidemiologists at a regional level) or electronically. An automated system can flag up potential errors or inconsistencies and prompt the user to make corrections.

Tables are a helpful tool for summarising errors by disease and detecting disease data which have a low validity (e.g. the number of false alarms or the number of recurring errors). As far as the validity of specific variables is concerned, the percentage of invalid records by variable can be used. Sometimes, the mean number of invalid data points by record for each disease is also used as a measure of validity.

Three blank tables which could be used to summarise internal validity are presented below (Tables 2.3f through 2.3h).

Table 2.3f. Number of recognised false alarms/errors by health outcome for a specific time period

Disease	Number of false alarms/errors
A	
B	
C	
...	

Table 2.3g. Validity results by health outcome and variable for a specific time period

Disease A (n = ?)*	% of records with invalid data
Variable a	
Variable b	
Variable c	
...	

** Number of records (notifications)*

Table 2.3h. Validity results by health outcome for a specific time period

Disease	Mean number of invalid values by record (SD)*
A	
B	
C	
...	

** SD = standard deviation*

2.3.4 Interpret the results of the measurement

Listing errors and their frequency is useful in order to interpret the results of a measurement and identify the diseases and variables with low validity.

The decision which diseases and variables should be corrected depends on:

- the objectives of the surveillance system (may be different for different diseases);
- the impact which each error has on public health interventions;
- the ability of the system to implement changes at a specific point in time (depends on the current system priorities); and

- the types of errors, and the reasons behind them (the type of error determines whether it can be easily fixed).

There are several tools that can be used to identify the most important sources of errors: the cause-and-effect or Ishikawa diagram, the Pareto chart, process mapping, etc. [17-19].

It can be helpful to list all factors that may have influenced the validity of the surveillance data; see Figure 2.3c. A listing of possible factors by disease/variable usually provides the most useful information.

Figure 2.3c. Surveillance systems-related factors which can negatively influence the validity of surveillance data

- Unclear case definitions
- Problematic reporting form (paper-based systems)
  - Unclear questions (e.g. when acronyms, complex words or abbreviations are used)
  - No logical order of questions (e.g. clinical and laboratory sections not clearly separated or mixed-up)
  - Meaning of question is unclear
  - Overlaps
- Problematic structure of database (electronic systems)
  - No logical order of fields
  - Unclear meaning of a field
  - Unclear data values in the pre-populated lists (e.g. when acronyms or abbreviations are used)
  - No pre-populated lists
  - Overlapping values in pre-populated lists
  - No logical order of values for discrete ordinal variables
  - Coding errors
- Problems during data transfer
  - Lack of training of data providers/people responsible for data entry
  - Lack of time (for example, reporting for surveillance purposes is unlikely to be a high priority for busy physicians)

Factors which only concern the surveillance system itself are relatively easy to correct. Factors which have to do with the healthcare system (e.g. lack of personnel, lack of training) or the operation of the healthcare services (e.g. lack of electronic systems/IT support or lack of structured forms for recording a patients' history) are more difficult to deal with and require the support of hospital administrations and often the national ministry of health.

2.3.5 Make suggestions for remediation

Error correction is essential during all stages of the data collection process.

Managing/correcting invalid data

Some errors can be corrected as soon as the epidemiologic data are received and processed. Often, checking back with the data providers, either manually or through an automated process, will lead to the desired results.

Manual process: If inconsistencies or errors are identified by the epidemiologists, direct communication with the data provider is required to correct inconsistent data. This approach is time-consuming, neither objective nor standardised, and is heavily influenced by the work load, training and experience of the receiving team. In addition, time is a major factor because it becomes increasingly difficult to correct data when too much time has passed after notification. However, if manual corrections are carried out regularly and consistently, the quality of data will be greatly improved.

Automated process: An electronic validation system can automatically identify inconsistencies and notify the epidemiologists, who then have to obtain corrections from data providers. Alternatively, the system directly informs the data provider of any inconsistencies in the data. Ideally, automated validity checks are created during the implementation phase of a surveillance system, but can be introduced at any point in time.

Which approach will be used depends on the capacity of the system. Using both approaches in parallel – and at different levels of reporting – is also possible.

Some errors, such as the ones caused by digit preference, can also be corrected at the time of analysis [13,14,20]. This would require data modelling, which is not always feasible because of system limitations.

Maximising internal validity

Performing a gap analysis is one of the methods that can be used to justify structural changes in the system and its components in order to increase the internal validity of the system. As described below, a gap analysis compares the difference/gap between the expectations regarding internal validity with the actual internal validity of a system.

The actual changes made in the system depend on: a) the results of the measurement, b) the factors that should be addressed, c) the available resources, d) the characteristics of the current system, and e) the methods already used to increase internal validity.

The most widely used measures/strategies to enhance internal validity of epidemiologic data include: a) training of data providers, b) modifying reporting forms (paper or electronic), and c) applying validity checks at the data entry level.

Training data providers

Training should be a constant process because there tends to be high personnel fluctuation. In addition, medical school in most countries do not teach surveillance concepts, and if they do, training takes place in the first semesters. Training should focus on:

- the definition and purpose of epidemiologic surveillance;
- types of surveillance systems;
- the health outcomes that are currently notifiable (all levels);
- reporting and legal framework; and
- currently used case definitions

All training courses should improve the motivation of the participating clinical doctors by convincing them that surveillance data are not merely gathered for statistical reasons but are essential for the timely implementation of public health measures.

If nurses or administrative personnel are responsible for data entry, they should also receive training.

It is preferable to provide training locally; trainers should be local public health specialists who are in contact with the central level and informed about the latest concepts in disease surveillance.

Modifying reporting forms (paper or electronic)

Reporting forms should be specific and comprehensive, easy to fill in, and self-explanatory. It is important to collect and use the feedback from data providers regarding the difficulties they had during the completion of the forms. Some basic concepts for creating a reporting form are included in Figure 2.3d.

Figure 2.3d. Basic concepts for creating a reporting form

- Use simple language.
- Use abbreviations judiciously, and only after spelling out the entire expression.
- Avoid epidemiologic jargon; if jargon cannot be avoided, provide explanations (e.g. 'high-risk group').
- Avoid using imprecise or ambiguous questions in reporting forms.
- Avoid questions that cover two or more topics at once (e.g. 'Did the patient present with vomiting and bloody diarrhoea?')
- Avoid long forms that require an unreasonable amount of time for completion.
- Include all possible answers to a question when providing predefined answers; make sure all predefined answers are mutually exclusive.

Forms should not use complicated or overly technical language or abbreviations because forms are not necessarily filled out by clinical doctors. Data entry operators may be unfamiliar with some of the terms, which could lead to misunderstandings, invalid data, and reduced completeness.

Applying validity checks to the system

Some examples of validity checks that can be applied are presented in Figure 2.3e.

Figure 2.3e. Examples of validity checks

- Date of onset cannot be later than date of diagnosis.
- Date of onset cannot be later than date of notification.
- Date of receipt of specimen at laboratory cannot be after the date of receipt at reference laboratory or before the date of onset.
- Invalid dates such as 01/01/2022 or 01/01/0011 should not be allowed.
- A patient with the same name or insurance number cannot be entered twice in the database for the same disease (for a specific period in time).

- If transmission mode is 'men who have sex with men', the gender must not be coded as 'female'.
- If gender is coded as 'male', pregnancy cannot be reported.
- For aggregated notifiable laboratory data: the weekly number of positive cultures (e.g. for *Salmonella*) cannot be higher than the total number of cultures processed in the same week.

Apart from the validity checks, the following good practices should be used:

Restrictions in the number of values/digits of the fields

For example, gender can take three specific values (0 for males, 1 for females, and 9 for unknown), and no other value will be accepted by the system. The variable 'age' is usually restricted to two digits, which would automatically rule out mistakes such as '152 years of age'.

Use of a unique identifier for each patient (mandatory field)

This identifier can be the insurance number or any other available number and can be used as the key index for avoiding duplicates and retrieving or checking reported information regarding the patient. It is important for this number to fall within a certain acceptable range of numbers; the data entry operator should not be able to skip this field.

Skipping fields

Adapting the database so some fields are automatically skipped if they are not applicable can make data entry easier and less time-consuming. For example, the field 'occupation' can be skipped when the patient is a child, and all fields that deal with 'residency at a hotel' can be skipped when a patient with legionellosis has not visited a hotel during the incubation period. When a culture result is negative or unknown, the fields 'pathogen' and 'resistance' can be skipped; when a case is autochthonous, the fields 'country of infection' and 'duration of travel' can also be skipped.

Automated help texts/notifications

Providing automated help texts/notifications with clarifications (e.g. of values, case definitions) is an easy and straightforward way to avoid misunderstandings. This can be useful for case classifications or for the clarification of terms (e.g. 'risk group' or 'imported case').

Value labels

It is also helpful to present associated labels when a value is selected. This helps to identify possible errors at the time of data entry.

Messages/notes about possible or frequently occurring errors

Providing the user with feedback on his/her actions during data entry helps to avoid common errors.

Appropriate coding of possible values and inclusion of the choice 'other' followed by a free-text field

Reported cases should be categorised by public health professionals at the local level rather than by medical practitioners to avoid any concern that the diagnosis is influenced by case definitions. The information needed for case classification (e.g. clinical manifestations, epidemiologic link, laboratory results) will be provided by medical practitioners.

2.3.6 Questions and answers

Q. What is the most important element of surveillance data quality?

A. Data quality depends on the completeness and validity of the gathered epidemiological data.

Q. What does low validity of surveillance data mean?

A. Validity (along with completeness) is one of the basic components of data quality. Low validity of surveillance data distort that the epidemiological characteristics of a disease.

Q. What criteria would you choose when monitoring variables for validity?

A. Include variables that a) are relevant (based on the objectives of the system), b) contain information regarding the case definition, and c) define the public health measures that need to be determined.

Q. How often should one monitor internal validity?

A. If a monitoring system is in place, monitoring is a continuous process.

Q. Who should be responsible for monitoring internal validity?

A. Monitoring is usually carried out by local public health departments (district or regional level) which helps to ensure that all public health measures are appropriate for the local situation. Data are sent to the central level once they have been cleansed.

Q. At what level should internal validity be monitored?

A. The most appropriate monitoring level is the intermediate level, i.e. the local public health authorities is. In some countries, this is not feasible (lack of personnel, lack of expertise or resources) and the monitoring system is maintained at the central level (e.g. the national public health authority).

Q. Is monitoring of validity necessary for a system?

A. Absolutely. Monitoring is necessary to avoid false alarms and unnecessary public health interventions. In addition, it is important to have valid surveillance data that correspond to the actual epidemiological situation in order to design appropriate and effective public health strategies.

Q. What is the most important consideration when designing a system to ensure validity?

A. Create reporting forms and databases as user-friendly, self-explanatory and simple as possible.

Q. What are the major factors you should take into account when suggesting changes to the surveillance system in order to increase its validity?

A. You should take into account:

- the results of the validity measurement;
- the factors that were identified as critical;
- the available resources;
- the characteristics of the current system;
- the methods/checks that are already in place.

Q. What are the three basic measures/strategies for enhancing internal validity?

A. Training of data providers, modifying reporting forms, and applying validity checks during data entry.

Q. What makes a reporting form efficient?

A. A reporting form should be simple, brief (only one page, if possible), clear and concise, self-explanatory, specific, logically structured, and not too demanding.

2.4 Monitoring timeliness

Dimensions of data quality, such as timeliness, depend on the purpose and scope of a surveillance system. There will be some 'high value' data items where timeliness is more critical than for others. (Similarly, there may be other items where completeness is more critical.) It is also important to remember that the value of any particular data item in a multi-user system may be different for different users.

The general approach for monitoring data quality, as described in Chapter 1, also applies to timeliness. A specific approach to monitor timeliness is presented in Figure 2.4a, Steps 1 through 3.

Figure 2.4a. Steps for monitoring timeliness

Step 1: Describe the system

Step 2: Plan the monitoring process

Step 3: Measure timeliness and summarise results

Step 4: Interpret the results of the measurement

Step 5: Suggestions for remediation.

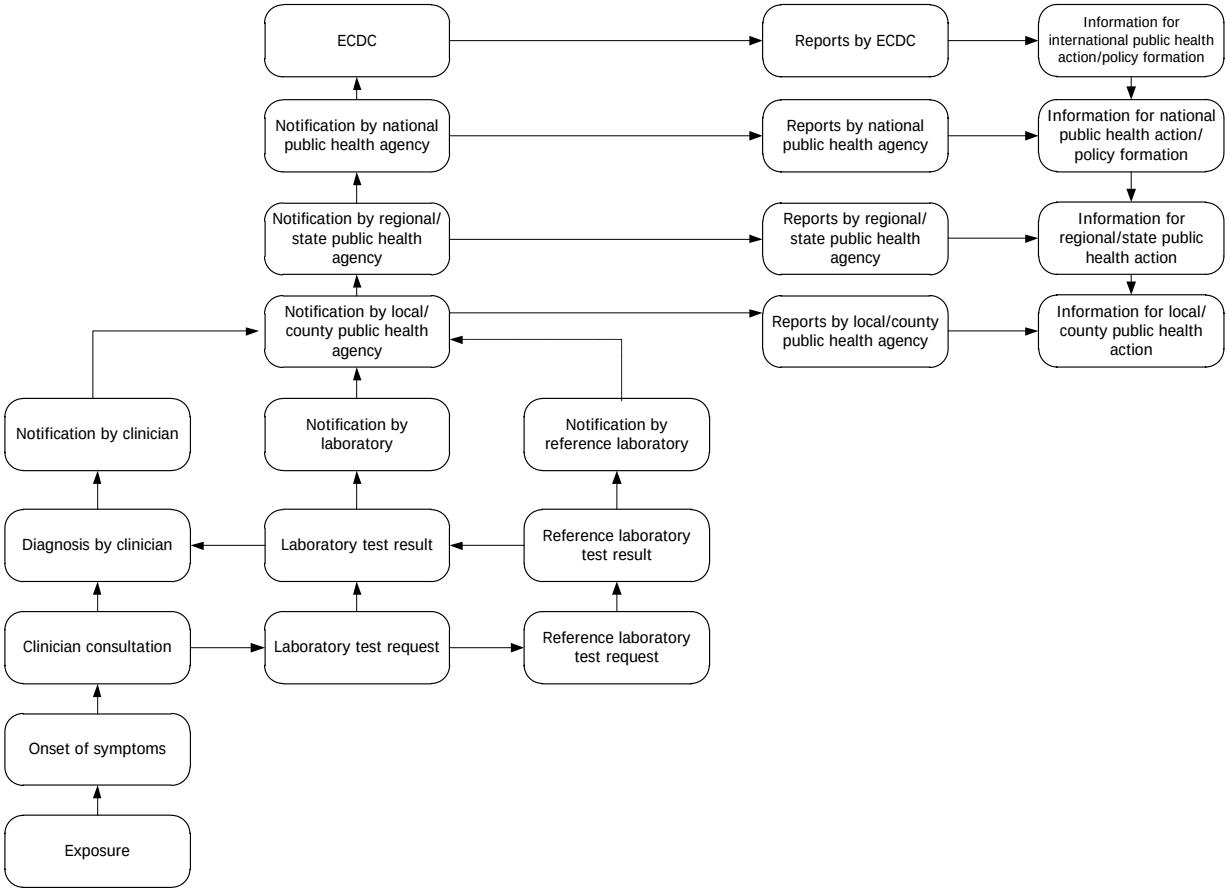
2.4.1 Describe the system

The purpose of an infectious disease surveillance/reporting system can be described by looking at its scope. The scope can vary substantially from system to system. Figure 2.4b presents an overview of the elements of an integrated surveillance system – including clinical and laboratory notifications – that can be used to measure timeliness and other attributes.

Surveillance systems with an extensive scope may include additional elements such as notifications from systems for disease-specific clinical management, tuberculosis, sexually transmitted infections, and infection

control/healthcare-acquired infection information. The scope of a disease surveillance system can also include support for local or regional case management, including preventative or prophylactic measures, corrections, and contact tracing.

Figure 2.4b. Scope of a reporting system to measure timeliness



2.4.2 Plan the monitoring process

Before data quality attributes on timeliness can be identified, it is necessary to identify the type information that is communicated through the infectious disease surveillance system.

Defining a list of data items

It is important to make sure that data items and the values they can hold are clearly understood. It is equally important that the information retrievable through report mechanisms is clearly understood and not misinterpreted.

Each data item in the information system needs to be assigned a value in relation to the scope and purpose of the system. It is important to realise that different users may rate the importance of data items differently and assign different values to it.

Equally, each data item incurs a certain cost, which is worth paying for if the item has been assigned a high value; if the value of the information does not warrant this expenditure, the data item should be dropped. The cost of a data item is not only defined by the effort and expense of collecting the related information but also by legal data protection requirements. This mechanism ensures that only data that are relevant and not excessive are collected.

The value of a given data item will determine the usefulness of monitoring or evaluating data quality attributes, for example completeness, validity, accuracy, and timeliness associated with that data item. It is important that the value of a data item determines whether a data quality attribute is monitored or evaluated, and not the ease with which it can be monitored or evaluated, i.e. the monitoring or evaluation process has to be useful and meaningful, and not opportunistic.

The below list is an example of data items for which timeliness has a high importance:

- Disease: to identify the disease being notified
- Surname: to identify the case
- County: to locate the case, for descriptive epidemiology

- Health board/region: to locate the case, for descriptive epidemiology
- Date of notification: for outbreak management, for descriptive epidemiology

For some data items, the level of importance of timeliness can be specified as high/medium and low. It is possible to recommend a maximum time span (number of days) to address specific objectives and public health actions. However, if these objectives and public health actions have target times, then the data timeliness has to support this.

Questions to be asked

The following questions (Figure 2.4c) should be asked when determining the value and cost associated with data items and who is responsible for the accrued cost.

Figure 2.4c. Steps for monitoring timeliness

- Is timeliness included in the scope of your surveillance system?
- Do you have a timescale for clinician reporting?
- Do you have a timescale for laboratory reporting?
- Do you have service level agreements (SLA) with regional public health departments which include information on the frequency or timeliness of notification to the national agency and/or other infectious disease information partners?
- Do you have SLAs with laboratories which include information on the frequency/timeliness of notification to public health officials?
- Do you have legal responsibilities or SLAs which stipulate the production of reports that will be distributed to infectious disease information partners?

2.4.3 Measure timeliness

The 'timeliness' of a surveillance system is defined as the time needed to communicate notifiable information between system partners who are required to receive this information.

This timeliness is determined by the time between the various steps in the information chain. The timeliness required for different infections or different data items will vary depending on the purpose for which the information is required. If information is required for immediate public health action in relation to priority diseases it is important that these reports are communicated as a matter of urgency. If reports are required for trend analyses (with a lower public health priority) or do not require rapid public health action, timeliness becomes less relevant.

For example, a case of meningitis or meningococcal disease in nursery, pre-school or school settings will typically require very prompt communication between clinicians, public health staff, and clinical and reference laboratory staff at the local level. Critical data items include contact information such the names of the affected individuals, their home addresses, and details on educational setting, clinical diagnosis, and laboratory investigations.

On the other hand, information required for a case of genital herpes simplex may be limited to core demographics.

