ecdc
EUROPEAN CENTRE FOR DISEASE PREVENTION AND CONTROL

TECHNICAL REPORT

Managing heterogeneity when pooling data from different surveillance systems

www.ecdc.europa.eu

ECDC TECHNICAL REPORT

Managing heterogeneity when pooling data from different surveillance systems

This report was commissioned by the European Centre for Disease Prevention and Control (ECDC), coordinated by Chantal Quinten, and produced by Anthony Nardone of EpiConcept.

Authors
EpiConcept: Esther Kissling, Anthony Nardone and Marta Valentiano.
ECDC: Julien Beauté, Bruno Ciancio, Joana Gomes Dias, Takis Panagiotopoulos and Chantal Quinten.
ECDC would also like to acknowledge the following ECDC staff for supporting the project and providing critical review of the documents: Mike Catchpole, Vicky Lefevre, Gianfranco Spiteri and Phillip Zucs.

Acknowledgements
The authors would like to acknowledge the support and guidance provided ECDC colleagues, especially Joana Haussig, Csaba Kodmon, Hanna Merk and Marianna Marozzi.
This guidance was produced under Framework Service Contract Number ECDC/2014/041 with EpiConcept.

Suggested citation: European Centre for Disease Prevention and Control. Managing heterogeneity when pooling data from different surveillance systems. Stockholm: ECDC; 2019.

Stockholm, October 2019

PDF
ISBN 978-92-9498-383-1
doi: 10.2900/83039
Catalogue number TQ-03-19-759-EN-N

Contents

Figures

Tables

Abbreviations

ECDC	European Centre for Disease Prevention and Control
EEA	European Economic Area
EU	European Union
HIV	Human Immunodeficiency Virus
I-TSA	Interrupted Time series Analysis
OR	Odds Ratio
ROR	Relative Odds Ratios

Executive summary

Pooling surveillance data from different surveillance systems has been undertaken extensively to inform public health action and to obtain estimates of health outcomes. However, pooling data from different surveillance systems poses a number of analytical and procedural problems that arise from the heterogeneity of these systems. Although a wide body of literature on managing heterogeneity when pooling data is available, only a few studies have focused on the particular issues that arise when pooling surveillance data from different surveillance systems.

The aim of this guidance is to support investigators in understanding the sources of and highlighting the approaches to managing heterogeneity when analysing pooled surveillance data from different surveillance systems. This will be done by:

- Defining the purpose of pooling data to achieve surveillance objectives, and listing the sources of heterogeneity that may arise (Section 1);
- Describing criteria to assess heterogeneity and approaches to minimise it for each stated surveillance objective (Section 2);
- For key sources of heterogeneity, describe statistical and procedural approaches to minimise or remove its impact (Section 3).

EpiConcept, in collaboration with the European Centre for Disease Prevention and Control (ECDC), undertook an iterative four-stage process to develop this guidance document, focusing on pooling data from different surveillance systems, which involved:

- List sources of heterogeneity that may occur when pooling surveillance data;
- Targeted literature review of statistical and procedural approaches to minimise heterogeneity;
- Development of a guidance to assess impact of and identify approaches to minimising heterogeneity;
- Piloting of guidance with experts based in ECDC.

Heterogeneity that arise when pooling data from different surveillance systems were grouped into the following three groups:

- Heterogeneity of surveillance systems: this may arise as a result of differences in the design and operation of surveillance systems and collection of data;
- Heterogeneity in disease determinants: this may arise as a result of true differences in exposure or vulnerability to disease between different populations under surveillance;
- Heterogeneity of data quality: this may arise as a result of either missing or erroneous data or a preponderance of data from a limited number of systems.

The impact of each of these sources of heterogeneity was assessed with regards to the following three surveillance objectives: trend analysis, risk factor analysis and burden of disease estimation. For major sources of heterogeneity, Section 3 provides case studies which describe in greater detail the operation of the source of heterogeneity, how to assess its impact, and statistical and procedural methods to minimise its impact.

The guidance will assist investigators to assess the validity of pooling data from different surveillance systems to derive point estimates, and describe a number of methodologies to minimise the impact of heterogeneity arising from pooling. This will allow public health specialists and researchers to answer key research and policy questions using available European data to the fullest possible extent.

Section 1. Introduction to heterogeneity and analysis of pooled surveillance data

1.1 Context

Pooling data from different surveillance systems, whether these systems are based on a national or local geography, has been undertaken extensively to inform public health action and to obtain estimates of health outcomes [1, 2]. Nonetheless, pooling data from different surveillance systems poses a number of analytical and procedural problems. Before undertaking a pooled analysis, due consideration must be given to the advantages and disadvantages of performing such an analysis. For some of the problems encountered when pooling surveillance data, a number of possible solutions may exist, whilst others cannot be resolved. For the latter, an awareness of the impact that these heterogeneities may have on estimates is needed to avoid erroneous conclusions being drawn and to decide on alternative analysis approaches.

The European Centre for Disease Prevention and Control (ECDC) contracted EpiConcept to prepare guidance for pooling surveillance systems, with a focus on pooling surveillance data from national surveillance systems.

