AUSTRALIAN COMMISSION
ON SAFETY AND QUALITY IN HEALTH CARE

INFORMATION
for clinicians

Validating infection surveillance data

Surveillance refers to the ongoing, systematic collection, analysis, interpretation, and dissemination of data about a health-related event for use for public health action to reduce morbidity and mortality and to improve health.[1] The main purpose of infection surveillance is to collect information about the infectious agent, its host, the healthcare environment, and other risk factors that contribute to the spread of infection.[1] This information can then be used to implement effective infection prevention and control (IPC) measures to reduce risk of infection.

Depending on the type of health service, infection surveillance may include:

- Monitoring processes that affect the risk of infection, such as hand hygiene, aseptic technique or correct use of personal protective equipment
- Monitoring for clinical outcomes, such as infection.

It is critical to ensure that infection surveillance data is high quality and reflects the true risk of infection in a population or setting before using it to inform IPC strategies. High quality surveillance data can be obtained by using standardised surveillance methods and definitions, and by data validation.

What is data validation?

Data validation refers to processes used to ensure that surveillance data is correct, complete, timely and plausible. If data is not validated, actions taken to respond to the data may be ill-informed and ineffective.

Data validation relies on the establishment of a set of criteria to assess the quality of the surveillance data. These validation criteria provide a reference point that defines:

- What data needs to be recorded
- How data needs to be presented
- What limits and parameters should be used to analyse the data.

Table 1 includes a list of checks that form the basis of validation criteria for infection surveillance.

Some validation criteria will need to be specific to the process or clinical outcome that is being monitored. For example, the criteria to validate *Clostridioides difficile* infection surveillance data will be different to the criteria used to validate surveillance data for infections related to total knee arthroplasty surgery.

Criteria may also change over time in response to changes in benchmarks, or due to changes in clinical and surveillance practices.[2]

Data validation should be done during surveillance data collection, data entry, and, crucially, before the data is analysed or used. Additional checking of the data may be required if:

- There has been a recent history of poor data quality
- A new ward, service or facility is opened or commissioned
- New surveillance staff have commenced
- A new data system contributing to surveillance data is in operation
- There are changes in laboratory testing or reporting practices that affect surveillance.

Ideally, independent audits of surveillance data (for example, spot checks) should also

be undertaken to complement routine data validation processes.[3]

Data validation can be done:

- By manually reviewing data
- Using automated processes that link information systems, such as surveillance databases, patient admission records and pathology systems, and use programming software to clean and check the data
- Using a combination of manual and automated processes.

Manual data validation can be labour intensive and require specialised surveillance skills and knowledge to review the data. However, errors are often easy to fix when found through manual data validation.

Automated data validation can provide more accurate data[4], however setup can be expensive and will require biostatistical and programming skills and software. There is also a risk that any errors made during the setup of the automated data validation can inadvertently introduce errors into surveillance data.

Organisations should consider the advantages and disadvantages of manual and automated data validation, relative to their infection surveillance needs.

Data validation in practice

CASE SCENARIO 1

A hospital's IPC team is preparing a twelve-month chart of bloodstream infections (BSIs). Based on the data that has been collected, at least eight infections occurred in each month, except for June, when there were two BSIs.

IPC teams have a good understanding and awareness of the surveillance data being collected and understand the usual infection trends in their organisation.

The two cases in June, in the context of the rest of the data, should alert the IPC team to crosscheck the surveillance data for the month of June to ensure that all BSI cases have been included in the surveillance data.

CASE SCENARIO 2

When an organisation submits data to the National Hand Hygiene Initiative Hand Hygiene Compliance Application (HHCApp), the HHCApp validation report will indicate if more than 700 moments against an auditor's name.

Data that has been collected in the HHCApp since 2010, provides the National Hand Hygiene Initiative with a data model on the usual auditing practice in Australian health service organisations. Based on this model, most auditors do not submit more than 600 moments for each audit period. This is an assumption that is built into the HHCApp data validation process to assist organisation administrators to check for inaccurate data. If this item is flagged at the end of an audit, the organisation administrator should check the session lists to ensure that the auditor's data has been entered correctly and that there is no duplication.