Figure 2.4d illustrates the most important time delays to be considered for monitoring timeliness

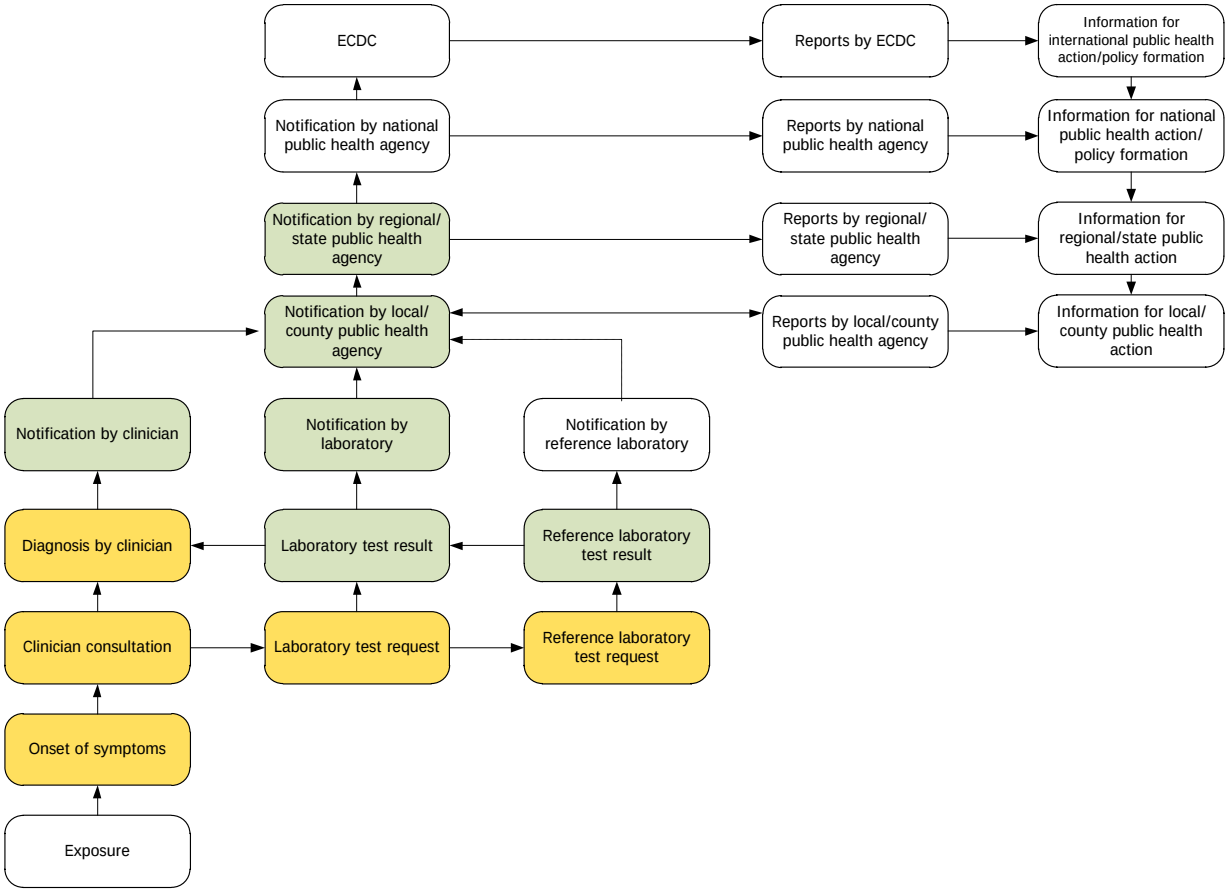
Figure 2.4d. Monitoring timeliness: main time delays

- Delay to onset = date of onset – date of infection
- Delay to consultation = date of clinician consultation – date of onset
- Delay of laboratory diagnosis = date of laboratory test result – date of laboratory test request
- Delay of diagnosis (clinical and laboratory) = date of laboratory test result – date of clinician consultation
- Delay to notification = date of notification (to local health authorities) – date of diagnosis (by the physician or the laboratory)

To correctly calculate time delays, the data quality of the recorded dates in the system is also important: errors to be avoided include date format errors and rounding errors.

Figure 2.4e is an example of important timeliness elements in the Computerised Infectious Disease Reporting (CIDR) system in Ireland. Input data are in orange, output data in green.

Figure 2.4e. Important timeliness elements in the Irish CIDR system



Relevant and measurable in the above example is the time taken from clinical notification until the time the case was entered into the surveillance system (by local or regional public health officials). Once the case is cleared – again at the local or regional level – the information is available for all authorised users of the online system. In order for the information to be distributed further, periodic file transfers from local/regional systems to a central national/federal system may be necessary.

Other timeliness elements may only be important for some diseases or in outbreak situations; these include the length of time from the onset of symptoms or from the date of diagnosis until the date of notification. This information may not be routinely recorded, except for during outbreaks or pandemics. For example, an increase of timeliness was observed during the Shiga toxin-producing *Escherichia coli* outbreak reporting system in Germany.

With regard to laboratory notifications, it is important to measure the time it takes a laboratory to notify the local/regional public health authorities, and then how long it takes the local/regional public health authorities to move a notification to the national level so that it can be viewed by all users of the system.

Other elements of timeliness may be important but not always be available or recorded. For example, the variable 'specimen date' may be a proxy/equivalent for the clinician consultation date. If the 'lab reporting date' – ideally recorded in the laboratory information management system – is not available, it may be functionally replaced by the variable 'date authorised' (by the laboratory) for viewing by public health.

In the Figure 2.4e the timeliness elements between the green boxes can be easily measured, which allows for excellent monitoring of the system's timeliness.

The timeliness elements between the orange boxes may be less readily measured as some of the information required may not be readily available.

2.5 References

1. World Health Organization. Communicable disease surveillance and response systems. A guide to monitoring and evaluating. Geneva: WHO; 2006.
2. Lescano AG, Larasati RP, Sedyaningsih ER, Bounlu K, Araujo-Castillo RV, Munayco-Escate CV, et al. BMC Proc. 2008 Nov 14;2 Suppl 3:S7.
3. [No authors listed] Overview of the WHO framework for monitoring and evaluating surveillance and response systems for communicable diseases. Wkly Epidemiol Rec. 2004 Sep 3;79(36):322-6. Available from: http://www.who.int/wer/2004/en/wer7936.pdf
4. Chrisman NR. The role of quality information in the long-term functioning of a GIS. Proceedings of AutoCarto 2006 06;2:303-321.
5. Checkland P, Poulter J. Learning for Action: A short definitive account of Soft Systems Methodology and its use for practitioners, teachers and students. New York: John Wiley & Sons; 2006.
6. Kroenke DM. Experiencing MIS. Englewood Cliffs, New Jersey: Prentice Hall; 2008.
7. Sekhobo J, Druschel C. An evaluation of congenital malformations surveillance in New York State: an application of Centers for Disease Control and Prevention (CDC) guidelines for evaluating surveillance systems. Public Health Rep. 2001 Jul-Aug;116(4):296-305
8. Pillaye J, Clarke A. An evaluation of completeness of tuberculosis notification in the United Kingdom. BMC Public Health. 2003 Oct 6;3:31.
9. Howell, DC. The analysis of missing data. In: Outhwaite W, Turner S, editors. Handbook of social science methodology. London: Sage; 2007.
10. Allison PD. Missing data. Thousand Oaks, CA: Sage Publications; 2001.
11. Schafer JL, Olsden MK (1998). Multiple imputations for multivariate missing-data problems: a data analyst's perspective. Multivar Behav Res 1998;33(4): 545–571.
12. Barladi AN, Enders CK. An introduction to modern missing data analyses. J Sch Psychol, 2010;48, 5-37.
13. Crawford SL, Johannes CB, Stellato RK. Assessment of digit preference in self-reported year at menopause: choice of an appropriate reference distribution. Am J Epidemiol 2002; 156: 676–83.
14. Camarda CG, Eilers PC, Gampe J. Modelling general patterns of digit preference. Statistical Modelling 2008; 8(4): 385-401.
15. Buckeridge DL. Outbreak detection through automated surveillance: a review of the determinants of detection. J Biomed Inform 2007; 40(4): 370-9.
16. Lescano AG, Larasati RP, Sedyaningsih ER, Bounlu K, Araujo-Castillo RV, et al. Statistical analyses in disease surveillance systems. BMC Proceedings 2008; 2(Suppl 3):S7
17. Tague NR. Seven basic quality tools – the quality toolbox. Milwaukee: American Society for Quality; 2004.
18. Ishikawa K. What is total quality control? The Japanese Way. Englewood Cliffs, New Jersey: Prentice-Hall; 1985.
19. Ridout MS, Morgan BJT. Modelling digit preference in fecundability studies. Biometrics 1991; 47: 1423–33.
20. Pickering R. Digit preference in estimated gestational age. Stat Med. 1992;11:1225–38.
21. Montgomery D. Introduction to statistical quality control. Hoboken, New Jersey: John Wiley & Sons; 2005.
22. Altmann M, Spode A, Altmann D, Wadl M, Benzler J, Eckmanns T, et al. Timeliness of surveillance during outbreak of Shiga toxin-producing *Escherichia coli* infection, Germany, 2011. Emerg Infect Dis. 2011 Oct;17(10):1906-9)
23. Centre for Disease Control and Prevention. Updated guidelines for evaluating public health surveillance systems. Recommendations from the Guidelines Working Group. MMWR July 27, 2001/50 (RR13);1-35

3 Evaluating surveillance systems

3.1 A stepwise approach to evaluation

When monitoring data quality, the goal is to continuously monitor few attributes to rapidly detect problems and take corrective measures; surveillance systems evaluations perform comprehensive measurements of relevant attributes (over time or in a single effort) in order to verify whether the system is still 'fit for purpose' and recommend improvements. Evaluating surveillance systems can be very challenging and can become very complex because of the growing use of electronic data, the linking of multiple sources, the extensive use of laboratory data, and financial constraints.

Each type of surveillance system will require a tailored evaluation effort (for example, by focusing on specific attributes) based on the objectives of the surveillance system.

The first step consists of describing the system, taking into account the different elements of the systems (Section 2.2; Figure 1.3a).

The second step, described in Section 3.2, consists of a detailed plan, including an overview of possible triggers for running an evaluation through the use of a decision tree (Table 3.2c.). The planning also makes provisions for the full engagement of the team and participating stakeholders.

Step 3 covers the implementation phase of the evaluation and includes two main components: the evaluation of performance and cost. Methods/approaches for evaluating the performance of surveillance systems by attributes are discussed in Sections 3.3 through 3.7, and the evaluation of cost is described in Section 3.8. Dealing with evaluations of each attribute is a complex task; this subsection of the chapter covers several attributes in detail and gives examples in order to help with the selection of an evaluation strategy.

Finally, step 4, offers recommendations on how to address the results of the evaluation. Recommendations immediately follow Sections 3.3 through 3.7.

Figure 3.1a. Evaluating surveillance systems

Step 1. Describe the system

Step 2. Plan the evaluation

Step 3. Perform the evaluation

Step 4. Propose an action plan

Persons who evaluate surveillance systems should always remember, however, that an evaluation is not complete until its results have been disseminated to the appropriate stakeholders of the system, and the stakeholders have responded to the results.

3.2 Plan the evaluation

The evaluation of a disease surveillance system is a complex task and requires careful planning and preparation.

3.2.1 Identifying triggers

Determining the most appropriate point in time for a system evaluation may seem a daunting task. Furthermore, narrowing down or focusing on particular problems of the evaluation process is a multi-step process which requires a sound plan.

Description of a trigger-oriented approach

What triggers an evaluation and what are the characteristics and operational peculiarities of the surveillance system in question?

There are various types of surveillance systems, and each will require different questions when planning an evaluation. For example, systems that are active and population-based may bring up the following questions:

- Does the system accurately measure incidence?
- Does it provide unbiased estimates of risk factors?
- Is it cost effective?

Passive, population-based surveillance systems may bring up the following questions:

- Does the system adhere to the principles of simplicity and acceptability?
- Does it lead to concrete public health actions?
- Does it provide good quality data for trend analyses?

Syndromic surveillance systems may warrant questions about the ability to detect outbreaks:

- Is disease detection timely enough?
- Are triggered actions cost-effective?

Triggers which can prompt a system evaluation are either external (outside of the system) or internal (from within the system) and usually related to data providers, changes in the structure of the system, or changes in organisations which use system outputs to improve public health.

Examples of external triggers include the emergence of new infections; changes in diagnostic techniques; concerns about the acceptability or representativeness of reporting from clinicians; failure of the system to detect an outbreak; delayed action for case management; and incapacity of the system to properly quantify a public health problem. Other external triggers which could lead to an evaluation are declining resources, or a requirement to link to other surveillance systems.

Although it is important to improve the internal functioning of surveillance systems through evaluation and quality improvement, external triggers should be paramount when determining the when and how of a surveillance system evaluation. An imperfect system that leads to effective public health action is better than a perfectly functioning system that is not being used to improve the public's health.

Table 3.2a. Examples of triggers leading to an evaluation

Trigger	Type of trigger	Evaluation response	Action
Missed outbreak	External (usefulness)	Evaluate completeness	Widen the scope of surveillance; add monitoring system for completeness
Increased disease importance and concomitant use of health services	External	Revisit aims of surveillance system and evaluate system to see if fit for purpose of indicating impact on healthcare services; also completeness	Ensure that data automatically reaches health policy decision-makers.
Errors in surveillance report used for policy (due to poor data quality)	Internal	Identify specific areas of error (e.g. validity)	Encourage reporting of valid data and incorporate continuous monitoring
Misrepresentation of disease measures in country comparisons across the EU due to incomplete or incorrect reporting to TESSy	External	Evaluate case definitions and data quality with respect to communication to TESSy	Add compatibility with TESSy and continuous quality measures

Summary of a literature review on triggers

The global consulting firm Public Health Practice conducted a systematic literature review on what situations could trigger a surveillance system evaluation and how such triggers would influence the choice of the evaluation method. The main research question asked for triggers which led to evaluations of public health surveillance systems between 1992 and 2012. To support the research question and to ensure the collection of all relevant data, several secondary research questions were developed:

- What were the purposes of the evaluations?
- What triggers and trigger categories were derived from those purposes?
- Do those triggers vary according to geography, time of evaluation, and type of surveillance?
- What guidelines were used during the evaluation process and how often were they implemented?
- If no guidelines were used, what evaluation methods and evaluation system attributes were used?

Defining the triggers

After step 1 (description of the surveillance system), step 2 consists of describing the triggers which can prompt a system evaluation. The description of triggers also initiates the planning process by linking triggers to the most appropriate type of evaluation model. Triggers are described and categorised in Table 3.2a, and further references are listed at the end of this chapter. The ability to identify what triggers an evaluation is important because the evaluation objectives and methods should be tailored according to the relevant trigger.

Table 3.2b. Triggers for the evaluation of surveillance systems: suggested definitions and categories

Triggers	Category	Underlying questions	Relevant references
Initial evaluation	General	First evaluation of the surveillance system	1–4
Compare systems	General	Comparison of different types of surveillance systems for the same disease to evaluate strengths and weaknesses of the current system	5–6
Regular periodic evaluation	General	Systematic evaluation is part of the surveillance process. A surveillance system owner may wish to regularly perform an evaluation to ensure the system is still fit for purpose.	7–9
Requirements	General	A public health institute is required to conduct a fundamental review of the surveillance system.	
Cost of surveillance	Economic changes	An evaluation of cost can lead to the prioritisation of surveillance components while other components will be scaled back (e.g. downsizing of a department; overall cost reductions). What is the cost of surveillance?	10–12
Data quality monitoring	Technical	Results of data quality monitoring indicate that there may be a problem with some aspects of data quality that needs to be investigated further.	13–16
New technology or other system innovations	Technical	New software or other technical innovations that require evaluation for effectiveness. Examples are the implementation of Health Level Seven International standards (see: http://www.hl7.org/index.cfm) or a new Windows service pack.	17–20
Change in demography	Technical	Investigate the representativeness of the surveillance system.	
New surveillance standards	New component	In an already existing surveillance system, standards should have been established (case definitions, surveillance objectives, list of diseases). Some countries can establish these standards through a regulatory process. When a new surveillance standard is implemented, public health managers need to know if it is well-accepted, useful, and relevant in terms of public health.	21–24
Change in case definitions	New component	Cases of communicable diseases reported to a surveillance system should be defined according to standard case definitions. A revised case definition will have an impact on sensitivity and specificity of surveillance.	25
Evaluating syndromic surveillance or switching to syndromic surveillance	New component	Switching to syndromic surveillance for some diseases can be less costly than indicator-based surveillance. Is the current syndromic surveillance system integrated with the regular public health surveillance system? Is the syndromic surveillance system useful for threat detection? Can it assess trends?	26
Mass gathering	Emergency, new component	Can the current surveillance system handle surveillance of a mass gathering event? If not, what components should be added? Surveillance should be evaluated during and after a mass gathering.	
Occurrence of a major/new public health event	Emergency	Can the current surveillance system handle a major public health event (e.g. an influenza pandemic)? The evaluation should focus on the early warning and response function of the surveillance system. An evaluation of a surveillance system for a new condition may also focus on other aspects (identification of risk factors), effectiveness of interventions, efficiency of a newly introduced surveillance system, and other relevant attributes).	
Preparedness	Emergency	Is the surveillance system prepared to cope with unexpected events in terms of workload and new information to be collected? For example, would an outpatient influenza surveillance system be able to cope during a severe pandemic? Would this system be able to provide information on risk factors for severity? To verify preparedness, one should identify the surveillance objectives for a predefined event and evaluate how efficiently the system meets these objectives.	
Change in public health policy	Public health measure	Ability of the surveillance system to measure the impact of changes in public health policy on the concerned population.	27–29
Introduction of control measures	Public health measure	Contact tracing for tuberculosis cases; measles (eradication): Is the surveillance system able to measure the impact of such interventions? Does the system provide accurate information for the purposes of programme monitoring?	

Triggers	Category	Underlying questions	Relevant references
Structural changes	Public health measure	Example: Change of healthcare structure, from secondary (mainly hospital healthcare) to primary care (general practitioners and other community-care service providers), in the United Kingdom.	

3.2.2 Selecting the evaluation type and method

Choosing an appropriate evaluation model and method can be facilitated by using a selection grid or decision tree.

Choosing the appropriate evaluation model

Choosing an evaluation model depends on three factors: frequency of evaluation, the type of protocol (light, medium, full), and cost. In Figure 3.2a, the y axis represents the frequency of evaluation and the x axis stands for the intensity of evaluation. Monitoring data quality is a continuous process, although the system can also be evaluated at different points in time (on a regular basis or ad hoc).

The degree of complexity of an evaluation depends on what triggered the evaluation as well as other parameters, for example the number of diseases to be included, the attributes to be evaluated, the number of stakeholders, and the coverage of the system. Indications for the three protocol types are described in Table 3.2b.

Figure 3.2a. Selecting an appropriate evaluation type

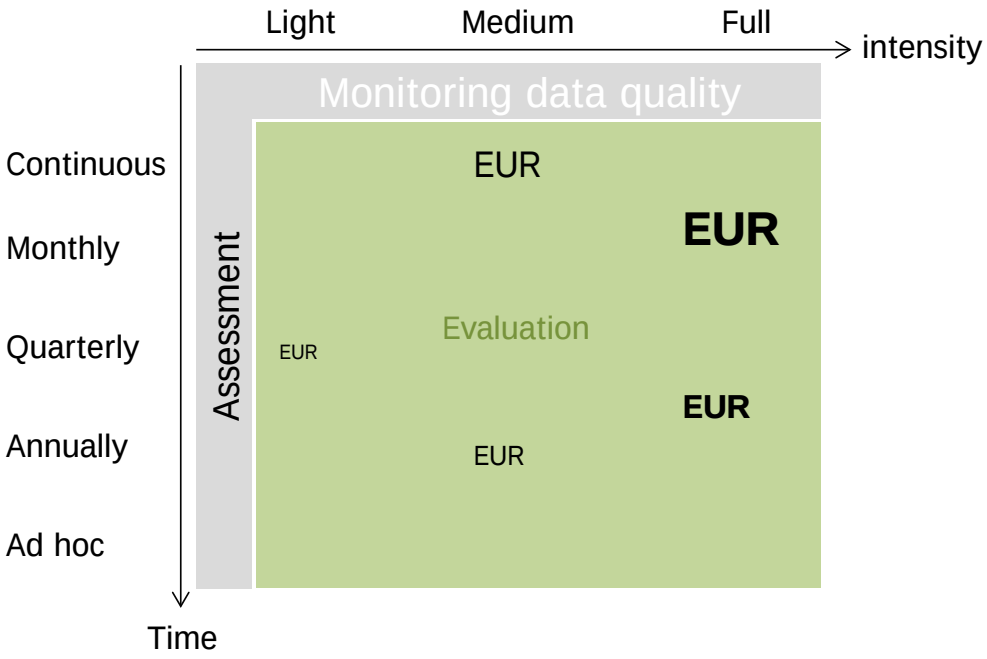


Table 3.2b. Types of protocols

Protocol type	Attributes (number and types)	Required resources (human, financial)	Stakeholders	Time required
Light	Limited number of attributes to be evaluated; used for data quality evaluation of internal completeness and/or internal validity rather than external data quality. Example: evaluation of a single attribute or a specific issue.	Limited budget	Limited number of stakeholders to be consulted.	Limited amount of time to set up and run the evaluation.
Medium	Depends on objectives; can include several attributes.	Substantial budget could be required, depending on the final evaluation scope.	May involve a large number of stakeholders to be consulted.	Time can be a limiting factor.
Full	Includes most or all attributes of a surveillance system. Can include external completeness/validity evaluation (e.g. performing ad hoc studies to assess certain attributes).	Substantial budget required to do a full evaluation.	May involve a large number of stakeholders to be consulted.	Substantial time is required.

Decision tree approach

An evaluation can require a complex decision process based on many parameters. Table 3.2c lists a number of (simplified) triggers, along with the attributes to be evaluated.

The decision on which attributes to evaluate and which type of protocol to use depends on several factors, including the disease, the objectives of the surveillance system, the type of surveillance system, etc. Some attributes can be mutually exclusive, e.g. a system is rarely sensitive and specific at the same time.

Table 3.2c shows examples of attributes to be evaluated. Attributes are listed with their corresponding triggers and the type of evaluation connected to these triggers. These examples can be used to develop a decision tree which helps to guide the decision for the evaluation and a related size.

By identifying similar categories and triggers, the decision tree helps to determine the most appropriate model for an evaluation and indicates whether a light, medium or full evaluation protocol is needed. It should be noted that this table is a simplification of a complex decision process, and it is only provided only as a general guidance.