1.2 Aims and objectives of report

The aim of the guidance is to support investigators in understanding the sources of heterogeneity and highlighting the approaches to managing it when analysing pooled surveillance data, with the focus on pooling data from different national surveillance systems. The objectives are to:

- Define the purpose of pooling national data to achieve ECDC surveillance objectives (Section 1);
- List, in an easily accessible checklist, sources of heterogeneity that arise from pooling of national surveillance data (Section 1);
- Describe criteria to determine if data can be pooled for the purpose of achieving specific surveillance objectives (Section 2);
- For key sources of heterogeneity, describe statistical and procedural approaches to remove or reduce unwanted heterogeneity when pooling national surveillance data and its impact on individual surveillance objectives (Section 3).

1.3 Methodology

EpiConcept, in collaboration with ECDC, undertook a four-stage process to develop this guidance document:

- Identification of sources of heterogeneity that may occur when pooling national surveillance systems were collected through interviews with experts and reviews of relevant literature;
- Targeted literature review to ascertain statistical and procedural approaches for minimising or controlling for heterogeneity when pooling national data;
- Development of a toolkit to collect expert opinions on the impact of each source of heterogeneity on stated surveillance and analysis objectives, and to identify the criteria and approaches for minimising or controlling for heterogeneity when pooling national data in light of the stated surveillance objective;
- Piloting and subsequent revision of the toolkit through its application to two EU disease surveillance systems coordinated by ECDC.

1.4 Reading the guidance

In this Section 1, we provide a general introduction to the issue of heterogeneity and its relevance to pooling of surveillance data from different national systems. Within this section we list in an easily accessible checklist, sources of heterogeneity that arise from pooling of national surveillance data.

In Section 2, we describe the importance and impact of the sources of heterogeneity when pooling surveillance data from different surveillance systems to achieve different surveillance objectives. We describe the approaches needed to control for them, which will vary according to the specific objectives (as listed below), of the analysis undertaken.

In Section 3, we present detailed case studies of some of the more important sources of heterogeneity and their impact on the surveillance objectives. These examples group some sources of heterogeneity as the issues raised and the approaches employed are very similar. For each example, the following structure is employed:

- surveillance objective;
- description of the source of heterogeneity;
- an example of the source of heterogeneity;
- the issues raised by the source of heterogeneity;
- assessment of the impact of the heterogeneity;
- statistical and/or procedural actions to reduce the impact of the source of heterogeneity and the limitations in applying these actions.

1.5 Definition of pooling

Pooling data is a useful exercise with application in a variety of fields in public health, including surveillance of infectious disease [1, 2], estimation of burden of disease [3, 4], vaccine effectiveness [5] and drug safety studies [6]. Estimates of health outcomes and effects from pooled data may be obtained in one of two ways: pooling of either individual data or of aggregate estimates derived from different sources or studies [7]. Although there are clear advantages in pooling data from different studies, it does give rise to three sorts of heterogeneity:

- clinical heterogeneity due to variability in the patient populations (e.g. age, baseline disease severity) and treatment protocols (e.g. frequency of dose);
- methodological heterogeneity due to variability in study design and risk of bias;
- statistical heterogeneity due to larger differences that could be expected from chance alone in the results of individual studies when measuring the same outcome. Methodological and clinical sources of heterogeneity may contribute to the magnitude and presence of statistical heterogeneity.

Ideally, data are pooled from studies and data sources that employ a standardised approach to data collection, thus reducing the impact of heterogeneity on subsequent estimates of health outcomes. However, the assumption of homogeneity can be questioned even in multi-site studies that employ a common protocol developed *a priori*, and study sites are often cited as a source of heterogeneity; such as sampling factors, particularly clinical characteristics of participants and, ad treatment protocol factors [8, 9].

For the purposes of this report, pooling refers to combining data from national surveillance systems to obtain European and regional estimates of health outcomes. In routine surveillance reports, such as ECDC's Annual Epidemiological Reports, data from national surveillance systems are pooled to provide European Union (EU) or EU/European Economic Area (EEA) estimates of disease, as well as analysis of trends and risk factors. Pooling of national surveillance data also occurs when undertaking non-routine analyses to answer specific research or policy questions.

Although efforts are made to standardise the collection of surveillance data in Europe (e.g. EU case definitions [10] and reporting protocols [11]), national surveillance systems remain largely heterogeneous. Therefore, heterogeneity remains a major concern and is likely to arise when pooling national surveillance data, which may impact on the precision and accuracy of estimates obtained and their interpretation.

1.6 Objectives of European surveillance of communicable diseases

ECDC performs indicator-based surveillance, which is complemented by event-based surveillance in all 28 EU Member States and three EEA countries (Liechtenstein, Iceland and Norway) [12].

Member States submit surveillance data on 56 communicable diseases and related health issues to the European Surveillance System, and ECDC manages, analyses and disseminates statistics and reports based on these data. ECDC and the European disease networks undertake validation of submitted data and strive for standardised reporting and data comparability across the EU through the use of common (externally quality-assured) diagnostic and typing methods, case definitions, metadata and reporting protocols.