CASE SCENARIO 3

A state health department discovered that two health networks had much lower healthcare-associated (HA) *Staphylococcus aureus* bloodstream infection (SABSI) rates compared to other health networks in the state.

An investigation revealed that the two health networks were using an incorrect definition that resulted in lower HA-SABSI rates. These networks were recording a HA-SABSI if the infection occurred 72 hours after admission or discharge. The other networks were using the national standard definition, for which a HA-SABSI is recorded if the infection is identified 48 hours after admission or discharge. The use of consistent national surveillance definitions allows for comparison across states and territories, and between organisations.[5]

The Commission has produced a series of national surveillance implementation guides to support a nationally standardised approach to the surveillance of these infections. The application of the surveillance case definition should be based on the results from the same diagnostic tests. If this is not possible, the different diagnostic tests must be shown to be comparable for surveillance data to be considered consistent and comparable for analysis.

Table 1. Data validation checklist for infection surveillance

Has all data been correctly recorded?
☐ Check for missing data entries
☐ Check for data that may have been recorded in duplicate by mistake
☐ Check the correctness of any data that has been copied from another data source
☐ Check for any data that has been recorded as '0' by mistake
☐ Check any data that has been generated by an automated formula
☐ Check that all recorded dates have occurred and are not in the future
☐ Check that all data directly relates to the surveillance that is being undertaken *e.g. For environmental cleaning audits, have visual inspection results been recorded instead of adenosine triphosphate (ATP) bioluminescence results?*
Is all data presented in the correct format?
☐ Check that data is recorded in the correct data field
☐ Check that the data is recorded consistently in same format
☐ Check that the data is presented consistent with historical or other reference data
Is there any unusual data?
☐ Check that data are within the expected or acceptable range
☐ Check data is consistent with historical or other reference data
☐ Check that the setup of automated data validation is correct
☐ Check that there is a plausible explanation for any outlying or unusual data
Are linked data logical and consistent?
☐ Check that linked data makes sense *e.g. Only infections from biological females are recorded for surveillance of caesarean-related surgical site infections*
☐ Check that data obtained from different data sources are in agreement
Has the surveillance definition been used consistently?
☐ Check that appropriate information has been used to apply the surveillance definition *e.g. Can pathology results alone be used to determine an infection or are other data sources needed?*
☐ Check that the data refers only to cases that occurred during the surveillance period
☐ Check that data is based on the collection of appropriate specimens, where relevant
☐ If multiple data sources are used for case finding, check that the data sources are used in a consistent order

For more information

The Commission has produced the following resources to support organisations with infection surveillance:

Implementation Guide for the Surveillance of *Staphylococcus aureus* Bloodstream Infection

Implementation Guide: Surveillance of Central Line-Associated Bloodstream Infection

Implementation Guide for the Surveillance of *Clostridioides difficile* Infection (CDI)

Surgical Site Infection Surveillance

Infection Prevention and Control eLearning Modules (Basics of Surveillance and Quality Improvement)

References

1. National Health and Medical Research Council. Australian Guidelines for the Prevention and Control of Infection in Healthcare. Commonwealth of Australia, 2019.
2. Mitchell BG, Gardner A. Addressing the need for an infection prevention and control framework that incorporates the role of surveillance: a discussion paper. J Adv Nurs. 2014 Mar;70(3):533-542.
3. Cato KD CB, Larson E. Data elements and validation methods used for electronic surveillance of health care-associated infections: a systematic review. Am J Infect Control. 2015;43(6):600-605.
4. Lasry O, Ailon T, Charest-Morin R, Dea N, Dvorak M, Fisher C, et al. Accuracy of hospital-based surveillance systems for surgical site infection after adult spine surgery: a Bayesian latent class analysis. J Hosp Infect. 2021 Jul 14;117:117-123.
5. Russo PL, Cheng AC, Richards M, Graves N, Hall L. Healthcare-associated infections in Australia: time for national surveillance. Aust Health Rev. 2015 Feb;39(1):37-43