Table 3.2c. Examples of attributes to be evaluated and possible type of evaluation for specific triggers

Trigger category	Trigger	Description, examples	Attributes to be evaluated, analysis to be performed	Type of protocol
General	Initial evaluation	Check if objectives can be achieved	Simplicity, flexibility, usefulness	Medium
	System comparison	Study underreporting by comparing mandatory notification and laboratory reporting	External completeness/validity, sensitivity	Medium
	Regular periodic evaluation	Evaluation of data quality component on a regular basis, or disease-specific evaluation	Can be targeted at all or selected attributes	Light, medium or large
	Requirement	A public health organisation may wish to conduct a fundamental review of the surveillance system to learn more about the data being collected by the system or the surveillance system itself.	All attributes	Full
Economic	Cost of surveillance	Comparison of performances of the surveillance system, taking into account the cost.	Usefulness, cost	Medium to large
Technical	Sudden increase in disease incidence	Is this increase real or due to reporting artefacts?	Completeness, validity, sensitivity, specificity	Medium
	Data quality monitoring	Example 1: Decrease in the number of specimen reported by laboratories for *Salmonella* detection: is this a real decrease of incidence, or simply underreporting? Example 2: The variable related to TB culture confirmation is only 70% complete: is this a problem of laboratory capacity?	Limited to data quality concern: completeness (internal/external), sensitivity, validity	Medium
	New IT technology	Implementation of HL7 messaging	Completeness, validity and timeliness	Light, IT driven
	New IT technology	Windows Service Pack causes system to crash	Timeliness	Light, IT driven
	Change in demographics	Increased case numbers during holiday season	Representativeness	Light
New surveillance component	New surveillance standards	New disease added to the list under surveillance because of public health concern (e.g. tick-borne encephalitis)	Acceptability, usefulness, public health importance	Medium
	New surveillance standards	New case definition	Completeness, validity, sensitivity	Medium
	Syndromic surveillance	Does the current syndromic surveillance system support the early warning function?	Timeliness	Medium

Trigger category	Trigger	Description, examples	Attributes to be evaluated, analysis to be performed	Type of protocol
	Mass gathering	Enhance surveillance during the Olympics or the Soccer World Cup	All attributes	Full
Public health emergency	Public health	Timely detection of an increase in a new VTEC strain	Measure delay in the system; timeliness, specificity and sensitivity, flexibility	Medium
	Preparedness	Influenza preparedness plan	flexibility	Full
Public health measure	Change in public health policy	Introduction of a new vaccination policy: evaluation of an intervention, before and after the introduction of the new vaccine.	Analysis of the objectives of the intervention: burden of disease (incidence/prevalence), effectiveness, efficacy, safety, and vaccination programme's impact on the disease.	Medium to large, disease specific
	Control measure	Example 1: Quarantine measure in case of an epidemic (e.g. 2009 influenza pandemic). Example 2: Contact tracing for cases of tuberculosis or measles (eradication)	Usefulness, public health importance Sensitivity, completeness, validity	Light or medium Medium

3.2.3 Engaging stakeholders

Evaluations usually involve several stakeholders. A stakeholder is defined as anyone who has a responsibility for the system, is in involved in the operation of the system, participates in the evaluation, or has some other type of interest in the system. Stakeholders can be heads of infectious disease departments, epidemiologists, data providers, analysts, administrators, data reporters, and can include institutions or organisations, for example a ministry of health.

Initial discussions should involve as many stakeholders as possible in order to cover a large number of viewpoints and interests. This is important for motivational purposes and also ensures that the evaluation effort remains transparent. While it may not be possible to involve every stakeholder, it could be helpful to meet with as many as possible to gain feedback.

It is important to make a clear distinction regarding the individual roles of each stakeholder. These could include the following:

- Evaluation lead. Leads the effort, coordinates planning and helps to strategise the completion of the evaluation. Will possibly manage staff who will be involved in conducting the effort.
- Evaluation staff. Scientists who will be conducting various aspects of the evaluation effort.
- Staff epidemiologists or scientists. The people who are at the frontline of the surveillance operation, interacting with datasets, regularly working on specific systems and serving in capacities that make their buy-in to the evaluation process essential.
- Unit heads or other leadership. These people often have a stake in the evaluation process, and as decision-makers they may help to identify specific concerns or key questions throughout the evaluation process. Furthermore, they may be interested in the results of the evaluation because they offer important information for decision-making and the prioritisation of efforts.

Since most evaluation activities will be limited to a single unit or involve a relatively small number of decision-makers, an evaluation project proposal may help to create commitment to the evaluation procedure and foster buy-in across all stakeholders. Such a proposal could be something as simple as a letter of intent, or a more thorough proposal listing such factors as those in described this document, and would provide a transparent overview of strategy, plans and expectations throughout the whole evaluation process.

3.2.4 Setting up a team for surveillance evaluation

The composition of the evaluation team should be related to the triggers and objectives of the evaluation. If possible, a team should be appointed and meet as often as possible to discuss all aspects of the evaluation effort (planning, implementation, results and recommendations). This helps to ensure the endorsement of recommendations, identify strengths and limitations, provide a clear rationale for recommendations, and promote a follow-up action plan. Integrating the team throughout the process will improve the chances that the evaluation will be completed as planned, and will also assist in later efforts to implement recommendations based on evaluation results. In short, the team approach maximises the likelihood that changes occur by creating buy-in across multiple interests.

The evaluation can be coordinated as a team effort. The composition and training of the evaluation team should use an interdisciplinary approach and integrate different job roles (e.g. evaluation managers, epidemiologists, data

specialists, and various stakeholder, where appropriate). This will ensure that epidemiologists with or without disease-specific expertise (depending on the surveillance system to be evaluated) can participate.

Training should be provided as necessary, particularly if surveillance specialists and other epidemiologists are not familiar with evaluation factors such as triggers and the various evaluation models. In addition to fully disclosing the evaluation strategy and involving everyone in the planning process who has a role in the evaluation, other types of training may be needed, not only for people who are assigned to the planning process, but also for those who are involved in the evaluation of specific attributes. For example, background reading and various examples for evaluating usefulness or representativeness of specific systems may be needed before a system evaluation can be undertaken.

3.3 Evaluation of external completeness

Key points on the evaluation of external completeness are listed in Figure 3.3a.

Figure 3.3a. Key points for the evaluation of external completeness

- Data must be complete in order to meet the required quality level, i.e. a surveillance system which is reliable and useful.
- Before evaluating a surveillance system for completeness (or any other attribute), a detailed protocol should be written; this protocol should specify all methods to be used.
- The objectives of the evaluation should be action-oriented. The evaluation team should be able to suggest correction methods and corrective approaches for increasing completeness at the end of the evaluation.
- The choice of an external data source depends on the specific conditions under which the evaluation is conducted.
- Underreporting can be significantly reduced by convincing physicians that all data collected by the public health authorities will be used to make informed decisions to promote public health.
- An evaluation of external completeness can be planned for diseases requiring a high level of completeness or in situations that have triggered an evaluation.
- Possible methods for evaluation of external completeness are the review of a sample of all surveillance data, a review of medical records, and capture–recapture studies.

3.3.1 Introduction

The process of evaluating data quality and implementing data quality assurance schemes can be summarised by a plan–do–check–act cycle [1]. Completeness of a system defines to an important degree the quality of the collected data. The terms 'external completeness' or 'completeness of reporting' refers to the absence of underreporting of diseases to surveillance systems. It expresses the ability of the surveillance system to capture the cases of disease that actually occur in the monitored population. External completeness requires that the surveillance system be inclusive of the whole population being monitored and that the reported cases represent the complete list of cases diagnosed and not just a fraction of them (Chapter 2). As specified in Section 2.1, external completeness refers to the case reporting process only, whereas sensitivity applies to both processes, case reporting and under-ascertainment (Figure 1.2a.).

Even the most efficient surveillance systems are affected by some underreporting. It is therefore important to quantify the degree of underreporting in order to:

- estimate the true burden of the disease in the population;
- take appropriate measures to improve the surveillance system; and
- set priorities regarding public health measures.

Studies have shown that reporting completeness varies greatly among diseases, surveillance systems, countries or even geographic regions in one country [2-5]. Often, only a low proportion of cases are reported for diseases that require immediate public health intervention. An additional consequence of a low reporting rate is a delay in implementing control measures and the delayed identification of disease clusters/outbreaks.

External completeness should be evaluated regularly to ensure that systems meet their intended objectives [6-8]. It is difficult, however, for surveillance systems to capture all cases of a disease, and underreporting is always assumed [9].

The different steps of performing an evaluation of a surveillance system described in Chapter 1 also apply to the evaluation of external completeness.

3.3.2 Describing the system

Among the key elements specified in Section 1.3, the following items are also relevant for evaluating a system's external completeness:

- Legislation (national/international levels)
- Main stakeholders
- Objectives
- Type of system
- Population under surveillance
- Time of data collection
- Data sources/data providers
- Type of data collected
- Reporting process, data flow
- Data management (e.g. data entry, editing, storage, back up, and transfer of data)
- Resources needed (types, amount, and timing)
- Data analysis and dissemination of results
- Patient privacy, data confidentiality, and system security
- Existing alternate data sources (external to the system) which could be used for comparisons

Using the data available in the surveillance system, the available resources, and the resources needed for evaluation (e.g. person-time, money, technical support), one can determine whether a system evaluation should be conducted, and if so, to what extent (light/medium/full intensity). Decisions should be based on a cost–benefit approach. When it is patently clear that completeness of the surveillance system is low, corrective measures can be taken without prior evaluation. However, it is also important to take a baseline measurement in order to later be able to assess the effects of the corrections/modifications made in the system. Economic factors should also be taken into account when deciding whether to continue or discontinue surveillance activities. The results of a gap analysis can provide a good basis for recommending changes without increasing costs (see Section 2.2).

The best sources of information about a surveillance system are most likely the persons regularly involved with the system (data providers and public health workers at the local, intermediate or central level), as the theoretical characteristics of a system often change when put into practice. Focused meetings with key stakeholders who are involved in surveillance at all levels of the reporting system can be useful for assessing the functionality of the system, the applied procedures, and the weak/strong links in the information flow.

3.3.3 Planning the evaluation

Selecting diseases

It is possible to measure external completeness for all diseases included in a surveillance system or decide to evaluate completeness for only a specific disease. It is important to distinguish between diseases for which a very high level of external completeness (or sensitivity of the system to rapidly detect and report single cases) is needed and situations (under-ascertainment) which require a general evaluation of external completeness, regardless of the nature of the disease (Table 3.3a).

Table 3.3a. Important information to be included in the evaluation protocol

Diseases requiring a high level of external completeness	Situations requiring an evaluation of external completeness
Highly contagious diseases leading to possible outbreaks (e.g. new influenza strains which can cause large outbreaks or even a pandemic). Public health measures will be taken upon detection of single cases (e.g. meningococcal disease and prophylaxis).	High incidence of a disease or recent increases of its occurrence but real reasons remain unclear (e.g. increased incidence of tuberculosis; sexually transmitted disease among refugees)
High case-fatality ratio associated with a disease (e.g. cholera)	Increased cost associated with increased hospitalisation rates or higher use of medication (e.g. increased cases of West Nile Virus infection that result in increased use of intensive-care hospital units)
Rare diseases (e.g. haemorrhagic fevers of various aetiologies)	A change in public health intervention/protocol (e.g. availability of a new vaccine for meningococcal disease)
Emerging or re-emerging diseases (e.g. malaria)	A change in diagnostic procedures (e.g. introduction of rapid tests for the detection of *Legionella* antigen in urine)
New strains/new antimicrobial resistance patterns (e.g. co-infection with multidrug-resistant tuberculosis and human immunodeficiency virus [HIV])	Cross-border threats (e.g. Q-fever and other zoonoses)
Diseases in the process of being eradicated or controlled (e.g. measles, brucellosis in Mediterranean countries).	Increase in the level of public/political interest (e.g. influenza leads to increased concern in the community)

Diseases requiring a high level of external completeness	Situations requiring an evaluation of external completeness
A previous intervention to increase completeness of reporting for a specific disease which has to be evaluated. In this situation, the current completeness of the system will be compared to the targets set when the intervention was designed.	Low completeness found by a previous evaluation or studies at the local level (compared to the anticipated/acceptable completeness).
	A general belief that a specific disease is underreported.

Regardless of the diseases that are selected for inclusion, the main objective of evaluation remains the same, i.e. estimating the degree of underreporting of diseases in the community in order to improve surveillance.

The number of diseases to be evaluated should not exceed the resources of the surveillance system. Some diseases may not be as important to be included as others, and prioritisation is essential. A written protocol should explain the rationale behind the selection of included diseases and state all reasons for the evaluation of completeness.

In practice, an evaluation of external completeness for all diseases is very unlikely. Underreporting usually has a variety of reasons. Depending on the disease, data sources will differ substantially and have to be assessed specifically for every disease.

Deciding on an acceptable level of underreporting by disease

Ideally, all cases of a reportable health condition should be reported, and recording of the input should be as complete as possible. In completely electronic systems this is easier to achieve than in manual ones, but one can always expect some number of missed cases because of the complexity of healthcare systems. In developing countries, there can be additional challenges, often due to the lack of resources and a degree of instability.

Below is a list of the basic factors that should be taken into consideration when determining the acceptable level of underreporting of a specific disease in a surveillance system:

- The objectives of the system (fitness for purpose). For example, a degree of underreporting is acceptable in a system that tries to monitor trends of disease incidence if the sensitivity of the system remains reasonably constant over time.
- Planned uses of the gathered data, including the need for public health action. For example, when a reported case of a disease requires specific and immediate public health intervention, such as meningococcal disease, it is unacceptable for any cases to be missed by the surveillance system. For diseases that do not require specific or urgent intervention, such as salmonellosis, underreporting may not be considered to be unsatisfactory.
- Potential for causing outbreaks, e.g. norovirus gastroenteritis cases among the inpatients of a hospital or another healthcare facility, should be always notified in order to take measures to prevent extended outbreaks.
- National priorities. For example, in an area in which an eradication programme for brucellosis is being conducted, 100% reporting completeness is needed; whereas in an area in which the same disease is endemic, a lower level of completeness of reporting may be acceptable.
- International priorities/standards. Although the internal needs and priorities of a country always come first, international standards (European or global) should be taken into account.

Defining expected outcomes and methods of feedback for the evaluation

Expected outcomes of the evaluation should also be included in the written protocol before the evaluation is done, as should be a description of the means of communicating results to the stakeholders (e.g. reports, newsletters, bulletins, supervisory visits).

Feedback may involve the following persons and organisations:

- Data providers at all levels
- Local public health authorities
- Public health authorities at a central level
- Ministry and other implicated stakeholders

Suggesting options for remediation

A step that is frequently omitted when an evaluation is being designed is foreseeing the possible approaches for improving completeness, based on the preliminary information collected. Possible interventions should be identified during the design phase, using a cost–benefit analysis, rather than after the evaluation is done. If a low level of data quality is identified as a result of the evaluation, it should be followed by a realistic recommendation for improvement.

In summary, the different steps/decisions that should be made when designing an evaluation include the following:

- Diseases to be evaluated
- Timing for the evaluation
- Methods/measurements to be used
- Acceptable level of external completeness by disease and by population/area
- Plan for analysis and display of results
- Expected outcomes
- Methods, timing, and direction of disseminating output (feedback)
- Possible approaches to more complete reporting

All the information above should be included in the protocol created during the planning segment of the evaluation. The checklist below (Table 3.3b.) can be used with all the basic information that a protocol should contain for an evaluation of completeness.

Table 3.3b. Important information to be included in the evaluation protocol

Information	Included in protocol
Objectives of the evaluation	Yes/no
Triggers	Yes/no
Attribute/s	Yes/no
Description of the surveillance system	Yes/no
Available resources	Yes/no
Diseases that will be evaluated and why	Yes/no
When the evaluation will be conducted	Yes/no
Sources of data	Yes/no
Ethical considerations	Yes/no
Methods/measurements to be used	Yes/no
Pilot testing of methods for a limited set of data	Yes/no
Acceptable level of external completeness by disease	Yes/no
Analysis plan and display of results	Yes/no
Means of disseminating the evaluation report	Yes/no
Estimated cost of evaluation (cost analysis)	Yes/no
Time table, like Gantt diagram	Yes/no
Feasible approaches for increasing reporting (estimated costs)	Yes/no
References	Yes/no

Selecting methods

After the diseases to be evaluated are selected, one has to decide on the methods that will be used. This decision should take into account the following:

- The sources of the data to be used
- The necessity of conducting a study
- The study design
- The type of data analysis
- Display of results/expected outputs (usually described with dummy tables or graphs)

3.3.4 Evaluating external completeness

An evaluation can provide insights into the causes for the underperformance of a certain attribute; data quality monitoring cannot offer a full analysis.

Review of a sample of all surveillance data

The evaluation team contacts data providers (e.g. clinical doctors, microbiologists) and requests the 'actual' number of patients they have treated/diagnosed with a specific health problem (and within a specified time period) to be evaluated. This number is then compared with the number of cases reported to the surveillance system.

One of the advantages of this approach is that one can ask directly why some cases are not reported and whether there are specific difficulties that the data providers encountered when reporting certain cases. This helps to identify recurring patterns that may prompt a change in the reporting process.

This approach can be problematic because doctors may be reluctant to admit that they did not report cases of a disease or that they reported fewer cases than they actually diagnosed (underreporting). Furthermore, a

prerequisite in order for doctors to report the 'actual' number of cases is that they have access to some kind of case registry (personal or hospital registry). The quality of these registries heavily influences the information provided by clinical doctors or general practitioners.

Review of medical records

Another approach is to select a random sample of medical records (e.g. from hospitals, private offices) and then compare the number of identified cases with the number of cases reported to the surveillance system. Disadvantages include the fact that hospital registries are often incomplete or not updated, especially when the diagnosis does not change the given treatment. Internal completeness is also an issue, and information needed for case classification may be missing. Registries may vary substantially from one hospital to another in structure and content. In addition, diagnosis and recording of diseases may be based on different definitions.

Furthermore, reviewing paper records requires substantial time and effort. After a certain time, records may be archived or even destroyed/deleted. Unclear handwriting on paper forms may be difficult for the evaluation team to decipher. Records may be ambiguous, e.g. the same patient may have more than one file (in the absence of electronic records) if he or she has been hospitalised twice for the same disease (e.g. for tuberculosis). Duplicate reports can also be present in the system when patients have received care in more than one hospital or were referred from one hospital to another without an accompanying paper trail.

Reviewing electronic records can also be challenging when hospitals do not use a common registry/database format. Coding and variables can differ dramatically, and reviewers need to be competent users of various software and coding systems.

Finally, some information may be confidential and therefore inaccessible.

When using medical records for the evaluation of surveillance systems or research, it should be kept in mind that the main objective of hospital records is to serve logistical, legal and practical purposes (such as the avoidance of errors in the treatment of patients): hospital records are not created for the purpose of assisting surveillance systems and their operators.

Capture–recapture studies

Another approach to assess external completeness of a surveillance system is to compare its data with data from a different system (or several systems) with a capture–recapture study. The capture–recapture method was initially developed to estimate the size of wildlife populations [13–15].

The core idea of these studies is to sample and identify cases of a disease in a population and then resample the population to see what fraction of cases in the second sample were also identified in the first one. In this way, the total number of cases in the population can be estimated, as well as the level of completeness of each data source (or combinations of them).

Capture–recapture analysis has been used for estimating underreporting of cancer and other chronic diseases, as well as for tuberculosis, meningococcal disease, HIV/AIDS and other infectious diseases [16-31]. Capture–recapture studies used in evaluating surveillance systems are usually two- or three-source studies (comparison of datasets from two or three different surveillance systems). For simplicity, this chapter covers only two-source studies. The benefit of having a third data source is also addressed.

Calculation of external completeness in a capture–recapture study

Completeness of each of the surveillance systems in a capture–recapture study corresponds to the number of cases reported to this system, divided by the estimated total number of cases in the study population.

To estimate the total number of cases of a given disease in the population, we need to know the following:

- The number of cases reported to the first system
- The number of cases reported to the second one
- The number of cases reported to both systems

Chapman's formula for two-source capture–recapture studies and Chao's lower-bound estimator formula using frequency data adapted for a two-source capture–recapture study appear below [32]:

Chapman's formula:

$$n = \frac{(L1+1)(L2+1)}{d+1}$$

$$95\%\ CI = n \pm 1.96 \sqrt{\frac{(L1+1)(L2+1)(L1-d)(L2-d)}{(d+1)^2(d+2)}}$$

where L_1 is the number of cases reported to the first surveillance system, L_2 the number of cases reported to the second one, and d the number of cases reported to both of them.