ECDC regularly disseminates summaries and analyses of submitted data through a variety of channels and media, which include the publication of regular routine outputs such as Annual Epidemiological Reports [13], as well as interactive online tools to analyse submitted data such as the Surveillance Atlas of Infectious Diseases [14]. Outputs based on data submitted to ECDC also include estimates of burden of disease, risk factor analysis and trends for national surveillance data, which can be pooled either at the European level for the EU/EEA, or in groupings of fewer countries. Furthermore, non-routine or *ad hoc* analyses on data collected by ECDC are often performed to answer specific policy and research questions.

The aim of the ECDC long-term surveillance strategy (2014-2020) is for the surveillance of communicable diseases in the EU/EEA to provide relevant data for the effective prevention and control of infectious diseases while minimising the burden on Member States [15]. The strategy identifies the following six objectives for the analysis of surveillance data:

- monitor trends in communicable diseases;
- detect and monitor any multinational communicable disease outbreaks;
- contribute to the evaluation and monitoring of prevention and control programmes;
- identify population groups at risk;
- contribute to the assessment of the burden of communicable diseases;
- generate hypotheses about determinants of disease epidemiology and their impact.

For the purposes of this guidance, these six objectives have been combined into three categories of analysis below which the pooling and analysis of surveillance data from different countries can better achieve:

- **Trend analysis**: Pooling of surveillance data from different countries enables epidemiologists to estimate European trends of disease incidence, and extrapolate likely future incidence of disease, either within the EU/EEA or regionally, contributing to the prioritisation of resources and identification of outbreaks. The comparison of national trends with both European trends as well as those of other countries can inform the evaluation of the impact of different policies and interventions to prevent and control disease. Pooling of national surveillance data may result in an increased power of data analyses if there is a decrease in variability, which can provide insights not available from national surveillance systems alone. Furthermore, pooling of national data may detect supra-national outbreaks of diseases that might have been missed by individual national surveillance systems.
- **Risk factor analysis**: Pooling national data will enable a larger sample size to be achieved and thus obtain greater statistical power to understand better known risk factors (e.g. age), identify unknown risks (e.g. in minority populations) and estimate the impact of less common risks and/or risk factors for rare diseases. Furthermore, pooling of data may also result in improved representativeness of the sample and thus wider generalisability of the subsequent results.
- **Burden of disease**: Pooling surveillance data from national systems will enable European estimates of the burden of disease to be calculated, either within the EU and EEA or regionally. A unique set of methodologies have been developed to estimate the burden of disease [3, 4, 16] which are important for assessing health progress (requiring comparison over time), and assisting in the prioritisation of resources at an international level.

1.7 Sources of heterogeneity

Pooling data from different national surveillance systems will lead to heterogeneity in the pooled estimate, despite efforts to reduce this through standardisation of national surveillance systems. We have developed a list of possible sources of heterogeneity that may arise when pooling data from different national surveillance systems based on the domains identified by the Cochrane Collaboration of methodological and clinical heterogeneity [17], as well as heterogeneity arising from data quality. We have grouped the sources of heterogeneity under the three groups listed below:

- **Heterogeneity of surveillance system structures and operations:** Heterogeneity of surveillance systems may arise as a result of differences in the design and operation of national surveillance systems. The sources of heterogeneity within this group are mostly methodological, as identified by the Cochrane Collaboration. Heterogeneity can arise from biases (e.g. selection and ascertainment biases; for further details see case study 4) which can lead to erroneous estimates and interpretation. Analytical and/or procedural approaches must be applied to remove or minimise these biases.
- **Heterogeneity in disease determinants:** Heterogeneity of measured health outcomes may arise as a result of differences between countries in the presence and the effectiveness of interventions (e.g. screening young people for Chlamydia infection), populations (e.g. number people who inject drugs), temporality (e.g. influenza seasons) or environments (e.g. West Nile Fever). This heterogeneity reflects the true disease pattern and should not be controlled for, as it is these differences that are of interest in the surveillance objectives (Section 1.5), informing research questions and developing health policy. Nonetheless, when reporting such differences, possible limitations of the results of pooled analysis need to be clearly stated.
- **Heterogeneity of data quality:** Heterogeneity due to differences in data quality may arise as a result of either missing or erroneous data. Missing data represent a loss of information which reduces the efficiency of the analysis, and the possibility of identifying effects sought in any analysis. Erroneous data reporting, differential missing data or reporting of data from a limited number of countries can lead to selection and ascertainment biases, which can result in the estimates of effect being over- or under-estimated. Finally, a limited number of data points can reduce the representativeness of the data and the subsequent generalisability of any interpretations.

For each of these three groups, we have prepared a checklist (Table 1.1) which identifies the possible sources of heterogeneity that may arise when pooling data from different national surveillance systems, and where there is possible overlap with other sources of heterogeneity.

Table 1.1. Checklist of sources of heterogeneity when pooling surveillance data from different surveillance systems.

Pooling issue	Source of heterogeneity	Link with other sources heterogeneity
1. Heterogeneity of surveillance system structures and operations		
1.1.1 System design: comprehensive, sentinel or survey	Different surveillance system designs will collect data from different populations, whether by size or other characteristics, as well as differing in coverage.	Geographical coverage (1.2.1) Population under surveillance (1.2.2)
1.1.2 Reporting mode: active or passive surveillance	Active surveillance systems will generally have higher levels of details than passive surveillance systems.	Case ascertainment (1.4) Missing covariate data (3.3)
1.1.3 Data format: Case-based or aggregate reporting	Data from aggregate surveillance systems are of generally poorer quality and include fewer variables, which will limit the flexibility and analyses that can be performed.	
1.1.4 Legal status: mandatory or voluntary	The imposition of mandatory reporting to surveillance systems may or may not result in improved reporting compared to voluntary systems [18;19].	
1.1.5 System permanence: routine or temporary	Surveillance systems may be established temporarily to cover periods of high incidence (e.g. during peak season or an emergency). Temporary surveillance systems will not provide data for all periods covered by routine systems.	Heterogeneity of time periods (2.3) Missing time period (3.2)
1.1.6 Data sources: population, general practice, laboratory or hospital	Cases reported from different data sources may represent different severity of diseases if hidden differences in the population are present.	Case definition (1.3) Case ascertainment (1.4) Intervention effectiveness (2.1)
1.2.1 Geographical coverage	Uneven geographical coverage of the systems (e.g. if surveillance is limited to a specific area/region) may impact the representativeness of the total national population.	System design (1.2.1)
1.2.2 Population under surveillance	Surveillance systems established in specific populations may not be representative of the general population; i.e. socio-demography (e.g. if surveillance is limited to a specific age group) or risk behaviours (e.g. if surveillance is limited to a people with a specific risk factor such as people who inject drugs).	Data sources (1.1.6) Heterogeneity of population characteristics (2.2.)
1.3 Case definition: variation in case definition, its status (i.e. confirmed, possible, probable or suspect) and/or whether ascertained through syndromic or diagnostic surveillance systems.	Surveillance systems, which employ different case definitions will ascertain cases with differing sensitivity and specificity.	Case ascertainment (1.4) Intervention effectiveness (2.1)
1.4 Case ascertainment	Health service configuration (type and number of clinical services available) and national policy (e.g. criteria to employ diagnostic tests, existence of a screening programme) can affect both access to services and reporting from services.	Data sources (1.1.6) Intervention effectiveness (2.1)
1.5 Timeliness of information flows	Reporting delay (the time from diagnosis to notification), can lead to heterogeneity if it differs between national surveillance systems. This can be due to both differences in the surveillance system structure (i.e. different timeliness requirements) or different data quality (see point 3 in the current table).	Missing time period (3.2) Missing covariate data (3.3)
2. Heterogeneity in disease determinants		
2.1 Intervention effectiveness	The level of disease in a population may vary by the presence and the types of interventions offered (e.g. different influenza vaccination options) or the delivery of the intervention on offer (e.g. coverage of immunisation programme).	Data sources (1.1.6) Case definition (1.3) Case ascertainment (1.4)
2.2 Heterogeneity of population characteristics	Disease incidence in a population may vary from that of another population (e.g. another country); if there are differences in the population characteristics that may be linked to the disease under study (e.g. age structure, prevalence of certain risk factors).	Population under surveillance (1.2.2)

2.3 Heterogeneity of time periods	Disease incidence may vary between countries if peaks occur in different seasons or time points.	System permanence (1.1.5) Missing time period (3.2)
2.4 Heterogeneity of environment	Disease incidence may vary due to different geography and/or climate.	
3. Heterogeneity of data quality		
3.1 Missing all reports from one or more countries	Missing case or event reports from one or more countries will impact the amount of information available (and thus the power of the analysis) and the representativeness of any estimates.	
3.2 Missing time period	Missing time period data (no data reported for a given time period e.g. year), will affect not only the power of the analysis and the representativeness of estimates, but also the validity of certain analyses such as time series analyses.	System permanence (1.1.5) Timeliness of information (1.5) Heterogeneity time periods (2.3)
3.3 Missing covariate data	Missing co-variate data will impact the amount of information available (and thus the power of the analysis) and the representativeness of any estimates.	Reporting mode (1.1.2) Timeliness of information (1.5)
3.4 Under-reporting of cases	The under-reporting of cases (e.g. cases diagnosed but not reported or notified to the system) will reduce the number of data points available and not represent the true estimates of disease.	Reporting mode (1.1.2) Legal status (1.1.4) Data sources (1.1.6)

1.8 Conclusion

Pooling data is a methodology often employed to better understand disease epidemiology and its public health implications. Much emphasis has been placed in understanding the impact of heterogeneity on pooling study data for meta-analyses [17], but much less critical thought is given to the impact of pooling data from different (e.g. national) surveillance systems to achieve key surveillance objectives.

Before pooling data from different surveillance systems, careful consideration must be given to whether this is appropriate in answering the research and policy questions posed, or whether there are clear reasons against undertaking such a process.