Chao's lower-bound estimator:

$$n = f1 + f2 + \frac{f1^2}{4\,f2}$$

$$95\%\ CI = n \pm 1.96 \sqrt{\frac{f1^2}{4f2}\left(\frac{f1}{2f2} + 1\right)^2}$$

where f_1 is the frequency of those reported to exactly one surveillance system (equal to $[L_1+L_2-2d]$ in Chapman's formula), and f_2 the frequency of reports to both systems (equal to [d]).

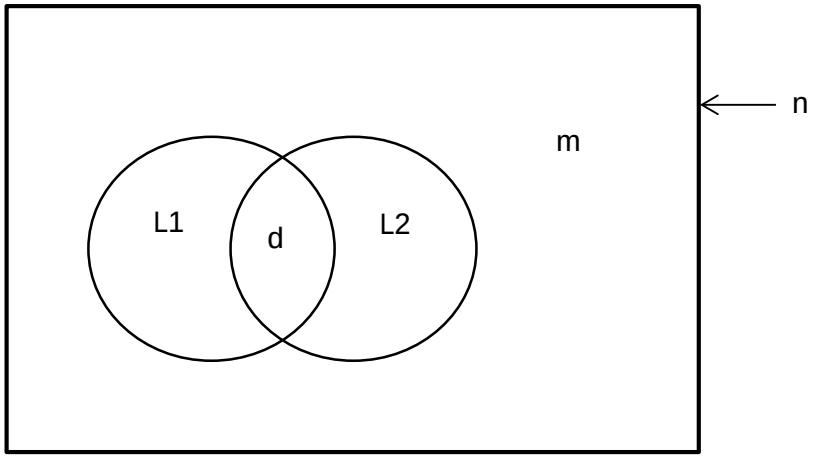
A Venn diagram is useful in an effort to provide a clearer picture of the measurement of underreporting with a capture–recapture study. In general, Venn diagrams (or set diagrams) show all possible logical relations (note: unless there are only two) between a finite number of sets (groups of objects). In other words, this diagram illustrates the relationships among sets of values. Venn diagrams usually comprise intersecting circles. To draw such a diagram, first a rectangle, which is called the 'universe,' is drawn, and then the intersecting circles are created. In a two-set Venn diagram, the following regions are produced:

- The interior of each circle, which represents the objects of the two sets
- The exterior of both circles, which represents the objects not included (empty set) in either of the sets
- The intersection of the two circles, which represents objects captured by both systems.

Similarly, a three-set Venn diagram consists of three symmetrically placed, mutually intersecting circles, for a total of eight regions.

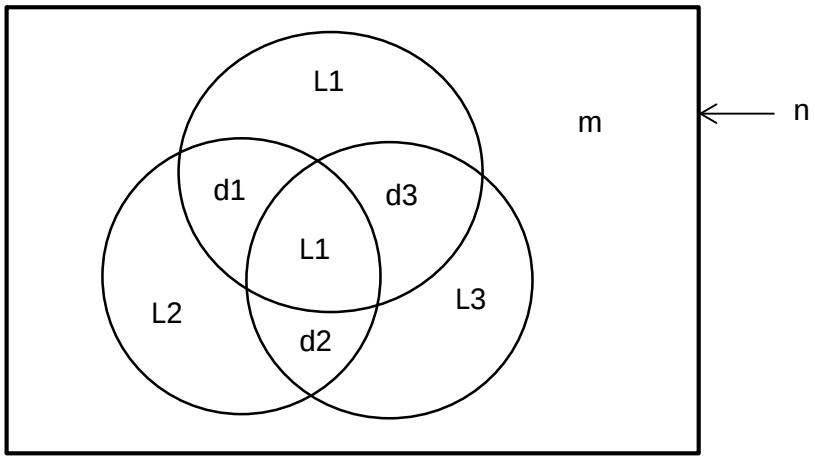
Illustrating the description of Venn diagrams, capture–recapture studies that include two and three surveillance systems are depicted in Figure 3.3b and Figure 3.3c, respectively. In both figures, the overlapping areas or intersections represent the cases that have been notified to both (or all three) systems.

Figure 3.3b. Venn diagram for a two-source capture–recapture study



In accordance with Chapman's formula, the rectangles in Figure 3.3b and Figure 3.3c represent the actual total number of cases (n) of the health problem, whereas L1, L2 and L3 indicate the number of cases in each list that derive from the parallel surveillance systems. Finally, *m* represents the number of the missing cases and *d* the duplicates.

Figure 3.3c. Venn diagram for a three-source capture–recapture study



Limitations of capture–recapture studies

In order to conduct a capture–recapture study, four criteria must be met.

The first criterion is that the population being studied – this method was first developed to estimate a population – is a closed, well-defined 'cohort' with no losses or new entries in a given time period. However, when we want to estimate the number of cases diagnosed during a time period in the past, this criterion is always met, because that number is fixed.

The second criterion is that the cases captured by both (or more) surveillance systems can be identified and matched. To meet this criterion, each case should have had at least one unique identifier when it was reported and entered into the database of the first system, and that unique identifier will be identical to a unique identifier from the second (or any additional) system. This common unique identifier ideally is usually a national identification number, such as the personal insurance number. In the absence of a unique national identifier or in case it cannot be used (e.g. because of confidentiality concerns), a combination of full name (initials), gender and age is usually sufficient, although adding the post code (if available) is useful.

The evaluation team should keep in mind that a) different types of identifiers may give different evaluation results; b) the more matching criteria you use, the better the matching; and c) when a common identifier is used, cases from two or more databases can easily be matched with a number of easily available computer software packages.

The third criterion in capture–recapture analysis is that each case in the study population has an equal chance to be reported to each of the systems. This means that capture probabilities are homogeneous across all individuals in the population (equal 'catchability'). Equal catchability is met when the probability of notification of one case to the surveillance systems is not influenced by the characteristics of the patient represented by that case, e.g. age, gender, place of residence, ethnicity, socio-economic status, severity of symptoms, stage of disease. This probability may vary among surveillance systems. In practice, this criterion is rarely met.

The fourth criterion refers to the independence of the data, which refers to the fact that the probability that each case will be notified to a surveillance system is not influenced by the probability that this case will be notified to another system [34-35].

The dependence that may exist for data of two (or more) surveillance systems can be 'negative' or 'positive' [35]:

- Negative dependence is present when cases reported to the first surveillance system are less likely than cases not reported to it to be notified to another surveillance system. In this situation, there is little overlap between the datasets, and the formulas described above will overestimate the actual total number of cases. The sensitivity of the system will then be underestimated.
- Positive dependence is present when cases reported to the first system are more likely than cases not reported to it to be also reported to another surveillance system. Here, the overlap between the two datasets is large, the total number of cases is underestimated, and sensitivity is overestimated.

Other characteristics of dependence also have to be taken into account:

- Data can never be totally independent among surveillance systems. Positive dependence occurs more frequently than negative.
- Since the probability that a case will be notified usually depends on covariates for the patient, such as severity of disease or age, proposed solutions to deal with confounding are:
  - to subdivide cases into different groups (strata) according to the covariate(s), to estimate the total number of cases for each stratum, and then to combine the estimates [36]. This technique is known as stratification.

– Often, covariate information is not available, but another approach can be used that involves at least three sources of data. Multiple different two-source estimates are made, considering each source against all of the others pooled [37]. This approach allows greater accuracy and reduces the effect of data dependence. Mathematical methods that are used to identify and deal with the interdependence of sources, such as log–linear modelling (usually using Poisson regression) or multinomial logit modelling, may be complex [34, 38-40], but their use has been greatly eased through computer software applications.

The preferred method for overcoming dependence of data among different surveillance systems is to include three or more data sources in the capture–recapture study, which makes it possible to adjust for dependence and variable catchability. When four or more data sources are used, the mathematical computations become increasingly complex, and little benefit may be obtained from the effort. The selection of three quality lists is the optimal choice.

The use of capture–recapture methods for estimating completeness of surveillance systems for infectious diseases has been criticised because the data from those systems carry an elevated risk of violating one or more of the four assumptions [36]. Consequently, the validity of capture–recapture studies requires critical scrutiny, and results must be interpreted with caution.

Some tips to keep in mind when using capture–recapture analysis to evaluate completeness are shown below:

- Most capture–recapture studies violate the underlying conditions to some degree.
- Capture–recapture studies are easy to perform and do not require much time because the needed data are already available. Despite their limitations, capture–recapture studies can be useful in evaluations of completeness.
- When you conduct capture–recapture studies, you can usually identify possible violations of the criteria discussed above and can try to predict their effect on the estimation of underreporting.

Other methods

Other methods for data collection (mainly of qualitative data) can also be useful in assessing the functionality of the system and in identifying problems that lead to data quality issues. Some examples of such methods are provided below [41]:

- Supervision
- Surveillance meetings and workshops
- Document review
- Focus group discussions
- Key informant interviews
- Analysis of strengths, weaknesses, opportunities and threats (SWOT analysis)

These are simple and direct methods which are helpful when applying case definitions, assessing the data provider's knowledge of surveillance, and examining how surveillance reporting is incorporated into the daily work of physicians. These approaches are mostly useful in the early stages of an evaluation, but when there is a need to quantify underreporting, methods such as the ones described earlier are preferable.

3.3.5 Interpreting the results of the evaluation

On the basis of the results of the evaluation, the researchers may conclude that completeness of the surveillance system is lower than acceptable. If this is the case, the reasons that have led to low completeness should be considered one by one when interpreting the results. A checklist of common reasons that can lead to low external completeness of a surveillance system is provided in Table 3.3c.

Table 3.3c. Checklist of possible reasons for low external completeness of surveillance systems

Possible reason (disease A)	Observed
Time-consuming notification process	Yes/no
Problems with hardcopy reporting form	Yes/no
Problems with the electronic surveillance form/database	Yes/no
Technical problems (e.g. lack of support)	Yes/no
Practical constraints (e.g. lack of resources)	Yes/no
Unclear notification process	Yes/no
Lack of adherence of data providers	Yes/no
Lack of training for data providers	Yes/no
Low acceptability of surveillance system by the data providers	Yes/no
Lack of appropriate feedback to data providers	Yes/no

Possible reason (disease A)	Observed
Confidentiality issues/concerns of data providers	Yes/no
Unclear to data providers who owns surveillance data and how data will be used (lack of knowledge on the objectives and usefulness of the surveillance system)	Yes/no

The checklist included in Table 3.3c can apply to all diseases, to a specific subgroup, or to a single one. Reasons for low external completeness can be grouped as follows:

- Factors that have to do with the healthcare system in general (e.g. lack of personnel)
- Factors that have to do with the data providers (e.g. lack of interest, training or proper supervision)
- Problems in the structure and functionality of the system

For example, long or complicated reporting forms discourage data providers from reporting regularly, especially when there is lack of personnel or even lack of space (a separate office/corner), where the electronic system can be used whenever it is convenient for them. In addition, doctors may not be aware of their obligation to report, the current list of notifiable diseases, or the exact reporting process (when to report and how). Physicians who treat HIV/AIDS patients may be reluctant to report cases because of confidentiality concerns. All these barriers – and a number that have not been discussed – may be relevant and should be taken into account when recommending measures that are aimed at improving the completeness of a surveillance system.

It is possible that the results of an evaluation show that completeness of the system is well within the acceptable values for some of the diseases. Possible explanations for satisfactory performance of the surveillance system may be useful for future strengthening of the system and for justifying requests for additional resources.

3.3.6 Improving external completeness

Several approaches have been proposed to increase external completeness of a surveillance system. None of them is suitable for *all* diseases and *all* systems. The approach you choose should take into account the capacities and the specific infrastructure of the system.

Changes proposed for a surveillance system in order to increase its completeness should be coherent with the infrastructure of the healthcare system.

Some common strategies that have been used in the past include the following:

- Making notification obligatory. This approach has been widely used, but studies have demonstrated low levels of reporting of several mandatory notifiable disease systems.
- Provide physicians with incentives (e.g. remuneration schemes) for motivating them to report cases [42]. Again, this approach tends not to be successful in the long term because the motivation for doctors to report soon decreases. In addition, it has several disadvantages because doctors may report cases in order to be paid without actually believing in the usefulness of reporting. This practice may lead to artificially high numbers of reported cases or false alarms. Also, with a fee structure, if payments cease for any reason, the entire reporting process may collapse.
- Sample-based or sentinel approaches. Resources can be invested in achieving higher reporting rates from a small number of data providers by providing them with training and support. A smaller group of physicians can be more easily reached to be given advice on notification and its benefits, feedback on surveillance results, and information on changes in reporting procedures and case definitions. However, this approach is suitable only for frequently occurring diseases for which the main purpose of surveillance is to record temporal changes in the occurrence of the disease and to detect outbreaks. For diseases that require public health action for every single reported case, such as meningococcal meningitis and tuberculosis, this is not a suitable approach.
- Creating a network of dedicated data providers for specific diseases. When healthcare (e.g. diagnosis, treatment) for a particular disease or group of diseases is provided in specific healthcare settings, surveillance can be limited to those sources. This approach can often achieve higher notification rates and require equal or less resources than traditional universal reporting.

Below are a few examples of approaches to increase external completeness. Some of these examples may also apply to internal completeness.

- Surveillance and follow-up of all acute flaccid paralysis (AFP) cases is a basic task of surveillance systems in Europe where poliomyelitis has been eradicated. Based on a World Health Organization programme, cases are reported weekly and are followed-up until their status is verified. In some countries, active surveillance of all AFP cases is in place: public health professionals call clinical neurologists once a week and ask about AFP cases.

- In most countries, patients with HIV/AIDS or other sexually transmitted infections are referred to specific laboratories and clinics/hospitals for diagnosis and to specialists (e.g. gynaecologists, dermatologists) for treatment. These doctors can establish a network for reporting these diseases.

- During an outbreak of Shiga toxin-producing *Escherichia coli* in Germany, nephrologists have contributed to the data for the surveillance of haemolytic uraemic syndrome. This was an additional source of information for the regular reporting system and was used for the evaluation of completeness.
- Active surveillance. Reporters are regularly prompted to report cases. This approach is mainly used for rare diseases. Sending out reporting cards to physicians to request that they report cases from a list of monitored diseases has been used to raise completeness in Britain [43].
- Reporting can make use of data collected for clinical or other operational purposes as part of routine clinical or administrative processes. Capturing data without requiring additional efforts by clinical staff includes making use of laboratory requests, insurance reimbursement forms, and pharmacy records (e.g. for dispensing vaccines or disease-specific medications). Again, using external sources of information for surveillance purposes can lead to false conclusions because documentation from such sources is designed to serve other purposes, mainly administrative.
- Information from one system can be used for increasing completeness of another system. For example, information on a positive blood culture with Gram-negative *diplococci* at the laboratory can be a trigger to remind clinicians to report meningococcal disease to the mandatory notification system. Even though this approach can be helpful, it is time-consuming and does not solve the problem of underreporting.
- Designing reporting forms and databases to be as user-friendly as possible. This approach has been identified as helpful in the past, especially when data providers were consulted about the problems they faced in completing reporting forms or entering data into the surveillance database.
  - Using short reporting forms that contain only essential information on no more than one page
  - Building automated 'jumps' into the data entry screens to skip over fields that do not apply to a reported case
  - Automated help messages during data entry for fields that may not be self-explanatory. For example, the field 'high-risk group' may need clarification because it may be perceived differently by different healthcare workers.
  - Automating alerts and notification requests can increase reporting rates of passive surveillance systems.
  - Implementing electronic reporting systems instead of paper-based ones. Published reports have shown that electronic reporting improves completeness of surveillance data [44-47].
  - Improving the process of providing feedback to data providers also enhances the commitment of healthcare professionals in reporting infectious disease and other health outcomes. It may also help public health authorities demonstrate that data are not only collected for administrative reasons or scientific publications but primarily for public health actions. It is important for data providers to see that their time and effort can lead to specific interventions.

3.4 Evaluation of sensitivity, specificity and positive predictive value

Evaluation of sensitivity, specificity and positive predictive value are linked to the evaluation of external completeness and validity. While sensitivity, specificity and positive predictive value refer to the reporting and the detection/ascertainment process, external completeness is limited to the reporting process. A summary of the information that should be included in an evaluation protocol can be found in Figure 3.4a.

Figure 3.4a. Key points for the evaluation of sensitivity, specificity and positive predictive value

- Evaluation of sensitivity, specificity and positive predictive value in a surveillance system is linked to the evaluation of completeness and validity.
- Sensitivity refers to the proportion of the actual cases that were reported to the system; specificity refers to the proportion of non-cases identified as such and not reported to the system.
- High sensitivity means that few cases are missed, and high positive predictive value that most reported cases are indeed cases of the specific disease.
- Improving one of these attributes usually has the opposite effect on other attributes of the system; surveillance systems with high sensitivity also have an increased number of false positive reports.
- Surveillance systems have to balance the risk of false alarms for cases/outbreaks with the value of early detection.

3.4.1 Factors which influence the sensitivity of a surveillance system

The following factors may influence the sensitivity of a surveillance system:

- The severity of the disease. Patients with mild symptoms do not usually seek medical attention, are under-diagnosed, and consequently underreported. Gastroenteritis is an example of a disease with a high under-ascertainment rate because of the mildness of symptoms and the fact that they usually resolve without treatment.
- The skill of the healthcare providers to diagnose the disease, which is especially important for rare, emerging, or re-emerging diseases. For example, a case of encephalitis caused by West Nile virus may not be diagnosed at all or may be diagnosed with a substantial delay if the physician does not suspect the specific virus as the aetiological agent.
- The sensitivity of the used diagnostic tests/screening.
- Access to the healthcare system. Difficulties and barriers that patients face when they seek medical help can lead to an increased number of undiagnosed cases. Geographic constrains and economic factors are examples of such possible barriers.
- The criteria used for reference for further laboratory testing or other exams. For example, stool specimens can be tested for *Campylobacter* spp. routinely, after a doctor's request, or in case of bloody diarrhoea.
- Laboratory capacity for the detection of different pathogens (for infectious diseases) or un-/availability of medical equipment/expertise necessary for diagnosis.
- Factors that create difficulties in reporting. These factors can be related to a) the surveillance system itself (long or complicated reporting forms, no electronic reporting), b) data providers (lack of interest or training), and c) the healthcare system (lack of personnel and/or technical support).

3.4.2 Methods

In order to assess sensitivity, specificity and positive predictive value, an external source of reference data is needed ('gold standard'). This source should be independent and as complete as possible. Cases (or outbreaks) included in the reference data will be classified as the true cases of the given disease (or the true outbreaks) in the population under surveillance in a specific time period. The calculation of the attributes mentioned below can be demonstrated with the use of a 2 x 2 contingency table.

Sensitivity

The actual frequency of the disease in the population under surveillance can be determined with the use of a reference source or alternatively through estimation of the total number of cases in the population under surveillance by using capture–recapture techniques.

The data collected by the system are validated.

Sensitivity (Se) based on the 2 x 2 table already presented is equal to

$$\mathrm{Se} = \frac{a}{a+c} \times 100 \text{ (95\% confidence interval)}$$

$$95\%\ \mathrm{CI} = \mathrm{Se} \pm 1.96 \sqrt{\frac{\mathrm{Se}(1-\mathrm{Se})}{a+c}}$$

Note: In this case it is assumed that most reported cases are correctly classified.

Se is the complementary proportion of underreporting (false negatives) as:

$$1 - \text{Se} = \left(1 - \frac{a}{c + a}\right) = \frac{c}{a + c}$$

High Se means that external completeness of the surveillance system is high, and low sensitivity means a high proportion of missed cases.

To confidently assess the sensitivity of a public health surveillance system, more than one measurement of an attribute is needed. For example, sensitivity could be determined for each data source, for combinations of data sources, or for specific time periods.

Specificity

According to the 2 x 2 table, specificity (Sp) equals:

$$\text{Sp} = \frac{d}{d+b} \times 100 \text{ (95\% confidence interval)}$$

$$95\%\ \text{CI} = \text{Sp} \pm 1.96 \sqrt{\frac{\text{Sp}\ (1 - \text{Sp})}{d + b}}$$

Specificity is the complementary proportion of the false-positives cases. The proportion of false positives is the proportion of the non-cases that were reported and equals b/(b+d).

An increase of the specificity usually results in a decrease of Se and an increase of PPV of the system.