We present elements in the following sections which allow the investigator to assess the impact and the validity of pooling surveillance systems. Pooling across national surveillance systems is used as an example, and a number of methodologies are described to minimise or remove the effect of heterogeneity arising from these different surveillance systems. This guidance will allow public health specialists and researchers to answer key research and policy questions using available European data to the fullest possible extent.

Section 2. Impact of sources of heterogeneity on surveillance objectives

2.1 Impact of sources of heterogeneity on trend analysis

2.1.1 Impact of heterogeneity of surveillance systems structures and operations on trend analysis

Preconditions for pooling national surveillance data: Data may be pooled for a trend analysis even if respective surveillance systems have major operational and structural differences, some of which may bias estimates (e.g. collecting data from different sources). Nonetheless, pooling of data is methodologically valid as long as each surveillance system contributing data has remained stable and consistent over the time period for which trends are analysed.

Approaches to enable pooling data: If national surveillance systems are not consistent during the time period of investigation, data may still be included in pooled estimates following an assessment of the anticipated and observed impact of these modifications on the levels of reported national data**.** If these changes remain minor, there are a number of analytical approaches that can be undertaken to the pooled data. A multiplication factor to account for differences in the new and old systems can be applied to national data to ensure consistency over time. The calculation of this factor can be undertaken through a variety of means including the use of serial capture-recapture studies to estimate the levels of under-reporting over time [20]. If the impact is deemed major, national data should not be included in pooled estimates and should be excluded from trend analysis or if possible presented stratified. Clear explanations must be given for the exclusion of the country data, and the trend data from that country should be presented separately (e.g. see case study 1).

Table 2.1. Impact of and approaches of accounting for heterogeneous surveillance system structures and operations when pooling data for trend analysis.

Source of heterogeneity		Impact of and approaches to accounting for heterogeneity of surveillance system structures and operations
1.1.1 System design	+++	Pooling data from differently designed surveillance systems is possible if their design is stable over the analysis period and if the systems target the same population type (e.g. size and characteristics). However, any alteration to system design (e.g. change in coverage) reported during the time period included in the trend analysis should be regarded as a major change. Thus, before pooling these data, qualitative and quantitative assessment of the impact of any change, and the application of statistical and procedural approaches are likely to be required. Further details are available in case study 1.
1.1.2 Reporting mode	++	Changes in reporting mode can be regarded as either major or minor and some assessment of the impact of the modification will be required.
1.1.3 Data format	+	As long as level of aggregation is still useful for trend of interest (e.g. trend by age group), pooled case-based and aggregated data trends can be followed over time and employed to detect outbreaks of disease. This approach can be applied to the evaluation of interventions but is limited by the inflexibility of aggregate data to control for possible confounding factors.
1.1.4 Legal status	+	Changes in legal status may have minimal impact on reporting, but this should be ascertained, especially if the health outcome is a stigmatised condition or linked to illegal activities.
1.1.5 Permanence of systems	+	Pooling data from surveillance systems with differing permanence to detect outbreaks is possible if data from the permanent systems are limited to the shortest common period. However, determining excess reporting thresholds will be statistically difficult.
1.1.6 Data sources	+++	If the data sources are different but remain stable during the study period, pooling from different systems following disease trends is methodologically valid and statistically feasible. It is nonetheless recommended to present trends by reporting source and compare them as severity of illness and/or populations may vary. Further details available in case study 1.
1.2.1 Geographical coverage	+	If the size of the population covered by the surveillance systems changes (i.e. moves from sub-national to national coverage), analyses should use robust population denominators to calculate rates.
1.2.2 Population under surveillance	+++	Pooling of data from different national surveillance systems established in different populations is not advised as combining data may obscure trends in one or the other population and bias estimates. Further details available in case study 1.

Source of heterogeneity		Impact of and approaches to accounting for heterogeneity of surveillance system structures and operations
1.3 Case definition	+++	Incompatible case definitions will lead to high levels of heterogeneity and pooling of data from different countries surveillance systems is therefore not recommended. If surveillance systems use the same case definition but different status (e.g. confirmed vs. probable), data can be pooled for trend analysis, but preferably stratified and presented by status. Further details available in =case study 2.
1.4 Case ascertainment	++	If major changes have occurred during the time period of investigation to national health service configuration or policy, national data should be excluded from the pooled estimate unless these changes can be corrected (e.g. underreporting) for.
1.5 Timeliness of information flows	++	Pooling of national data without reviewing and, if appropriate, adjusting for reporting delays is not advised. Delays to the notification of cases most often occur in the more recent years which will impact on trend analysis. An *a priori* threshold of excessive reporting delay for which adjustment is considered necessary must be considered for each disease. For example, if more than 5% of HIV cases are notified with a delay of two or more quarters in excess of the minimal truncation time, data should be adjusted for reporting delay [21]. Most adjustment techniques rely on estimation of the delay distribution independently of the diagnosis rate which is used to estimate the proportion of cases already reported [21,23].