In summary, four percentages can be calculated when 'reading' the 2 x 2 table vertically:

- % of true positive cases (Se)
- % of false positive cases (1 – Sp)
- % of true negative cases (Sp)
- % of false negative cases (1 – Se)

Positive predictive value

- Positive predictive value (PPV) equals:

  $$\text{PPV} = \frac{a}{a+b} \times 100 \text{ (95\% confidence interval)}$$

  $$95\%\ \text{CI} = \text{PPV} \pm 1.96 \sqrt{\frac{\text{PPV}\ (1 - \text{PPV})}{a + b}}$$

- PPV is the complementary probability of positive predictive error (PPE)

  $$\text{PPV} = 1\text{-PPE}$$

- PPE is the proportion of reported cases that are not true cases of the disease

  $$\text{PPE} = \frac{b}{a+b} \times 100 \text{ (95\% confidence interval)}$$

  $$95\%\ \text{CI} = \text{PPE} \pm 1.96 \sqrt{\frac{\text{PPE}\ (1 - \text{PPE})}{a + b}}$$

Negative predictive value (NPV)

- NPV is the proportion of not reported cases that actually do not have the disease and equals

  $$\text{NPV} = \frac{d}{c+d} \times 100 \text{ (95\% confidence interval)}$$

  $$95\%\ \text{CI} = \text{NPV} \pm 1.96 \sqrt{\frac{\text{NPV}\ (1 - \text{NPV})}{c + d}}$$

- NPV is the complementary probability of negative predictive error (NPE)

  $$\text{NPV} = 1\text{-NPE}$$

- NEP is the proportion of not reported cases that are actually cases of the disease and equals to

  $$\text{NPE} = \frac{c}{c+d} \times 100 \text{ (95\% confidence interval)}$$

$$95\%\ \mathrm{CI} = \mathrm{NPE} \pm 1.96 \sqrt{\frac{\mathrm{NPE}\ (1 - \mathrm{NPE})}{c + d}}$$

NPV depends on the Se of the system; the more sensitive the system, the easier it captures cases of the disease. Thus, when a case is not reported, it is probably truly not a case. NPV also depends on disease prevalence because high prevalence of the disease leads to more false negative cases and low prevalence to more true negative cases.

When 'reading' the 2 x 2 table horizontally we can calculate:

- Positive predictive value (PPV)
- Positive predictive error (PPE)
- Negative predictive value (NPV)
- Negative predictive error (NPE)

In case of categorical variables with more than two values, the approach is similar. In this case, *n x n* tables are collapsed to *n (2 x 2)* tables.

3.4.3 Results

Table 3.4a shows a blank table for the documentation of Se and PPV.

Table 3.4a. Sensitivity (Se) and positive predictive value (PPV) of surveillance system X

Disease	Number of cases – reference source	Number of notified cases	Se (95% CI*)	PPV (95% CI*)
Disease A				
Disease B				

** CI: confidence interval*

3.4.4 Interpreting the results of an evaluation

Ideally, all cases of a particular disease are reported to the surveillance system, and all non-cases remain unreported. However, this rarely happens. For the purpose of evaluation, an acceptable or desired level (cut-off point) of Se, Sp and PPV is set. This level is compared with the results of the evaluation, and changes are then proposed to improve the system. Setting an acceptable level for these attributes is challenging, and their effect on the system needs to be balanced.

Factors to take into account when deciding on an acceptable level of Se, Sp and PPV include the following objectives of the surveillance system:

- The perceived threat of a disease
- The system's possibility to detect an outbreak
- The community value attached to early detection
- The availability of resources
- The cost of investigations (which may be substantial, especially in case of outbreaks)

As these factors can vary substantially over time, defining the right balance between Se, Sp and PPV for the purpose of an evaluation is difficult.

Sensitivity

Sensitivity depends on the surveillance objectives. For example, if the objective of rabies surveillance is to propose prophylaxis to exposed persons, sensitivity should be as high as possible. Only when monitoring trends is the main objective, sensitivity can be lower. At the other side of the spectrum, high-sensitivity systems aim at identifying every single case of a disease in the population so specific public health measures can be taken for every case.

If sensitivity is too low, the detection of outbreaks and the implementation of interventions will be delayed, while an overly high sensitivity level may lead to unnecessary, resource-intensive investigations.

Specificity and PPV

When resources are limited, it can be advisable to set the acceptable level for Sp and PPV to a higher level in order to preserve resources. Low Sp and PPV are associated with responding to false notifications and false outbreak signals. They may also cause unnecessary public concern and hamper the system's credibility. In other words, it may be acceptable to miss small outbreaks or cases of less severe diseases with a small impact on morbidity in order to save resources for other public health purposes.

The mathematical relations between predictive value, Se and Sp are given below:

$$\text{PPV} = \frac{\text{Se} \times \text{P}}{\text{Se} \times \text{P} + (1 - \text{Sp})(1 - \text{P})}$$

$$\text{NPV} = \frac{\text{Sp} \times (1 - \text{P})}{\text{Sp} \times (1 - \text{P}) + (1 - \text{Se}) \times \text{P}}$$

P = Prevalence of the disease

3.4.5 Enhancing sensitivity, specificity and positive predictive value

Steps taken to improve one attribute can often affect the other and therefore it is suggested to proceed with caution when implementing changes. Some actions will affect PPV (e.g. broadened case definitions) whereas other actions may not affect PPV (improved reporting).

The sensitivity of the surveillance system can be improved by the following actions:

- Broaden the case definition.
- Encourage the reporting of possible cases without laboratory confirmation.
- Conduct active surveillance (contact doctors at hospitals, private doctors, laboratories).
- Implement an electronic reporting system.
- Remove barriers of reporting (technical difficulties, lack of knowledge, etc.).
- Track all reported 'possible cases', investigate them, and dismiss those which are not true cases.
- Monitor the diagnostic effort (e.g. track submission of laboratory requests for diagnostic testing) and give recommendations for diagnostic procedures and diagnostic algorithms based on evidence-based medicine.
- Train data providers in notification processes and sensitise them by providing regular feedback on their performance.

Additional measures that can be taken to improve the sensitivity of the system during an outbreak include the following:

- Implement syndromic surveillance.
- Set lower investigation and alarm thresholds.

The measures taken to increase the PPV of a surveillance system are the opposite of the measures taken to increase a system's sensitivity: the PPV of the surveillance system is increased if public health authorities, for example, use narrower case definitions, employ strict decision rules about electronic records, encourage clinicians and labs to report only confirmed cases, and review all cases before entering the data into the system. During outbreaks, public health authorities may raise investigation and alarm thresholds in order to improve the PPV of the system.

Modifications to improve performance by changing a single attribute usually leads to deterioration in another attribute. Only changes that relate to the system as a whole (e.g. by adding a data type or applying a new detection algorithm) will not affect the balance of the system.

3.4.6 Questions and answers

Q. What is the difference between completeness and sensitivity of a surveillance system?

A. The sensitivity of a surveillance system includes a case detection/diagnosis component and a disease-reporting component, while external completeness refers only to the proportion of diagnosed cases that are reported to the surveillance system.

Q. Can you give examples of diseases for which a high level of sensitivity of the system is important?

A.

- SARS (an emerging disease, contagious, high fatality rate)
- Poliomyelitis (severe disease, under eradication)
- Haemorrhagic fever (immediate public health measures required)
- Meningococcal meningitis (measures for preventing cases among the close contacts of the patient)

Q. Can you give examples of diseases for which an increased PPV is important?

A.

- Gastroenteritis outbreaks (common disease, a degree of certainty is required for initiating investigation)
- Pulmonary anthrax (possibility of bioterrorism, increased public and political concern)

- Dengue fever (in a county which is affected for the first time: many people will be involved in case management and prevention of new cases, public concern will be high, information has to be shared with other countries)

Q. How are measures to increase sensitivity likely to affect other attributes of the surveillance system?

A. Specificity and positive predictive value will be decreased, and negative predictive value will be increased.

Q. Are measures to increase positive predictive value likely to affect the other attributes of the system?

A. Sensitivity will be decreased and specificity will be increased.

Q. Can you list the most important criteria for adjusting the balance between sensitivity and positive predictive value?

A.

- The objectives of the surveillance system
- The perceived threat of a disease
- Its potential to cause outbreaks
- The community value attached to early detection
- The availability of resources
- The cost of investigations and public health interventions

Q. What are the main components of a surveillance system that need to be taken into account before evaluating data quality?

A.

- Main stakeholders
- Objectives
- Type of system
- Population being monitored
- Data sources
- Data providers
- Types of data collected
- Data flow
- Available resources
- Data confidentiality
- Legislation

Q. Can you list the main factors that are taken into account when selecting diseases to be included in an evaluation?

A.

- Objectives and priorities of the system
- Feasibility of evaluating a specific disease
- Available resources
- Triggers that may indicate the need to include a disease

Q. Can you refer to examples of diseases or health problems for which it is important to evaluate completeness and explain why?

A. Completeness is an attribute that is important for all notifiable diseases. However, there are diseases where even a single case can cause a major public health problem. For example:

- Meningococcal disease. Missed cases may lead to additional transmission of cases among the population. All cases should be reported so that healthcare providers can give appropriate prophylaxis to close contacts of cases (e.g. family, friends, classmates).
- Tuberculosis. Identification of the source of the infection and of other cases (symptomatic and asymptomatic) has become extremely important because of the high incidence of the disease, increasing antibiotic resistance, and co-infection with HIV.
- Measles. The implementation of control and preventive measures for each diagnosed case is essential because measles spreads easily, causes clusters of cases or even outbreaks among unvaccinated people; in addition, there is the goal to eliminate the disease worldwide.

Q. Which external data sources can be used for evaluating completeness, and which one is the most appropriate?

A. Medical records, laboratory registries, insurance files, death certificates, lists of prescriptions, association membership lists, registries of non-profit organisations are only a few of the data sources that can be used for this purpose.

- All types of data sources have advantages and disadvantages, and the selection of the most appropriate ones depends on the context in which evaluation occurs.
- Medical records are widely used as an external data source.

Q. Can you list some disadvantages of using medical records to evaluate completeness?

A.

- Not always complete or updated
- Important information may be missing
- Heterogeneity among hospitals
- Use of different definitions
- Confidentiality issues

Q. Given that the assumptions of a capture–recapture analysis are unlikely to be totally valid when looking at epidemiologic data, why use this method?

A. Capture–recapture studies are easy to perform and do not require much time. Despite their limitations, they are a useful epidemiologic tool. When we use this kind of study, we must identify possible violations of the four key assumptions in order to describe the study limitations and recommend caution when interpreting results.

Q. Which assumptions do we make when we apply the capture–recapture method?

A. There are four main assumptions:

- The number of diagnosed cases is a static figure.
- The cases captured by two or more surveillance systems can be identified and matched.
- Each case has an equal chance of being reported to each of the systems (equal catchability).
- The probability that each case will be notified to a given surveillance system is not influenced by the probability that it will be notified to a second system (independence of data).

Q. Name some factors that may lead to low external completeness/sensitivity of surveillance systems and have to do with the surveillance system itself?

A. Low completeness may result from factors that have to do with the healthcare system (e.g. lack of personnel), the data providers (lack of interest or training), and also with structural problems of the system such as any of the following:

- Time-consuming notification
- Complicated procedure (Who sends data? Where? How?)
- Problems with hard-copy reporting forms (e.g. use of abbreviations, unknown terms)
- Problems with an electronic surveillance system (e.g. difficulties in retrieving records and saving new ones)

3.5 Evaluation of external validity

Key points on the evaluation process for external validity are listed in Figure 3.5a.

Figure 3.5a. Key points for evaluating external validity

- External validity is important for the data quality of surveillance systems.
- In order for surveillance data to be valid they should correspond to 'true' data.
- Invalid information in a surveillance system can lead to false 'signals' or 'alarms' and unnecessary public health interventions. This also increases the risk of missing a real alert.
- Low concordance between surveillance data and reference source data indicates low data validity.

Evaluating external validity of a surveillance system usually involves one or more studies using different methods and a combination of several data collection methods. This technique of combining different methods facilitates validation of data through cross-verification from different data sources. The term triangulation is used to describe this technique.

In most cases, surveillance data are compared with the data of another source that is considered to be the 'gold standard'. Which evaluation method is selected depends on:

- the types of errors to be addressed (duplicate records, case misclassification, etc.);
- the types of variables that have been selected to be evaluated (categorical, non-categorical);
- availability of external data sources; and
- cost (associated with the various methods).

In addition, the evaluation team has to choose between retrospective (more common) or prospective evaluation methods. Retrospective methods use data which were collected earlier, while prospective methods collect data after the notification of cases. Prospective studies may lead to more accurate estimates, especially in the absence of a quality reference data source, but they need more resources and time.

3.5.1 Identifying duplicate records

Duplicate records are more frequent than one may think and can be identified by creating a string variable, using the last name (text variable), the date of birth (date variable), and the gender (binary variable) for each person listed. If the names of the patients cannot be accessed, reported cases with the same identification number can be used for the same purpose.

Identifying duplicate case records can be difficult, and a common set of criteria needs to be employed to delete the duplicates. For example, when two cases of tuberculosis with identical identification numbers appear in a dataset, but with different dates for symptom onset, it can be difficult to decide whether this is a duplicate. In this case, a minimum time interval between notifications (cut-off point) needs to be set to determine whether the second entry should be listed as a separate case.

The need to define disambiguation criteria (e.g. cut-off points) depends on certain aspects of the reported disease (immunity, frequency, or relapse). These criteria should be determined before the evaluation and depend on available system variables, the availability of a unique identification number, and the overall completeness of information.

3.5.2 Calculating concordance through record linkage

Concordance is a measure of the agreement among data sources, in this case between surveillance data and data from some external source. High concordance indicates increased validity of surveillance data, while low concordance indicates low validity and the need for intervention. Concordance can be estimated via a record linkage between surveillance data and data from an independent dataset. This is a frequently used method, especially for categorical variables.

Several data sources can be used for this purpose, such as medical records, records of reference laboratories, hospital discharge papers, vaccination cards, or death certificates. Because it is unlikely that a complete data source is available, in practice the most complete data source one can access is used.

Medical records of reported cases are the most common data source for estimating the validity of surveillance data (see Section 3.3). To assess the accuracy of the information included in medical records, the evaluation team should become familiar with the different processes of recording data at different healthcare settings. To complicate matters further, the retrieval of medical records – and subsequent requests for clarification – is not always an easy task.

In order to calculate concordance, a sample of the cases reported to the surveillance system is selected, and medical records are retrospectively reviewed. The number of the cases one needs to include in the study (sample size) is easily calculated with a standard statistical software package. After selecting an adequate number of cases, the records of the surveillance data set are linked to the appropriate medical records, and data are compared. The agreement between the two data sources is determined, and discrepancies are identified. The level of agreement expresses the degree of validity. Data are compared by variable of interest (stratification of cases by variable). Figure 3.5b describes the typical steps of this calculation.

Figure 3.5b. Schematic presentation of the steps of a study of concordance between surveillance data and medical records (external data source)

Step 1. Calculate the required number of reported cases (sample size).

Step 2. Select cases reported to the surveillance system via a stratified design.

Step 3. Identify the appropriate medical records and have the hospital retrieve them.

Step 4. Record linkage and conduct in-depth review of the records (one or more experts).

Step 5. Compare data from the two sources by variable.

Step 6. Recognise discrepancies by variable and disease.

Step 7. Calculate concordance of data sources at 95% CI (confidence interval).

Step 8. Display results (usually in a table).

Data should be reviewed simultaneously by two or more experts in order to minimise the effect of subjectivity. The need for more than one expert to review the same data – and the type of expertise needed – depends on:

- the kind of external data source that will be used;
- the data quality of the external source;
- the data format (electronic or paper-based); reviewing paper-based data is prone to be more subjective because of differences in the type of documentation, handwriting and wording.
- the type of information the study focuses on; and
- the availability of human resources.

Physicians with a good knowledge of the evaluated disease and epidemiologists with experience in medical record abstraction and surveillance can be chosen to independently examine medical records and identify possible discrepancies. This approach may be useful when evaluating case definitions. Determining which cases in the medical records meet the case definition can be difficult, especially for diseases with complicated case definitions (clinical manifestation, laboratory and epidemiologic criteria).

The concordance between the two data sources is calculated as the proportion of cases in agreement by characteristic (variable). The proportion of cases in agreement can take any value between 0 and 100%.

In summary:

- The higher the concordance, the higher the validity of surveillance data.
- Increased proportions of different values means that errors were made during the reporting process (provided that medical records were correct).

Results can be displayed as exemplified in the sample table below (Table 3.5a).

Table 3.5a. Evaluation of validity of data in a surveillance system, comparison between surveillance data and data from a reference source

Variable	Number of cases – reference data source	Number of notified cases	% of cases in accordance (95% CI*)	% of contradictory cases (95% CI*)
Gender				
Nationality				
Vaccination status				
Outcome				

** CI: Confidence interval*

3.5.3 Calculating concordance via follow-up of notified cases

An alternative way to evaluate concordance is to select a sample of reported cases and actively follow them up to verify the information included in the notification. Already collected information can be verified by personal visits to the hospitals, reviews of charts and medical records, contacts with doctors or public health professionals, or via personal interviews with the patients. The main disadvantages of this approach are that it is costly and time-consuming, and raises confidentiality issues.

3.5.4 Validity checks during statistical analysis

The validity of specific variables can be calculated during statistical analysis by, for example, calculating the absolute mean difference between the two sources for date variables. Normally, one would expect the difference to be zero, which is hardly ever the case. Small differences usually do not change the plan of action, but for diseases that require direct intervention, these discrepancies are relevant. The results of such calculations can be summarised as presented in Table 3.5b.

Table 3.5b. Absolute mean difference between dates in the notification dataset and in the medical record (or alternative data source)

Date*	Absolute mean difference (days/months)	Coefficient of variation %
Date of birth		
Date of onset of symptoms		
Date of laboratory confirmation		
Date of treatment initiation		
Date of death		

** Date formats: month/day/year or month/year*

Similar calculations can be conducted for ages, titres and other non-categorical, continuous variables of interest.

Further examples of logical validity checks during statistical analysis in order to recognise errors in surveillance data:

- Selection of all cases of congenital toxoplasmosis and identification of errors regarding age.
- All confirmed salmonellosis cases should have a positive value for *Salmonella* spp. stool culture. Errors are immediately detected if confirmed salmonellosis cases are connected with records of negative stool culture.
- For typhoid fever, the serotype of *Salmonella* in laboratory data must always be *S. typhi.* If there are records of typhoid fever which list serotype *S. typhimurium*, one variable contains invalid data.

The main disadvantage of this approach is that in some cases the correct value cannot be easily found and data correction may be difficult or impossible.

3.5.5 Validity of outbreak notifications

Evaluation of validity, as previously discussed, is important because of the substantial cost of false alarms and the delay in the investigations of actual public health emergencies.

The number of actual outbreaks, the proportion of false alarms, and the proportion of missed outbreaks during a specific time period can be used to assess validity. The number of possible outbreaks corresponds to the number of statistical aberrations detected at a set threshold over a defined period of time. Some of these outbreaks prove to be real outbreaks, and some do not. The number of false alarms can be estimated with the registries kept by local public health authorities on investigations (if available). The lower the percentage of false alarms, the better the quality of the surveillance data. Data regarding the validity of reports of outbreaks can be summarised as shown in Table 3.5c.

Table 3.5c. Validity of reports of outbreaks of specific diseases

Disease	Number of alarms – signals	Number of true alarms (%)	Number of false alarms (%)	Number of missed outbreaks (%)
Salmonellosis				
Measles				
Legionellosis				

3.5.6 Indirect methods to evaluate validity

The quality of data can be influenced, for example, by ambiguously phrased questions in surveillance forms, poorly trained data providers, the absence of validity checks, poor communication between data providers and public health professionals, and the lack of effective communication channels.

These indicators can be assessed, for example by asking data providers to share their personal opinions on the used reporting forms or by requesting doctors to participate in a mock scenario in order to identify reporting errors.

These indicators may be strong enough to trigger an evaluation of the validity of a surveillance system; they should, however, not be used to draw definitive conclusions on data validity.

3.5.7 Ethical considerations

The evaluation protocol should address ethical considerations. The evaluation team should make sure that information will only be used for the purposes of the evaluation and that personal data are only used if justified by the objectives of the evaluation.

Disease data are often highly confidential, which makes ethical considerations even more important. Patient consent is an important issue, especially when the medical records of patients are used. Medical confidentiality implies that:

- national and international standards should always be followed, and all measures necessary to protect the confidentiality of personal data should be followed when evaluating surveillance systems;
- records should only be linked by means of a common, non-personal unique identifier (e.g. record number; encrypted identifiers are easily created in electronic systems);
- only personnel essential to the evaluation process should have access to the data; and
- specific protocols for data transfer should be used.

Given the fact that the purpose of the evaluation remains strictly focused on improved public health and not on research, individual informed consent of patients is usually not required. However, evaluation methods must be in compliance with the World Medical Association's *Declaration of Helsinki* on 'Ethical principles for medical research involving human subjects'[1].

3.5.8 Interpretation and modifications

After discrepancies and invalid data are identified and summarised, all possible causes of errors will be listed, which will facilitate the interpretation of the data. The evaluation team can then suggest modifications to increase data validity. Some errors can be corrected by going back to the data provider or patients, while contacting data providers may also provide useful information to fully understand how these errors occurred.

The process of data analysis requires experience and knowledge of the surveillance system and its components. This phase (i.e. data analysis by a team of experts) is decisive because all subsequent suggestions to modify the system depend on it. Interpretation can be difficult, and occasionally further studies may be needed to identify the actual sources of specific errors.

The acceptable level of errors depends on the disease or the variable under evaluation, the type of error (random/systematic), the impact of low validity on data interpretation, public health interventions, and the resources and capacity of a system to be modified. The acceptable level may change over time as priorities of the surveillance system change.

For example, an error regarding the age of a patient with salmonellosis has less of an impact than misclassifying a case of poliomyelitis. In the first case, the acceptable level of errors can be raised; in the second example, a validity of 100% is required.

Since most surveillance systems cannot avoid duplicates, epidemiologists usually set an acceptable level of duplicates (expressed as a percentage of duplicates in a dataset), based on the characteristics of the surveillance system and its objectives. The acceptable level of duplicates may differ among diseases. For HIV, for example, there should be validation checks during data entry to completely avoid duplicates in the dataset, while for salmonellosis the acceptable level may be set at <5% of recorded cases.

An acceptable level may be thought of as a cut-off point (or threshold) that determines whether action is required; this threshold plays a substantial role in determining the diseases and the variables (by disease) that will be

[1] The World Medical Association (WMA) has developed the Declaration of Helsinki as a statement of ethical principles for medical research involving human subjects, including research on identifiable human material and data. Available from: http://www.wma.net/en/30publications/10policies/b3/17c.pdf

addressed. A list of possible errors by disease and variable can provide important information before applying corrective measures.

Note: Cases may be mismatched or misclassified, usually reflecting missing values in the system rather than an incorrect data entry. Such issues refer to data completeness and should be differentiated as they affect data interpretation.

It may be useful to produce a short summary for all errors connected to a variable:

- Recognised errors and the frequency of occurrence
- Exact phase in the reporting process during which the errors occurred
- Possible causes of these errors
- Proposed modifications to eliminate errors

A sample table for summarising the conclusions of the validation process is presented in Table 3.5d.

Table 3.5d. Summary of the validation process

Disease/variable	Errors identified	When did the error occur?	Frequency	Cause	Suggested modification
Tuberculosis	Duplications	Data entry	5%	No validity check during data entry	Use a unique identification number. Use validity checks during data entry.
Tuberculosis/case definition	Not in compliance with clinical and laboratory data	Case classification	16.5%	• Recent change of case definition • Lack of training	Train data providers.

Errors may occur during all phases of work:

- Data collection – paper-based system
- Data collection – electronic system (data entry)
- Classification of cases (application of case definitions)
- Statistical analysis
- Transfer of data, coding – encryption

Possible causes of errors:

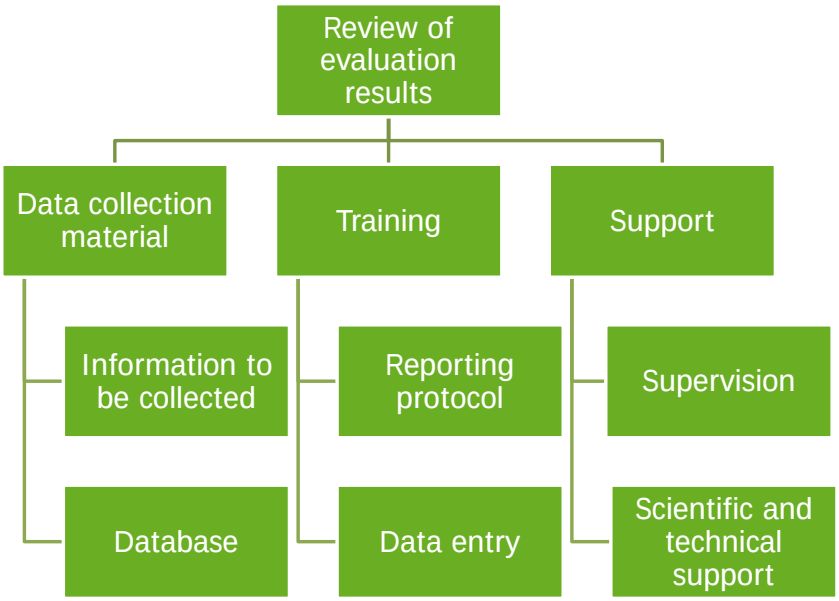
- General problems of the healthcare system (lack of time, personnel, and technical support)
- Problems with the reporting tools (forms/databases)
- Unavailability of an electronic system
- Problems with the reporting process (lack of training, recent changes in the reporting process)
- Inadequate statistical analysis of data

Possible steps to reduce errors and increase data validity:

- All healthcare services should use the same structure for their medical records.
- All local public health authorities should receive the same standard case investigation questionnaire.
- Development of best practices for record keeping and documentation at all levels of medical care (medical reports for diagnosis, laboratory reports, outcome)
- Implementation of electronic notification systems and integration of data collection tools into existing health information systems
- Revision of reporting forms and databases and validity checks
- Help functions and support for system users, either directly in the information system or online, to ensure the uniform distribution of revisions
- Systematic training of data providers at all levels; supervision and guidance
- Information technology and statistical support
- Two-pass verification, also known as 'double data entry', when entering data (useful and effective approach to identify and correct errors; costly)

Approaches that can be used to increase validity are usually connected to: a) the data collection formats, b) the training of data providers, and c) the enhancement of supportive mechanisms.

Figure 3.5c. Approaches to increase data validity in a disease surveillance system



The results of an evaluation may show that most errors occur only during a specific point of the reporting process. In this case, the implemented measures can be restricted to the precise level of the surveillance system which is actually affected. This will save both time and money.

Changes and modifications to the surveillance system should be done carefully and only when absolutely necessary. It is important to inform all stakeholders in the system of impending changes.

The final results of the evaluation need to be disseminated to the appropriate recipients. Recipients should be chosen on the basis of the proposed remediation approaches (e.g. ministry of health, local public health authorities, data providers, or hospital administrators). For the communication of results to be successful, the written report should be short and concise, with clear suggestions, and directly addressed at the target audience (language, terminology, content).

3.5.9 Questions and answers

Q. Could you give an example of a disease for which data validity is of major importance?

A. A false notification of a pulmonary anthrax case would be very unfortunate because it is a disease connected to bioterrorism. The surveillance system has to be 100% accurate in such a situation. Another example of a lack in validity would be the consistent overreporting of probable measles cases as confirmed because of a misleading or misinterpreted case definition, because this could lead to changes in vaccination policies.

Q. What is triangulation and how it is used in the evaluation of validity?

A. Triangulation is the term used to describe the technique of combining several different methodologies and different studies (data collection methods) for the purposes of evaluation.

Q. Can you list possible external data sources that can be used to validate surveillance data (data concordance)?

A. Medical records, data from reference laboratories, discharge papers, vaccination cards, death certificates, study results, results of active follow-up of cases, etc.

Q. How can duplicate entries be identified?

A. They can be identified by creating a matching string variable to identify pairs, using the last name (text variable), the date of birth (date variable), and the gender (binary variable) of each case.

Q. Which three basic parameters do you need to address in order to increase external validity?

A. Errors are usually connected to: a) data collection formats (notification forms and databases), b) training of data providers (constant changes in surveillance practices over time), and c) supervision and data management support (at all levels).

Q. Is 100% validity always required?

A. No, 100% validity is only needed for variables that directly affect the course of action that will be taken. Instead, validity is usually set to a cut-off point or an acceptable level of false reports that will be allowed while still meeting the system's target for validity.

Q. What further information do you need for each identified error?

A. You need to know when each error usually occurs, how, and why, and also how it can be corrected.

3.6 Evaluation of usefulness

The surveillance attribute 'usefulness' is defined in Section 1.2.4. As specified in the CDC guidelines for evaluation of surveillance systems, the measurement of usefulness can be inexact.

3.6.1 Introduction

The usefulness of a surveillance system can be perceived differently by different stakeholders. For example, government institutions are likely to value systems which avoid false alarms and only report verified health threats to the community, or provide data on declining disease incidence. When a public health threat occurs (e.g. an outbreak in a neighbouring country) the existence of a sound surveillance system is reassuring. The scientific community will find it useful when disease surveillance data offer new insights in disease epidemiology, or reveal new phenomena, which will help to generate new hypotheses. A health security officer will be primarily interested in a system that is able to timely detect and assess health events; a disease control officer will appreciate data that deepen the understanding of patterns of transboundary disease spread, as well as the impact of control strategies; and a health planner will appreciate reliable data that will help to prioritise and quantify the needed resources.

Evaluating the extent to which a disease surveillance system meets these criteria of usefulness can help to clarify why certain user groups have reservations towards the system. On the other hand, satisfied users of the system will offer arguments for its relevance.

3.6.2 Identifying expected uses of a disease surveillance system

Usefulness can be assessed by looking at the public health functions that are supported by the surveillance system:

- Detection of diseases or other health hazards for timely public health responses
- Estimation of disease impact
- Description of the natural history of diseases
- Establishment of disease distribution, spread modality, risk factors
- Hypotheses to stimulate research
- Results of control measures
- Guidance for public health planning
- Identification of trends and detection of outbreaks

Some of the above functions may be formally stated goals of the system; it is also possible that some functions are implicitly carried out by the system without having been added to an official list of public health functions.

Before proceeding, it should be established which public health functions, if not all, are expected to be supported by the surveillance system, and in which order of importance.

3.6.2 Creating a user list

A subsequent step is the creation of a list with all persons/projects/institutions using and/or accessing the system's surveillance data. The list can be based on documents produced by the system, monitoring reports, input from system operating staff, health authorities, funding institutions, NGOs, representatives of target populations, the private sector, academics, scientists, and published literature, and should specify the data types used by each institution (see Table 3.6a.).

Table 3.6a. Sample user list for data generated by the health surveillance system

Institution	Type of user	Data use detected
National Hygiene and Anti-Epidemic Centre	Public health agency	Signal analysis and notification of public health events
Communicable Diseases Department	Public health agency	Data analysis and publication of weekly epidemiological report
Communicable Diseases Department	Public health agency	Detection and response to [disease] outbreak in June 2009
Communicable Diseases Department	Public health agency	Health education campaign on [disease]

Institution	Type of user	Data use detected
Communicable Diseases Department	Public health agency	Planning of routine school vaccination
WHO	Public health agency	Management of [disease] pandemic
[Name] University, School of Public Health	Academic researcher	Epidemiological study on [disease] disease
[Name] University, School of Medicine	Academic researcher	Evaluation of impact of [disease] screening on specific mortality
[Name] University, School of Veterinary Medicine	Academic researcher	Study on breeding practices and zoonosis incidence in the general population
National Congress	Government (non-health sector)	Law on mandatory [disease] certification of food producers
Ministry of Finance	Government (non-health sector)	National financial plan 2012
The [name] weekly magazine	Media	Articles on health issues
[Name] Institute	Private researcher	Update on prevalence of [disease] condition, by province
[Name] Foundation	Private researcher	Study on incidence of [disease] event
[Name] Pharmaceutical Company	Private researcher	Study on incidence of [disease] disease

3.6.3 Documenting system usage

A third step can be the exploration of the various sources listed below. A documentation of system usage during a given time period can be conducted by literature searches, reviewing documents and interviewing users.

The type of evidence sought should include the following:

- Quantification of instances of use of surveillance data leading to improved public health:
  - Number of public health events that were detected through the system
  - Number of bulletins, scientific and press articles that used system data
  - Number of thematic/surveillance reports that were made available to decision-makers
  - Number of data items and indicators that were used by health planners
  - Number of data items and indicators that were used to assess disease control strategies
- Qualitative information on the use of surveillance system data by users (see Table 3.6a), for example:
  - Policies that were formulated or revised on the basis of system data
  - Urgent measures
  - Crisis management
  - Quality improvement
  - Savings in time, effort, funds
  - Accuracy of public communication (press, reports, websites, etc.)
  - Assessment of prevention and control strategies
  - Research projects
  - Academic courses
- Indication of the possibility to obtain the same data from another source, or to arrive at the same results without using system data, for example:
  - Other existing surveillance systems
  - Other means to reach equivalent results
- Identification of useless surveillance system data:
  - Data elements that were not useful
  - Reasons and factors that prevented the use of data

3.6.4 Remediation

Depending on the findings of the assessment of usefulness, different kinds of recommendations may be issued.

In usefulness was rated 'modest' due to data quality issues, corrective measures will seek to improve the key attributes as identified in the evaluation, generally including completeness, validity and sensitivity. All of these attributes potentially affect usefulness.

If the evaluation reveals a lack of visibility of the surveillance system and its data, data presentation should be improved, and dissemination should be directly related to the intended target groups. Improvements should focus on a distribution list, innovative methods of data sharing, and suggestions on how to explore the data.

If some data elements were found to be not useful now and in the future, it may be proposed to discontinue their collection, a measure that would improve simplicity and may have a positive impact on acceptability and data quality. Finally, it may be found that the system, or parts of it, is redundant because the target functions are fulfilled by other sources, or it has lost usefulness because of context changes. If resources were used ineffectively, restructuring might be the only option.

3.7 Evaluating the costs of disease surveillance

Public health surveillance is a critical component of public health. Yet implementing and operating a surveillance system consumes resources. To ensure that the public health problem under surveillance is being efficiently and effectively monitored, periodic evaluations are essential, especially when the system has been running for many years.

3.7.1 Background

In a resource-constrained environment, the estimation of the cost of the resources used to run a surveillance system should be part of the overall evaluation framework of the system [72,73].

Currently, most budgets for surveillance, and for that matter, other public health functions (e.g. case management, outbreak investigation) are allocated and managed by accountants, not epidemiologists or public health subject matter experts. Accounting offices are often under a different management or organisational substructure from the programmes they fund and support. Additionally, budgets are usually categorised into line items (e.g. personnel, equipment, travel), and financial resources are not directly linked to public health activities.

Surveillance evaluations rarely evaluate the costs of surveillance. The World Health Organization (WHO) has issued a report on the challenges in evaluating the costs and benefits of national surveillance systems [74]. Economic evaluations in public health are often conducted for public health interventions such as screening programmes, mass vaccinations, treatment campaigns, or healthcare services, but hardly ever for surveillance activities. Yet, surveillance is foundational to all of these activities.

Cost–benefit analyses can be helpful tools in the decision-making process as they address one of the main challenges of the evaluation process: they estimate a system's impact (averting deaths and disabilities), assess its social and economic benefits, while at the same time calculating the costs of surveillance. This may require the use of different approaches such as modelling or panel expertise (experts' consultation/opinion) [74].

This section focuses on estimating the baseline cost of a surveillance system. When evaluating the costs of disease surveillance, the rationale (e.g. cost comparison, formal evaluation of the entire surveillance system, financial constraints) and the objectives of the analysis should be clearly stated. An analysis plan, including a timeframe for the study and surveillance system activities for cost analyses, should also be defined. Each resource/item should be identified and a monetary value should be assigned to it.

3.7.2 Classification of costs of public health surveillance

Cost can be defined as 'as the amount of expenditure incurred on or attributable to a particular good or activity (e.g. production or service)' [75,88]. Several cost classification systems can be used to support cost estimates depending on the nature, the objectives, and the methodology of the cost analysis. Some technical terms defined below were taken from a review of literature on 'costing healthcare' service [75].

Classification according to traceability

Direct costs are the costs which can be directly linked or attributed to the use of particular resources or cost objects. A cost object can be a good, a service, or a programme. Directs cost can be materials, labour or drugs. Indirect costs have no direct relationship to the cost object. They can also be material, labour or expenses. For instance, the cost of catering or cleaning in a hospital is usually classified as an indirect cost. Indirect costs are also called overhead costs.

Classification by type of cost behaviour

According to this classification there are two major types of costs: variable and fixed costs. Variable costs change with the volume of total activities (number of patients treated). Fixed costs do no vary with the level of activity or service output.

Classification by frequency of expenditures

This classification identifies two types of cost: capital costs and recurrent costs.

Recurrent costs are consumed within one financial year while capital costs are the costs of acquiring fixed assets that last more than one financial year.

Classification in health programme evaluation

Several authors have suggested classifying health programme costs as follows [75]:

- Costs related to healthcare services
- Costs not related to healthcare services

A similar suggestion establishes four categories:

- Programme cost (full cost of providing a particular service)
- Non-programme costs (including cost of uncovered healthcare services and any resulting effects such as savings)
- User (patient) and family costs
- Productivity costs

In the first classification, health services cost encompass cost of illness, future costs and direct medical costs, which include overheads and capital costs.

3.7.3 Cost analysis

In order to perform a cost analysis in the context of evaluating a public health surveillance system the following steps are recommended:

Step 1. Describe perspective

The first step of an economic analysis is to describe the perspective of the cost evaluation, outlining the categories which need to be considered. A cost analysis can be performed from a patient, employer, private/public purchaser, government, or societal perspective. The latter looks at all the costs relevant to society as a whole and is the perspective most frequently applied in health policy [74,75]. Based on the perspective chosen, the evaluation will focus either on direct costs (i.e. resources directly required to operate the system) or will include an estimation of indirect costs (e.g. productivity or travel). In a public health context, the direct costs of surveillance activities [1,2] are usually covered by state funds administered by the national ministries of health.

Step 2. Define scope

Another important step in the cost analysis of surveillance is to define the included activities. Public health surveillance is the systematic ongoing assessment of the health of a community based on the collection, interpretation and use of health information and data, and it provides information essential to public health decision-making [76]. Buehler JW et al. have designed a conceptual framework of public health surveillance which categorises surveillance activities into eight core functions and four supporting activities, including the provision of resources. This framework is a good starting point to list the activities of a surveillance system and their respective costs [77] (Table 3.8a).

Activities that relate to response and preparedness need to be included in the inventory of surveillance costs. As public health surveillance aims to control or prevent major outbreaks, it seems relevant to include these activities in the cost analysis [74,77].

Step 3. Identify resources

The total cost of surveillance includes all resources necessary to operate the system, and a checklist of the basic activities (+/- response) can be developed. Alternatively, a modelling approach can be developed in which all events that could occur in a surveillance and response system should be considered.

In a study aimed at estimating the cost of an integrated disease surveillance and response system for priority infectious diseases in Africa, cost estimations were performed for each of the core functions of surveillance (detection, reporting, analysis, investigation, response and feedback) and for support functions (training, supervision, communication). For each of these activities, the resources were grouped into seven major categories such as personnel, laboratory or transportation [78].

Another study performed in Mexico aimed at comparing two alternative logic models for malaria surveillance. The authors used a logic model approach integrating the following steps: programme, action, activities, tasks and input. Actions were the interventions necessary to complete the programme, activities were the concrete procedures to carry them out, and tasks were the individual units of work (smallest division for each activity). Inputs were the necessary resources to perform each task [79].

Step 4. Attach monetary value to resources

Different methods exist for assigning a monetary value to each resource or item in surveillance programmes [75]. Direct costs are used when data are available from the surveillance programme budget. An ad hoc study could obtain additional information on specific elements (e.g. time spent by personnel on a specific activity).

The following thee approaches are possible:

- The bottom-up approach (micro- or activity-based costing) estimates the cost of each resource or item used. Assigning a financial cost to each resource unit or item is the simplest way to estimate cost. The cost of an activity would then be equal to the resource unit cost multiplied by the quantity of resources used. The overall costs of surveillance would amount to the sum of the different activities valued. If the goal is to compare costs between countries, monetary amounts have to be first converted to a common currency. This approach can be time-consuming and expensive.
- The top-down approach calculates the surveillance costs at a higher level than the surveillance programme itself. This global cost is then broken down at the level of the surveillance programme by allocation (see below). Although cheaper than the first approach, the top-down approach is less precise.
- The mixed approach can either use one of the above-mentioned techniques for specific activities (e.g. activity-based costing for laboratory supplies and top-down approach for personnel).

When the direct measurement of costs is not possible because of insufficient data, standard costs could be used based on WHO estimates [80].

3.7.4 Discussion

Practical remarks

One common problem is how to allocate overhead (indirect) costs. Resource unit measurements should be chosen carefully. For example, personnel wages are broken down into hours spent in the programme. The cost of utilities and maintenance could be estimated in relation to the size (occupied space) of the surveillance offices.

Different allocation methods exist [75,76]. Direct allocation methods proportionally allocate the final total cost of surveillance activities to a specific programme by putting total personnel time (or total occupied space) in relation to personnel time devoted to the surveillance programme (or space occupied by the programme).

Another issue of concern is capital costs (e.g. buildings, vehicles, office equipment). Buildings and vehicles have a long life span (e.g. 50 years for a building, five years for a vehicle) and their value depreciates over time. The costs are usually annualised using a formula that takes into account the depreciation and the opportunity cost of investment. For the analysis, economic guidelines recommend that the total costs should be adjusted to the year used in the analysis, and costs incurred in the previous year have to be discounted (usually at 3%) [74,81]. Unavailable costs have to be assumed. Sensitivity analyses could be used to assess the effect of these assumptions on the cost estimates. Likewise, if samples are selected to estimate costs, an extrapolation to a larger population may need statistical adjustment (at 95% confidence interval).