Definitions of grading: +Limited impact and bias unlikely; ++Limited impact but bias possible; +++Major impact and bias likely

2.1.2 Impact of heterogeneity in disease determinants on trend analysis

Preconditions for pooling national surveillance data: Data may be pooled if the effect of a public health intervention or other disease determinants are the same across countries, i.e. the observed differences are due to chance. Nonetheless, if there are different interventions, for example two different vaccines strategies adopted in Europe, data can be pooled to assess if the effect of the two strategies is significantly different. In similar fashion, data from surveillance systems with heterogeneity in population, temporal or environmental profiles can be pooled for a trend analysis. However, if the observed differences cannot be attributed to any specific factor that can be adjusted for, then the results should be presented separately and pooling is not advised.

Approaches to enable pooling data: It is advised however that further information should be collected regarding risk factors such as age, temperature and external events such as atypical or novel outbreaks, or introduction of key public health interventions that may bias trends. Controlling for these confounding variables in a trend analysis is recommended when different subpopulations present different patterns over time.

Table 2.2. Impact of and approaches to accounting for heterogeneity in disease determinants when pooling data for trend analysis.

Source of heterogeneity	Impact and approaches to accounting for heterogeneity in disease determinants	
2.1 Heterogeneity of intervention effectiveness	+	When analysing and interpreting trends, external information should also be collected. For example, the implementation of a screening programme in one country may result in a preponderance of cases from that country in the pooled dataset. Thus, pooling data from countries with different public health interventions may obscure important national trends. Therefore, trends should be stratified by countries with different interventions to evaluate their impact.
2.2 Heterogeneity of populations characteristics	+	When analysing trends, important socio-demographic or risk factors should be included as covariates to identify important trends that may affect sub-populations.
2.3 Heterogeneity of time period	+	Pooling data from countries with different peak seasons of disease incidence require that data are either aggregated at time unit level so that the peaks occur in the same time units of analysis (e.g. year) or that data are presented separately. When analysing trends, the inclusion of atypical years (e.g. because of outbreaks or seasons) may need to be controlled for or eliminated depending on the research question.
2.4 Heterogeneity of environment	+	Data from countries with different environments that may impact disease incidence may be pooled, although it is best practice to stratify trends by environment.

** Definitions of grading: +Limited impact and bias unlikely; ++Limited impact but bias possible; +++Major impact and bias likely*

2.1.3 Impact of heterogeneity of data quality on trend analysis

Preconditions for pooling national surveillance data: Data from different surveillance systems with different data quality accuracy may be pooled for trend analysis. However, whether the observed trends reflect the true trend depends on whether or not the trend is statistically significantly different for each level of the covariate (e.g. confounding factor) under investigation, and the reason for missingness. For example, when the covariate data is missing completely at random, the observed trend reflects the true trend. Trend analysis is not advised if pooling surveillance data involves missing information over a time period from a country, or when a preponderance of cases is from one or a few countries.

Approaches to enable pooling of data: If data are missing for a specific time period under investigation, either all data from that country should be excluded, at the cost of efficiency and representativeness, or missing data for one year can be extrapolated from the available time periods for that specific country. With necessary caution, missing country data can be imputed with data from neighbouring countries or countries with similar characteristics.

Table 2.3. Impact of and approaches to account for heterogeneity of data quality when pooling data for trend analysis

Source of heterogeneity		Impact of and approaches to accounting for heterogeneity of data quality
3.1 Missing all reports from one or more countries	++	Missing country data will impact on the amount of information available (and thus the power of the analysis) and the representativeness of any estimates. Pooling data to which one or more countries have not contributed may give rise to bias if these countries share a common factor (e.g. increasing or decreasing trend). Missing country data have been imputed by comparison with statistical and geographic neighbours, although such imputations tend to be performed only for smaller countries [3].
3.2 Missing time period	+++	In some cases, missing time period data from a single country have resulted in the exclusion of that country's data from any trend analysis [21, 25]. However, incomplete time series analyses can be biased and also lead to a loss of power and precision for the proposed trend analysis. Missing time period data can be imputed. Further details are in case study 3.
3.3 Missing covariate data	++	Pooling national data with missing covariates will not invalidate the trend analysis but will limit further interpretation and identification of trends for important sub-groups for which there are high levels of missing data. Further details in case study 3.
3.4 Under-reporting of cases	+	Trends can be analysed by pooling data from national surveillance systems despite varying levels of under-reporting as long as these have remained stable over the time period of investigation. However, analysts should also consider that differential under-reporting of cases by one or a group of countries may obscure or exacerbate reported trends. Correction factors can be applied to national data to account for differing levels of under-reporting.

** Definitions of grading: +Limited impact and bias unlikely; ++Limited impact but bias possible; +++Major impact and bias likely*

2.2 Impact of sources of heterogeneity on risk factor analysis

2.2.1 Impact of heterogeneity of surveillance system structures and operations on risk factor analysis

Preconditions for pooling national surveillance data: Data from different national surveillance systems with different structures and operations may be pooled to generate hypotheses and identify risk factors under the assumption that the presence of a case in each data source is not dependent on the risk factor under investigation (i.e. there is no selection bias in the reporting systems). It may be difficult to envisage the existence of such a bias for some of the sources of heterogeneity (e.g. active or passive surveillance systems). However, other sources of heterogeneity, such as the legal status of the system may cause a selection bias if the risk factor or hypotheses under investigation are linked to a behaviour or activity that may result in under- or over-reporting (e.g. illicit drug use) and consequent bias in the estimate of the effect.