Costs versus benefits

Public health authorities may have to choose between alternative ways of spending money and maximising the efficiency of programmes. Consequently, determining the benefits of surveillance is another important aspect of cost evaluation. An analysis of the costs and benefits of surveillance 'requires an estimate of the cost of illness and [should] answer the question of how many cases to a particular disease need to be prevented by the surveillance system to be exactly equal to the expenditure on the system' [82]. Benefits can also be evaluated by estimating the number of cases or deaths averted by public health intervention. This can be achieved by using logic models or calculating the outbreak case fatality ratio and subtracting the case fatality ratio in the absence of treatment.

Few studies have been conducted in this area, and they mostly address surveillance and response systems [82,83]. A study in the United Kingdom [83] has estimated the cost–benefit ratio of an early intervention for controlling a nationwide outbreak of salmonellosis. According to this study, the investigation yielded a 3.5-fold rate for the public sector and a 23.3-fold rate for the society. Another study estimated the cost-effectiveness of the integrated disease surveillance and response for meningitis in Burkina Faso. Assuming that all benefits observed after five years of programme involvement were actually due to the programme, they compared health outcomes such as prevented deaths or the number of outbreaks before and after the programme. Results indicated that the surveillance and response programme reduced the number of meningitis cases by 43/100 000 population [84]. The cost-effectiveness of the programme was USD 23 per case averted and 98% per meningitis-related death averted. Usually, estimating the direct benefits of a surveillance programme is not as unequivocal, in particular for surveillance systems which employ no immediate control measures, e.g. HIV/AIDS surveillance.

Another study compared the cost and performance of two surveillance systems for tuberculosis control [85]. By developing a checklist of standardised performances indicators and estimating the costs for each of the two surveillance and control programmes, the authors were able to calculate a performance score measured against costs. Figures could then be mapped in a four-quadrant display showing performance and cost. Benchmarking, as demonstrated in this study, can be a helpful tool when dealing with funding requests. Ideally, one should always compare two types of surveillance programmes to identify the best cost option. A study conducted in Mexico did exactly that: the costs of two surveillance schemes for malaria were compared using different laboratory tests [86].

Cost analyses published in the literature mainly aim at evaluating public health interventions or treatment options. Even if surveillance activities are part of these interventions, surveillance costs are rarely integrated. Guidelines for the evaluation of a surveillance system should include cost estimations as a criterion, and systems should be evaluated with regard to their objectives and usefulness. A recent literature review on the evaluation of surveillance systems on veterinary and human public health has shown that only five (of 99) studies – two in veterinary health and three in human health – have included cost in the evaluation [87].

Due to budgetary constraints in many European public health agencies, cost–benefit analyses for surveillance will become increasingly important.

Table 3.8a. Conceptual framework of public health surveillance

Activity	Comprising…
Core activities for surveillance	Detection
	Registration
	Confirmation
	Reporting
	Analysis
	Feedback
Public health action	Acute (epidemic-type) response
	Planned (management-type) response
Support activities	Communication
	Training
	Supervision
	Resource provision

Source: McNabb et al. [7]

3.8 References

1. Erhart A, Thang ND, Xa NX, Thieu NQ, Hung LX, Hung NQ, et al. Accuracy of the health information system on malaria surveillance in Vietnam. Trans R Soc Trop Med Hyg. 2007 Mar;101(3):216-25.
2. Chadee DD. Evaluation of malaria surveillance in Trinidad (1988–1998). Ann Trop Med Parasitol. 2000 Jun;94(4):403-6.
3. Gazarian M, Williams K, Elliott E, Chant K, Longbottom H, Mellis C, et al. Evaluation of a national surveillance unit. Arch Dis Child. 1999 Jan;80(1):21-7.
4. Baldissera S, Campostrini S, Binkin N, Minardi V, Minelli G, Ferrante G, et al. Features and initial assessment of the Italian behavioral risk factor surveillance system (PASSI), 2007-2008. Prev Chronic Dis. 2011 Jan;8(1):A24.
5. Arscott-Mills S, Holder Y, Gordon G, System JIS. Comparative evaluation of different modes of a national accident and emergency department-based injury surveillance system: Jamaican experience. Inj Control Saf Promot. 2002 Dec;9(4):235-9.
6. Ritzwoller DP, Kleinman K, Palen T, Abrams A, Kaferly J, Yih W, et al. Comparison of syndromic surveillance and a sentinel provider system in detecting an influenza outbreak – Denver, Colorado, 2003. MMWR Morb Mortal Wkly Rep. 2005 Aug;54 Suppl:151-6.
7. World Health Organization. Assessment of epidemiological disease surveillance system in Mozambique, 13 Nov–4 Dec 2006. [no place]; 2006.
8. Watkins RE, Martin PAJ, Kelly H, Madin B, Watson C. An evaluation of the sensitivity of acute flaccid paralysis surveillance for poliovirus infection in Australia. BMC Infect Dis. 2009 Sep 30;9:162.
9. World Health Organization, Regional Office for South-East Asia. Comprehensive assessment of national surveillance systems in Sri Lanka. Joint assessment report 4–13 March 2003. New Delhi: WHO Regional Office for South-East Asia; 2003.
10. Furuno JP, Schweizer ML, McGregor JC, Perencevich EN. Economics of infection control surveillance technology: cost-effective or just cost? Am J Infect Control. 2008 Apr;36(3 Suppl):S12-7.
11. Van Benthem BH, van Vliet JA. Reflections on an evaluation of the Dutch Infectious diseases Surveillance Information System. Euro Surveill. 2008 Mar;13(11). PubMed PMID: 18768126.
12. Somda ZC, Meltzer MI, Perry HN, Messonnier NE, Abdulmumini U, Mebrahtu G, et al. Cost analysis of an integrated disease surveillance and response system: case of Burkina Faso, Eritrea, and Mali. Cost Eff Resour Alloc. 2009;7:1.
13. Chotivichien S, Tharmaphornpilas P, Sinawat S. Evaluation of growth monitoring and promotion (GMP) surveillance system Thailand 2002. Asia Pac J Public Health. 2006;18(2):42-5.
14. Davila JC, Payne DC, Zhang Y, Rose CE, Aranas A, Hill AN, et al. Quality assessment of nonanthrax vaccination data in the Defense Medical Surveillance System (DMSS), 1998–2004. Vaccine. 2008 Mar;26(12):1577-84.
15. Freitas FT, de Sena Neto SA, Grossi MA, Macário EM, da Nóbrega AA. Evaluation of leprosy surveillance system in the Amazon region, Brazil, 2001–2007. Lepr Rev. 2012 Mar;83(1):52-63.
16. Sekhobo JP, Druschel CM. An evaluation of congenital malformations surveillance in New York State: An application of Centers for Disease Control and Prevention (CDC) guidelines for evaluating surveillance systems. Public Health Reports. 2001 Jul–Aug;116(4):296-305.
17. Ang BC, Chen MI, Goh TL, Ng YY, Fan SW. An assessment of electronically captured data in the patient care enhancement system (PACES) for syndromic surveillance. Ann Acad Med Singapore. 2005 Oct;34(9):539-4.
18. Parrella A, Dalton CB, Pearce R, Litt JCB. ASPREN surveillance system for influenza-like illness – A comparison with FluTracking and the National Notifiable Diseases Surveillance System. Aust Fam Physician. 2009 Nov;38(11):932-6.
19. Huaman MA, Araujo-Castillo RV, Soto G, Neyra JM, Quispe JA, Fernandez MF, et al. Impact of two interventions on timeliness and data quality of an electronic disease surveillance system in a resource-limited setting (Peru): a prospective evaluation. BMC Med Inform Decis Mak. 2009;9:16.
20. Peterson KE, Hacek DM, Robicsek A, Thomson RB, Peterson LR. Electronic surveillance for infectious disease trend analysis following a quality improvement intervention. Infect Control Hosp Epidemiol. 2012 Aug;33(8):790-5.

21. Assessment of the national communicable disease surveillance and response system, Ethiopia. Wkly Epidemiol Rec. 2001 Jan;76(2):9-16.

22. Reeder B, Revere D, Olson DR, Lober WB. Perceived usefulness of a distributed community-based syndromic surveillance system: a pilot qualitative evaluation study. BMC Res Notes. 2011;4:187.

23. Meynard JB, Chaudet H, Varcol C, Decker E, Andresen S, Holterm HU, et al. Evaluation of the NATO Disease Surveillance System by its users in Kosovo in 2008. Mil Med. 2010 Jul;175(7):466-8.

24. Soto G, Araujo-Castillo RV, Neyra J, Fernandez M, Leturia C, Mundaca CC, et al. Challenges in the implementation of an electronic surveillance system in a resource-limited setting: Alerta, in Peru. BMC Proc. 2008;2 Suppl 3:S4.

25. Klevens RM, Fleming PL, Li JM, Gaines CG, Gallagher K, Schwarcz S, et al. The completeness, validity, and timeliness of AIDS surveillance data. Ann Epidemiol. 2001 Oct;11(7):443-9.

26. Hope KG, Merritt TD, Durrheim DN, Massey PD, Kohlhagen JK, Todd KW, et al. Evaluating the utility of emergency department syndromic surveillance for a regional public health service. Commun Dis Intell. 2010 Sep;34(3):310-8.

27. Abbott C, Tiede B, Armah G, Mahmoud A. Evaluation of cost-effectiveness of live oral pentavalent reassortant rotavirus vaccine introduction in Ghana. Vaccine. 2012 Mar;30(15):2582-7.

28. Guerra J, Mayana B, Djibo A, Monzo ML, Llosa AE, Grais RF. Evaluation and use of surveillance system data toward the identification of high-risk areas for potential cholera vaccination: a case study from Niger. BMC Res Notes. 2012 May;5(1):231.

29. Betanzos-Reyes AF, Rodríguez MH, Duran-Arenas LG, Hernández-Avila JE, Méndez-Galván JF, Monroy OJ, et al. Comparative analysis of two alternative models for epidemiological surveillance in the Mexican Malaria Control Program. Health Policy. 2007 Mar;80(3):465-82.

30. Decker MD. Continuous quality improvement. Infect Control Hosp Epidemiol 1992;13(3):165-9.

31. Mor Z, Migliori GB, Althomsons SP, Loddenkemper R, Trnka L, Iademarco MF. Comparison of tuberculosis surveillance systems in low-incidence industrialised countries. Eur Respir J 2008;32(6):1616-24.

32. Holman N, Lewis-Barned N, Bell R, Stephens H, Modder J, Gardosi J, et al. NHS Diabetes in Pregnancy Dataset Development Group – Development and evaluation of a standardized registry for diabetes in pregnancy using data from the Northern, North West and East Anglia regional audits. Diabet Med 2011;28(7):797-804.

33. Centers for Disease Control and Prevention. Assessing completeness of perinatal hepatitis B virus infection reporting through comparison of immunization program and surveillance data – United States. MMWR Morb Mortal Wkly Rep 2011:60(13):410-13.

34. Salemi JL, Tanner JP, Kennedy S, Block S, Bailey M, Correia JA, et al. A comparison of two surveillance strategies for selected birth defects in Florida. Public Health Rep 2012;127(4):391-400.

35. Centre for Surveillance Coordination Population and Public Health Canada: Framework and Tools for Evaluating Health Surveillance Systems. Ottawa: 2004. Available from: http://publications.gc.ca/collections/Collection/H39-4-46-2004E.pdf.

36. Romaguera RA, GRKD: Evaluating public health surveillance. In Principles and practice of public health surveillance. Edited by: Teutsch SM, Churchill RE. New York: Oxford University Press. 2000:176-93.

37. German RR, Lee LM, Horan JM, Milstein RL, Pertowski CA, Waller MN. Updated guidelines for evaluating public health surveillance systems: recommendations from the Guidelines Working Group. MMWR Recomm Rep 2001;50:1-35.

38. Doyle TJ, Glynn MK, Groseclose. Completeness of notifiable infectious disease reporting in the United States: An analytic literature review. Am J Epidemiol 2002;155:866-74.

39. Sekar C and Deming EW. On a method of estimating birth and death rates and extent of registration. J Am Stat Assoc 1949;44:101-15.

40. Le Cren ED. A note on the history of mark-recapture population estimates. J. Animal Ecol 1965;34:453-4.

41. Pollock KH. Modeling capture, recapture and removal statistics for estimation of demographic parameters for fish and wildlife populations: past, present and future. J Am Stat Assoc 1991;86:225-38.

42. McClish D, Penberthy L. Using Medicare data to estimate the number of cases missed by a cancer registry: a 3-source capture–recapture model. Medical Care 2004;42:1111–6.

43. Peragallo MS, Urbano F, Lista F, Sarnicola G, Vecchione A. Evaluation of cancer surveillance completeness among the Italian army personnel, by capture–recapture methodology. Cancer Epidemiol 2011;35(2):132-8.

44. Tilling K, Sterne JA, Wolfe CD. Estimation of the incidence of stroke using a capture–recapture model including covariates. Int J Epidemiol 2001;30:1351–9.

45. Mahr A. et al. Prevalence of polyarteritis nodosa, microscopic polyangiitis, Wegener's granulomatosis, and Churg-Strauss syndrome in a French urban multiethnic population in 2000: a capture–recapture estimate. Arthritis Rheum 2004;51:92–9.

46. Dunbar R, van Hest R, Lawrence K, Verver S, Enarson DA, Lombard C, et al. Capture–recapture to estimate completeness of tuberculosis surveillance in two communities in South Africa. Int J Tuberc Lung Dis 2011;15(8):1038-43.

47. Bassili A, Grant AD, El-Mohgazy E, Galal A, Glaziou P, Seita A, et al. Estimating tuberculosis case detection rate in resource-limited countries: a capture–recapture study in Egypt. Int J Tuberc Lung Dis 2010;14(6):727-32.

48. van Hest NA, Smit F, Baars HW, De Vries G, De Haas PE, Westenend PJ, et al. Completeness of notification of tuberculosis in the Netherlands: how reliable is record-linkage and capture–recapture analysis? Epidemiol Infect 2007;135(6):1021-9.

49. Schrauder A, Claus H, Elias J, Vogel U, Haas W, Hellenbrand W. Capture–recapture analysis to estimate the incidence of invasive meningococcal disease in Germany, 2003. Epidemiol Infect 2007;135(4):657-64.

50. de Greeff SC, Spanjaard L, Dankert J, Hoebe CJ, Nagelkerke N, de Melker HE. Underreporting of meningococcal disease incidence in the Netherlands: results from a capture–recapture analysis based on three registration sources with correction for false positive diagnoses. Eur J Epidemiol 2006;21(4):315-21.

51. Lievre L, Deveau C, Gerbe J, Enel P, Tran L, Castro Nd, et al: Yearly number of patients diagnosed with primary HIV-1 infection in France estimated by a capture–recapture approach. AIDS 2006;20:2392-4.

52. Huang WT, Huang WI, Huang YW, Hsu CW, Chuang JH. The reporting completeness of a passive safety surveillance system for pandemic (H1N1) 2009 vaccines: a capture–recapture analysis. Vaccine 2012;30(12):2168-72.

53. Vong S, Goyet S, Ly S, Ngan C, Huy R, Duong V, et al. Under-recognition and reporting of dengue in Cambodia: a capture–recapture analysis of the National Dengue Surveillance System. Epidemiol Infect 2012;140(3):491-9.

54. Staat MA, Rice MA, Donauer S, Payne DC, Bresee JS, Mast TC, et al. Estimating the rotavirus hospitalization disease burden and trends, using capture–recapture methods. Pediatr Infect Dis J 2010;29(12):1083-6.

55. Cathcart SJ, Lawrence J, Grant A, Quinn D, Whitty CJ, Jones J, et al. Estimating unreported malaria cases in England: a capture–recapture study. Epidemiol Infect 2010;138(7):1052-8.

56. Van Hest NA, Hoebe CJ, Den Boer JW, Vermunt JK, Ijzerman EP, Boersma WG, et al. Incidence and completeness of notification of Legionnaires' disease in The Netherlands: covariate capture–recapture analysis acknowledging regional differences. Epidemiol Infect 2008;136(4):540–50.

57. Gallay A, Vaillant V, Bouvet P, Grimont P, Desenclos JC. How many foodborne outbreaks of salmonella infection occurred in France in 1995? Application of the capture–recapture method to three surveillance systems. Am J Epidemiol 2000;152:171–7.

58. McCarty DJ, Tull ES, Moy CS, Kwoh CK, LaPorte RE. Ascertainment corrected rates: application of capture-recapture methods. Int J Epidemiol 1993;22:559–65.

59. Brenner H. Use and limitations of the capture-recapture method in disease monitoring with two dependent sources. Epidemiol 1995;6:42–8.

60. Hook EB, Regal RR. Capture–recapture methods in epidemiology: methods and limitations. Epidemiol Rev 1995;17:243–64.

61. Hook EB, Regal RR. Effect of variation in probability of ascertainment by sources ('variable catchability') upon 'capture recapture' estimates of prevalence. Am J Epidemiol 1993;137:1148-66.

62. Cormack RM. Log-linear models for capture-recapture. Biometrics 1989; 45:395–413.

63. Alho JM. Logistic regression in capture–recapture models. Biometrics 1990;46:623–35.

64. Tilling K, Sterne JAC. Capture–recapture models including covariate effects. Am J Epidemiol 1999;149:392–400.

65. World Health Organization. Communicable disease surveillance and response systems – Guide to monitoring and evaluating: Geneva: WHO; 2006. Available from: http://www.who.int/csr/resources/publications/surveillance/WHO_CDS_EPR_LYO_2006_2.pdf

66. McCormick A. The notification of infectious diseases in England and Wales. Communicable Disease Report. CDR Review 1993;3(2):R19-25

67. British Paediatric Surveillance Unit. 19th Annual Report 2004–05. London: Royal College of Paediatrics and Child Health; 2005. Available from: http://www.bpsu.inopsu.com/publications/annual_reports/annualreport_2005.pdf

68. Effler P, Ching-Lee M, Bogard A, Ieong MC, Nekomoto T, Jernigan D. Statewide system of electronic notifiable disease reporting from clinical laboratories: comparing automated reporting with conventional methods. JAMA 1999;282:1845-50.

69. Panackal AA, M'ikanatha NM, Tsui FC, McMahon J, Wagner MM, Dixon BW, et al: Automatic electronic laboratory-based reporting of notifiable infectious diseases at a large health system. Emerg Infect Dis 2002;8:685-91.

70. Ward M, Brandsema P, van SE, Bosman A. Electronic reporting improves timeliness and completeness of infectious disease notification, the Netherlands, 2003. Euro Surveill 2005;10:27-30.

71. Wurtz R, Cameron BJ: Electronic laboratory reporting for the infectious diseases physician and clinical microbiologist. Clin Infect Dis 2005;40:1638-43.

72. Romaguera RA, German RR, Klauck DN. Evaluating public health surveillance. In: Principles and practice of public health surveillance. Teutsch S, Churchill RE, editors. 2nd edition. New York: Oxford University Press; 2000: 176-193.

73. German RR, Lee LM, Horan JM, Milstein RL, Pertowski CA, Waller MN. Updated guidelines for evaluating public health surveillance systems: recommendations from the Guidelines Working Group. MMWR Recomm Rep. 2001 Jul 27;50(RR-13):1-35; quiz CE1-7.

74. Evaluating the costs and benefits of national surveillance and response systems – Methodologies and options. Geneva: WHO; 2005. Available from: http://www.who.int/csr/resources/publications/surveillance/WHO_CDS_EPR_LYO_2005_25.pdf

75. Mogyorosy Z, Smith P. The main methodological issues in costing health care services. A literature review. York, UK: Centre for Health Economics at Alcuin College University of York; 2005. Available from: http://www.york.ac.uk/che/pdf/rp7.pdf

76. Drummond F, Stoddart G, Torrance G. Methods for the economic evaluation of health care programmes. New York: Oxford University Press; 1997.

77. Teutsch S. Considerations in planning a surveillance system. In: Principles and practice of public health surveillance. Teutsch S, Churchill RE, editors. New York: Oxford University Press; 2000: 17-29.

78. McNabb SJ, Chungong S, Ryan M, Wuhib T, Nsubuga P, Alemu W, et al: Conceptual framework of public health surveillance and action and its application in health sector reform. BMC Public Health 2002, 2:2.

79. Somda ZC, Meltzer MI, Perry HN, Messonnier NE, Abdulmumini U, Mebrahtu G, et al. Cost analysis of an integrated disease surveillance and response system: case of Burkina Faso, Eritrea, and Mali. Cost effectiveness and resource allocation 2009, 7:1.

80. World Health Organization. Cost effectiveness and strategic planning (WHO-CHOICE). [Internet]. 2012 [cited 2014 Sep 12]. Available from: http://www.who.int/choice/costs/en/.

81. World Health Organization. WHO guide to cost effectiveness analysis. Geneva: WHO; 2003. Available from: http://www.who.int/choice/publications/p_2003_generalised_cea.pdf.