Approaches to enable pooling of data: To avoid potential selection bias, the investigator must first assess the existence and size of the bias through the calculation of relative odds ratios (ROR) in the group of interest and the source population [26, 27]. Selection bias can be controlled for by adjusting for the covariates linked to the bias [27] or inverse probability weighting [28]. When this is not possible, pooling should be avoided or analyses restricted to the countries with unbiased data. More details in case study 4.

Table 2.4. Impact of and approaches to accounting for heterogeneity due to surveillance system structures and operations when pooling data for risk factor analysis

Source of heterogeneity		Impact of and approaches to accounting for heterogeneous surveillance system structures and operations
1.1.1 System design	+	If inclusion in each different system design is not dependent on the risk factor under investigation, data can be pooled from different data sources to generate hypotheses or identify risk factors.
1.1.2 Reporting mode	+	If the distribution of risk factors is dependent on the reporting mode, then the heterogeneity in the reporting mode should not have an impact (e.g. selection bias) on the risk factor analysis under investigation.
1.1.3 Data format	+	If aggregate data sources report on the risk factor under investigation, transformation of case-based to aggregate data will enable risk factors to be identified, although the reduced ability to control for confounders may limit the applicability of this approach. The limited flexibility and analytical options available when using aggregate data can reduce the utility of such data for hypothesis generation.
1.1.4 Legal status	++	Pooling data from surveillance systems with differing legal status to generate hypotheses is statistically feasible and methodologically valid. The presence of possible selection biases because of different legal status should be ascertained, especially if the hypothesis involves behaviours (e.g. illegal activity) that may result in under-or over-reporting to one or the other system. In case of selection bias, the steps outlined in case study 4 should be considered.
1.1.5 Permanence of systems	+	Pooling data from surveillance systems with differing permanence to generate hypotheses is statistically feasible and methodologically valid. Consideration should be given to the existence of possible selection biases if there is a suggestion of seasonality in the outcome or the putative risk factor. If a selection bias seems likely, either limit data from the permanent surveillance systems to cover the shortest common period, or include seasonality in multivariable models.
1.1.6 Data sources	+++	Pooling data from surveillance systems with differing data sources to identify risk factors is statistically feasible and methodologically valid. However, as some data sources are more likely to include cases with different risk profiles, following the steps described in case study 4 is recommended.
1.2.1 Geographical coverage	+++	Pooling data from surveillance systems with differing geographical coverage for hypotheses generation is statistically feasible and methodologically valid. Careful consideration should be given to the possible operation of selection biases. If sub-regional systems cover populations that differ from the national population (e.g. by age, socio-demography or risk factor under investigation), the size of the effect may be either under- or over-estimated. In such instances, it is recommended to follow the steps described in case study 4.
1.2.2 Population under surveillance	+++	Surveillance systems established in a specific population may not be representative of the general population, be it by socio-demography (e.g. age), risk behaviours (e.g. people who inject drugs) or severity of disease. Combining such data with that from general population surveillance systems can obscure trends in one or the other population and bias estimates. Thus, pooling of data from national surveillance systems established in different populations is not advised. If the investigator wishes to pool data, the steps outlined in case study 4 should be followed and reported upon.
1.3 Case definition	++	Incompatible case definitions in different surveillance systems may lead to high levels of heterogeneity and pooling of data from these systems is often not advisable. Before deciding on pooling, the investigator should assess the nature of the differences and judge if this can lead to substantial misclassification of cases. If surveillance systems use the same case definition but different status (e.g. confirmed *vs.* probable), data can be pooled, but preferably stratified and presented by status.
2.2 Case ascertainment	+++	Pooling data to generate hypotheses is not advised if national surveillance systems have differing case ascertainment. If there are major differences in case ascertainment between countries (e.g. the existence of a screening programme [29]), data from those countries should either be stratified by country or excluded from pooling.
1.5 Timeliness of information flows	++	Pooling of national data without reviewing and, if appropriate, adjusting for reporting delays is not advised. If reporting delays occur most often in cases with a particular socio-demography or risk, especially if linked to the hypotheses under investigation, data should be analysed as outlined in case study 4.

** Definitions of grading: +Limited impact and bias unlikely; ++Limited impact but bias possible; +++Major impact and bias likely*

2.2.2 Impact of heterogeneity in disease determinants on risk factor analysis

Preconditions for pooling national surveillance data: Data from surveillance systems with heterogeneity in disease determinants may be pooled for a risk factor analysis. Moreover, generation of hypotheses by pooling data from countries with heterogeneous health outcomes can provide further insights and understanding of disease transmission and risk factors by analysis of extra covariates such as types of interventions. There are limited occasions when this heterogeneity would impact the validity of the pooled estimates. Therefore, it rarely requires statistical approaches to control for or remove this heterogeneity.