82. Nsubuga P, White ME, Thacker SB, Anderson MA, Blount SB, Broome CV, et al. Public Health Surveillance – A tool for targeting and monitoring interventions. In: Disease Control Priorities in Developing Countries. Jamison DT, Breman JG, Measham AR, Alleyne G, Claeson M, Evans DB, et al. Washington. DC: International Bank for Reconstruction and Development/The World Bank Group; 2006.

83. Roberts JA, Sockett PN, Gill ON. Economic impact of a nationwide outbreak of salmonellosis: cost-benefit of early intervention. BMJ. 1989 May 6;298(6682):1227-30..

84. Somda ZC, Perry HN, Messonnier NR, Djingarey MH, Ki SO, Meltzer MI. Modeling the cost-effectiveness of the integrated disease surveillance and response (IDSR) system: meningitis in Burkina Faso. PloS one 2010, 5(9).

85. Phillips VL, Teweldemedhin B, Ahmedov S, Cobb J, McNabb SJ: Evaluation of program performance and expenditures in a report of performance measures (RPM) via a case study of two Florida county tuberculosis programs. Evaluation and program planning. 2010, 33(4):373-378.

86. Betanzos-Reyes AF, Rodriguez MH, Duran-Arenas LG, Hernandez-Avila JE, Mendez-Galvan JF, Monroy OJ, et al. Comparative analysis of two alternative models for epidemiological surveillance in the Mexican malaria control program. Health policy (Amsterdam, Netherlands) 2007, 80(3):465-482.

87. Drewe JA, Hoinville LJ, Cook AJ, Floyd T, Stark KD. Evaluation of animal and public health surveillance systems: a systematic review. Epidemiol Infect. 2012 Apr;140(4):575-90.

88. Lucey T. Costing. 7th edition. Stamford: Cengage Learning; 2009.

4 Intervention

Often, surveillance systems are unable to improve data quality without a series of interventions and modifications.

4.1 Motivation of data providers for improving data quality

Despite their apparent professionalism, some data providers are less forthcoming than others. Motivating your counterparts is an essential skill to secure high-quality data.

Figure 4.1a. Key points

- Improving data quality is necessary as public health information systems inform the public, policy makers and administrators. Therefore, it is important to know to what extent data can be trusted.
- Five components can increase the willingness of data providers to collect valid and reliable data: explanations, consultations, collaboration, ergonomics, and incentives.

4.1.1 Why improve data quality?

Data from health information systems inform the public, policy makers, administrators and healthcare workers about the distribution and determinants of health conditions and the impact of interventions. Based on these data, a series of decisions are made [1]. It is important to be aware of the strengths and weaknesses of the data, and to know to which extent the data can be trusted before making a decision. Only good-quality data deserve the trust of decision-makers [2].

4.1.2 How to motivate data providers to deliver good quality data

It is critical that data providers know how collected data are used and processed before the data collection process is implemented.

Give explanations

Surveillance system staff should explain the system's objectives, what will be done with the data, and how better data can lead to a series of public health benefits. Data providers are likely to take more care in the data collection process when it is clear that the data are used for a meaningful purpose. Holding regular meetings and training courses will reinforce their commitment and help to develop skills needed to better collect and interpret data.

Data providers need to know their role and responsibility in the data collection process. Their role needs to be clearly defined in a document which describes job responsibilities, data collection procedures, regular training activities, and concepts of data interaction. A good example can be found on the website of the Dutch National Institute for Public Health and the Environment (RIVM) under 'Preventie van ziekenhuisinfecties door surveillance' (www.prezies.nl), which offers protocols, procedures for hospital validation visits, and definitions of hospital infections [9]. The general idea is to motivate data providers and show them that they are a vital link in the chain and, as such, hold a responsibility for the entire surveillance system.

Consult and collaborate

It is of great importance that system operators collaborate actively with their data providers. Before requesting information, data recipients should also consider 1) which data can be easily and quickly retrieved by the data providers, 2) how often data should be sent when considering relevance and practicability, and 3) the data quality level required to meet surveillance objectives. For some analyses, an estimated onset-of-disease date may be just as valuable as the precise – but harder to get – date for onset of disease. Furthermore, one should carefully consider the moment when data are requested. Requests received when the case report is already closed, for instance for 'date of death', may result in inefficient provision of data.

Regular (validation) meetings with data providers are a good way to discover which part of the data collection process is working out and which not. These meetings can also be used to give the data providers some feedback on data quality and improvement options. Data collection requires planning, constant attention, and fine-tuning, which in turn requires monitoring, error checking, and training. It is important that data collection procedures reflect all possible changes.

Providing guidance to data providers will increase their commitment and ability to provide quality data. This can be achieved with meetings, hands-on demonstrations, tutorials, software help texts, frequently asked questions, personal contact points (at least two), or online helpdesks.

Ergonomics

Once data providers are willing to send their data, data recipients should facilitate the process as much as possible. The reporting process should be simple, fast and efficient. Data providers will provide data more willingly if it is not overly time-consuming. One good way to achieve this is through computerised, web-based surveillance data entry, which accelerates data collection and improves timeliness. A well-designed and appealing graphical user interface may encourage data providers to use the provided tools more often.

Resources

The only way to get good quality data is to make data collection a priority, e.g. by hiring extra staff, and allocating additional time and money. Data providers cannot be expected to monitor data quality in addition to their normal duties if resources are insufficient.

Incentives

Incentives could be deployed to ensure further participation. In the Dutch sentinel influenza surveillance scheme, general practitioners receive a financial bonus which increases with the number of samples sent. In the Dutch STI surveillance scheme for high-risk patients, financial compensation is restricted to the notification of cases: the STI clinics only receive compensation when a high-risk patient is reported. In the Belgian National Reference Centres Network, reference centres receive additional funding in order to gather more data and relay them to the relevant authorities.

Another way to motivate data providers is to share the gathered data. Summary statistics, graphics, quality indicators, reports can be send back to the data providers, perhaps along with some personalised advice on methods and tools to gather, store, process and supply data [3].

In some scenarios, money can actually worsen data quality, for example if funding is suddenly stopped and the data provider's motivation drops. If financial incentives are merely an external motivation to collect data, data quality usually suffers.

4.2 Improving surveillance system quality

The quality of a disease surveillance system depends on constant improvements.

Figure 4.2a. Key points

- For real change to occur, all people involved in surveillance need to be willing to improve their skills and accept changes. Both data providers and society need to commit to the objective of improving surveillance.
- Implementing a continuous monitoring process is a good way to regularly review the surveillance process.
- Assessing the quality of a surveillance system is easier if it is computerised and web-based.

4.2.1 How to motivate for surveillance system improvement

Improving surveillance systems is a process that involves assessment, action planning, implementation, and monitoring and evaluation. For real change to occur, many people involved in surveillance need to change their behaviours. In order for these interested parties to change, they must first commit to the objective of improving surveillance [4].

How to commit data providers

The following actions can strengthen the commitment of data providers to improve the surveillance system.

1. Recognising the benefits of improved surveillance

Data providers may need to be persuaded to conduct an assessment of their surveillance system or data collection process. It is essential to help them understand the importance of good surveillance systems. Conducting an assessment is the first step to improve quality.

2. Getting the right team

The local assessment team plays a central role in conducting the assessment and moving from assessment results to action planning and implementation. Having skilled and competent people involved from the beginning will help pave the road for implementation.

3. Turning assessment team members into surveillance advocates

Local assessment team members can be trained to advocate the improvement of surveillance systems to a critical audience of colleagues. They should strengthen their data providers' sense of ownership by explaining the assessment, how it is part of a process to improve national surveillance, and describe the benefits of a well-functioning system.

4. Carefully considering all feedback
Assessment team members should carefully consider the feedback received from local data providers when reviewing assessment data, drawing conclusions, and making recommendations. Incorporating such ideas helps to build the data providers' ownership of assessment recommendations.

5. Identifying opportunities to improve national surveillance systems
In presenting preliminary findings to a broad group of stakeholders at a post-assessment workshop, assessment team members must show their audience the specific problems uncovered by the assessment and opportunities for improvement.

6. Securing commitment to action
During the post-assessment workshops, key actors in the surveillance system should be encouraged to commit themselves to actions that support improving the surveillance system.

How to commit society

Improving data quality requires commitment from all professionals who collect disease surveillance data. Individuals are usually reluctant to compromise their privacy, even if society benefits from this. To prevent that individual health information becomes unavailable to public health practitioners, the state must ensure that privacy information is safeguarded and national surveillance is not unnecessarily intrusive. If society wants to benefit from public health surveillance capacities, individuals should go beyond their distrust of the government and focus on the administration's ability to ensure the confidentiality of information.

Sweden and Great Britain are examples of countries which use a national health identification number to link various databases but at the same time have been able to ensure individual privacy and data integrity [5].

Another possibility is to ask the general public to support surveillance. In the Netherlands, people can voluntarily register influenza-like symptoms on a dedicated surveillance website (https://www.degrotegriepmeting.nl).

4.2.2 Reinforcing monitoring

Monitoring of surveillance systems can be reinforced by implementing a continuous monitoring process, which is both costly and time-consuming but will eventually yield research questions, which in turn will bring about improvements in the system. The process which reinforces the monitoring of the surveillance system contains six steps (Figure 4.2b.).

Risk assessment

Data should be reviewed on a regular basis. Data reflect how well the programme works. By critically reviewing their data, system operators can identify shortcomings. For instance, in measles surveillance, information about the vaccination status is often missing because people cannot remember if they were vaccinated. This makes it impossible to calculate the vaccination rate and difficult to interpret surveillance data on measles.

Prioritisation

Problems should be first identified and then prioritised. For instance, missing vaccination data make it impossible to calculate the vaccination rate in the various age groups. One explanation could be that municipal health services do not have access to historical data of primary healthcare clinics.

Surveillance system

Issues that need to be addressed should be treated as if they were a research project; research questions with measurable outcomes should be developed. This will ensure that changes in procedures actually lead to the desired results, i.e. did an agreement between primary healthcare clinics and municipal health services on the delivery of vaccination status data lead to fewer missing data regarding vaccination status?

Plan of action

A data analysis plan should be developed to monitor the effects of all modifications. For the above example, this could mean the production of weekly frequency tables on the vaccination status data after the implementation of the new procedure for vaccination status data.

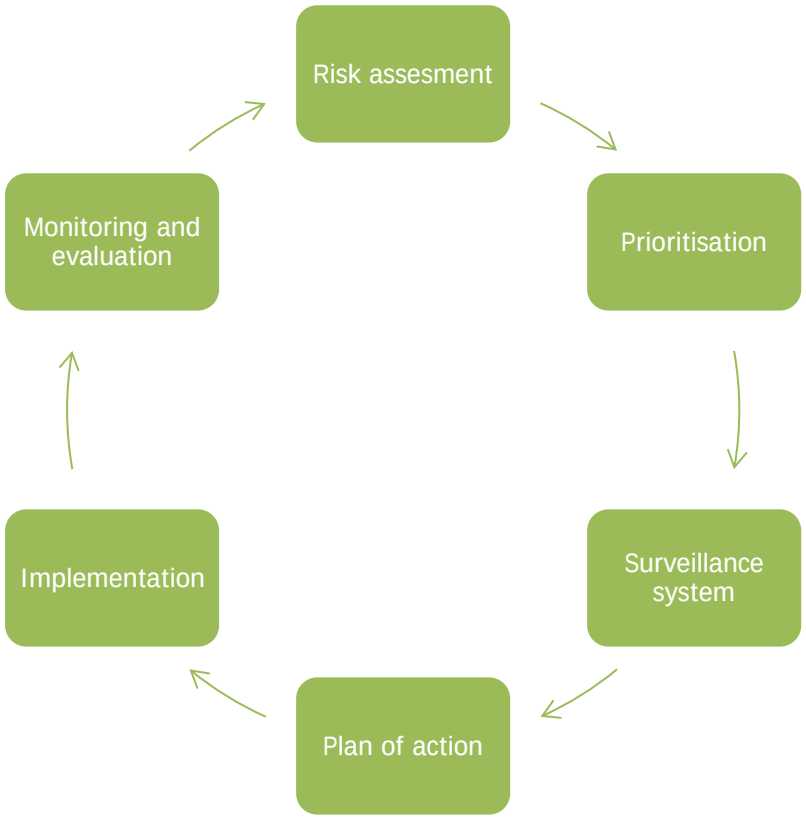
Implementation

An implementation plan should include the resources, staff and procedural changes needed for the implementation. For the above example, this would involve an agreement on data exchange between the primary healthcare clinics and municipal health services, complete with procedures describing how municipal health services can retrieve the necessary data from the clinics.

Monitoring and evaluation

Collecting more data and reviewing the data again after the modification is essential for the assessment of changes in the programme: Does the weekly report reflect a lower percentage of missing data on vaccination status? [3]

Figure 4.2b. Process to reinforce monitoring of data quality of surveillance system



The government may play an in important role in the reinforcement of monitoring national surveillance systems and in initiating activities to strengthen the local capacity to conduct evaluations (e.g. by organising training activities or providing evaluation assistance).

4.2.3 Structuring output activities

Usefulness of computerised surveillance systems

Evaluating the quality of a surveillance system is easier with computerised surveillance systems. In a computerised system, standardised checks can be implemented to detect missing data and dates. In addition, logical checks can identify incorrect or inconsistent data items. This will increase the quality of the data. Computerised data are easily accessed with statistical packages like SAS, SPSS or R, which can handle large amounts of surveillance data.

Reporting tools can be included in surveillance registration systems. The Dutch, web-based registration system for notifiable diseases, named Osiris, includes a reporting tool in which notifications that meet certain conditions can be easily selected and analysed.

Computer Assisted Search for Epidemics (CASE) is a framework for computer-supported outbreak detection, developed at the Swedish Institute for Communicable Disease Control. The system uses data from the national registry of notifiable diseases (SmiNet). The system supports several statistical algorithms and allows for different parameter settings for the various diseases. An automated analysis is performed daily. If CASE detects a potential outbreak, an email is sent to the epidemiologists in charge. CASE is available as open-source software (https://smisvn.smi.se/case/).

Advantages of an internet-based surveillance system

Faster notification process

An internet based-surveillance system will speed up the notification process and give easy access to up-to-date reports on data quality. When clinicians, laboratories or local health department investigators enter data securely over the internet, information is immediately available to state or local health departments, avoiding delays caused by mailing paper forms or data entry processing. As a result, reports are produced more quickly and can be made available sooner.

An internet-based surveillance system may improve communication on notifications, which may result in better quality of the provided information. In the Dutch online surveillance system Osiris, data recipients can add

comment fields anywhere in a surveillance questionnaire if they need to clarify the data provider's answers. The data provider can then respond by replying in another comment field.

Automatic updates of reports on quality

When high-speed processing of notifications is in place, a reporting tool can be used for the automatic updating of reports on data quality. In the Netherlands, the reporting tool Webfocus produces daily updates on notifiable diseases; a second type of report is used as quality reference for virological laboratories in virological surveillance. Furthermore, Webfocus is used to provide current epidemiological information for reports on the website of the Dutch National Institute for Public Health and the Environment (RIVM) (http://www.rivm.nl/Onderwerpen/Onderwerpen/M/Meldingsplicht_infectieziekten).

4.3 References

1. Wuhib T, Chorba TL, Davidiants V, Mac Kenzie WR, McNabb SJ. Assessment of the infectious diseases surveillance system of the Republic of Armenia: an example of surveillance in the Republics of the former Soviet Union. BMC Public Health 2002,2:3. Available from: http://www.biomedcentral.com/1471-2458/2/3

2. USAID. Performance monitoring and evaluation – TIPS – Conducting data quality assessments. Washington, DC: USAID; 2010. Available from: http://pdf.usaid.gov/pdf_docs/pnadw118.pdf

3. US Department of Education. Guide for improving NRS data quality – Procedures for data collection and training. August 2002. Available from: http://www.nrsweb.org/docs/dataquality.pdf.

4. Scribner S, Miller Franco L, Ram S. Improving surveillance systems: building support for implementation – A user's guide. Washington, DC: US Agency for International Development; [no year]. http://pdf.usaid.gov/pdf_docs/PNACM403.pdf

5. Thacker SB, Stroup DF. Future directions for comprehensive public health surveillance and health information systems in the United States. Am J Epidemiol. 1994 Sep 1;140(5):383-97.

Glossary

Carrier: a person or animal that harbours the infectious agent for a disease and can transmit it to others, but does not demonstrate signs of the disease.

Cohort: a well-defined group of persons who have had a common experience or exposure and are then followed up, as in a cohort study or prospective study, to determine the incidence of new diseases or health events.

Culture result: whether or not the test was carried out for diagnosis of the disease and the result of the test. Culture done from any specimen should be reported.

Effectiveness: the ability of an intervention or programme to produce the intended or expected results in the field.

Efficacy: the ability of an intervention or programme to produce the intended or expected results under ideal conditions.

Efficiency: the ability of an intervention or programme to produce the intended or expected results with a minimum expenditure of time and resources.

Equal catchability: when capture probabilities are homogeneous across all individuals in the population. Equal catchability is met when the probability of notification of one case to the surveillance systems is not influenced by the characteristics of the patient represented by that case.

False alarm: signal indicating the reporting of a case or an outbreak when it has not occurred.

Fit for purpose: a public health information system is fit for purpose if it enables users to make an accurate assessment or prediction of the reported diseases.

Gantt diagram: type of bar chart, developed by Henry Gantt in the 1910s, that illustrates a project schedule. Gantt charts illustrate the start and finish dates of the terminal elements and summary elements of a project. Terminal elements and summary elements comprise the work breakdown structure of the project. Modern Gantt charts also show the dependency (i.e. precedence network) relationships between activities.

Gold standard: reference dataset to which surveillance data can be compared when conducting an evaluation of a surveillance system.

High-risk group: a group of persons whose risk for a particular disease, injury, or other health condition is greater than that of the rest of their community or population.

High-risk patients: a group of patients whose risk for a particular disease, injury, or other health condition is greater than that of the rest of their community or population.

Incubation period: the time interval from exposure to an infectious agent to the onset of symptoms of an infectious disease.

Interoperability: ability of a system or a product to work with other systems or products without special effort on the part of the user.

Place of exposure: place where the infection of a reported case probably took place.

Place of residence: place of residence of a reported case at the time of disease onset.

Public health emergency: in general terms, an emergency is a sudden occurrence of a threat that, due to its scale, timing and unpredictability, requires immediate action with increased resources. Events become emergencies based on two factors: the volume of work to be performed and the time in which the organisation must respond. Emergency is a concept that relates to the context in which it occurs.

Public health event: a manifestation of disease or an occurrence that creates a potential for disease.

Public health threat: a condition, agent or incident which may cause, directly or indirectly, ill health.

Quality assurance: quality assurance is the intended or regular actions required to provide enough confidence that a product or service will assure the given requirements.

Quality control: quality control is the ongoing attempt to maintain the integrity of a process in order to maintain the reliability of an outcome, or the techniques and actions necessary to ensure quality.

Quality improvement: quality improvement is the purposeful change of a process to improve the reliability of achieving an outcome.

Quarantine: the restriction of activities and/or separation from others of suspect persons who are not ill or of suspect baggage, containers, conveyances or goods in such a manner as to prevent the possible spread of infection or contamination.

Random sample: a sample of persons chosen in such a way that each one has the same (and known) probability of being selected.

Risk assessment: the qualitative and quantitative estimation of the likelihood of adverse effect that may result from exposure to specified health hazards or from the absence of beneficial influences.

Risk factor: an aspect of personal behaviour or lifestyle, an environmental exposure, or a hereditary characteristic that is associated with an increase in the occurrence of a particular disease, injury, or other health condition.

Risk: probability that an event will occur (e.g. that a person will be affected by, or die from, an illness, injury, or other health condition within a specified time or age span).

Sample size: the number of cases that need to be included in the study.

Semantic interoperability: the ability of computer systems to exchange data with unambiguous, shared meaning. Semantic interoperability is a requirement to enable machine-computable logic, inferencing, knowledge discovery, and data federation between information systems.

Surveillance systems evaluation: a comprehensive measurement of relevant attributes (over time or in a single effort) that have been running for some time in order to verify whether the system is still 'fit for purpose' and to give recommendations for guiding improvements.

Surveillance attribute: defined set of characteristics used to evaluate an existing system or to conceptualise a proposed system.

Syndromic surveillance: the real-time (or near real-time) collection, analysis, interpretation and dissemination of health-related data to enable the early identification of the impact (or absence of impact) of potential human or veterinary public-health threats which require effective public health action.

Triangulation: technique of combining several different methods (when evaluating external validity of a surveillance system) to facilitate validation of data through cross-verification from different data sources.

Vehicle: an inanimate object that can carry an agent from a reservoir to a susceptible host (e.g. food, water, blood products, and bedding).

Venn diagram: diagram that shows all possible logical relations between a finite collection of sets. Venn diagrams were conceived around 1880 by John Venn. They are used to teach elementary set theory, as well as illustrate simple set relationships in probability, logic, statistics, linguistics and computer science.