Approaches to enable pooling of data: When pooling data with heterogeneity to assess the effectiveness of interventions, it is important to collect other external information that may be valuable in the interpretation of any results (e.g. existence and coverage of a screening or immunisation programme).

Table 2.5. Impact of and approaches to accounting for heterogeneity in disease determinants when pooling data for risk factor analysis

Source of heterogeneity		Impact and approaches to accounting for heterogeneity in disease determinants
2.1 Heterogeneity of intervention effectiveness	+++	Pooling data of countries with major heterogeneity of service effectiveness is not advised unless this heterogeneity is included as a covariate for multivariable analysis.
2.2 Heterogeneity of populations characteristics	+	Generating hypotheses by pooling data from countries with different population structures is statistically feasible and methodologically valid, but important socio-demographic or risk factors must be included as covariates in order to control for confounders that may affect subpopulations.
2.3 Heterogeneity of time period	+	Pooling data from countries with different peak season of disease incidence will require that data are either pooled, so that the peaks occur in the same time units of analysis (e.g. year), or are presented separately.
2.4 Heterogeneity of environment	+	Data from countries with different environments that may impact disease incidence may be pooled, although it is best practice to include this as a covariate in a multivariable analysis.

** Definitions of grading: +Limited impact and bias unlikely; ++Limited impact but bias possible; +++Major impact and bias likely*

2.2.3 Impact of Heterogeneity of data quality on risk factor analysis

Preconditions for pooling national surveillance data: Data with heterogeneity due to incomplete data for covariate data may be pooled for risk factor analysis. Whether the observed effects reflects the true effects depends on whether or not the observed outcome is statistically significantly different for each level (e.g. male vs female) of the risk factor under investigation, and the reason for missingness. For example, when the covariate data is missing completely at random, the observed trend reflects the true trend. If data are missing at random, this results in a loss of power in hypothesis generation and true associations may be missed. Furthermore, the complete missingness of important covariate data will limit the multivariable analysis and the control of important confounders.

Approaches to enable pooling of data: Statistical approaches are available to eliminate potential limitations driven by incomplete covariate data by imputing the missing values. If data for the time period or the risk factor under investigation are completely missing for a specific country, either all data from that country should be excluded, with a concomitant loss of efficiency and representativeness, or data should be imputed with information from other countries with similar risk profiles.

Table 2.6. Impact of and approaches to accounting for heterogeneity of data quality when pooling data for risk factor analysis

Source of heterogeneity		Impact of and approaches to accounting for heterogeneity of data quality
3.1 Missing all reports from one or more countries	++	Missing country data will impact on the amount of information available (and thus the power of the analysis) and the representativeness of any estimates. Pooling data to which one or more countries have not contributed may give rise to heterogeneity if these countries share a common factor (e.g. populations at risk). If the majority of cases is reported by one or a group of countries, this may bias the information and reduce the representativeness and generalisability of any analyses. Under such circumstances, pooling data is not advised as it will not add value to analyses already performed by the Member State.
3.2 Missing time period	++	Missing time series data will impact on the amount of information available (and thus the power of the analysis) and the representativeness of any estimates.
3.3 Missing covariate data	+++	Pooling data with missing covariate data is possible, although the reduced information available can reduce the power of the analysis to identify risk factors as well as introduce biases. Multiple approaches to assess the nature of missing covariate data and to imputing missing data are described in case study 5. However, if there are high levels of missing covariate data, it is not advised to pool data as it may have an adverse impact on generation of hypotheses and may compromise the interpretation of any analysis.
3.4 Under-reporting of cases	+	Hypotheses can be generated using pooled data from national surveillance systems with varying levels of under-reporting as long as the under-reporting is considered or demonstrated to be random. If under-reporting is linked to a factor that may contribute to the outcome, a selection bias may be in operation. Further details are available in case study 4.

** Definitions of grading: +Limited impact and bias unlikely; ++Limited impact but bias possible; +++Major impact and bias likely*

2.3 Impact of sources of heterogeneity on burden of disease estimates

2.3.1 Impact of heterogeneity of surveillance system structures and operations on burden of disease estimates

Preconditions for pooling national surveillance data: Pooling data from different surveillance systems has been undertaken extensively to obtain global and regional estimates of disease, but these estimates will only be representative if there is no case ascertainment bias and the population characteristics are comparable across the pooled surveillance systems. When these preconditions are met, pooling data is valid in obtaining an estimate of the total burden of a disease. However, depending on the research question and consequently the purpose of pooling, it might be appropriate to have only a representative number of cases within a particular type or population (e.g. high risk population) under surveillance to represent the burden within that population of interest. As a consequence, the preconditions may differ depending on the research purpose.

Approaches to enable pooling data: There are a number of statistical options to address the heterogeneity rising from different data collection systems. The most common methodology is that described by the Global Burden of Disease study in which meta-regression modelling is used to account for differences in reporting by system design [4].

Table 2.7. Impact of and approaches to accounting for heterogeneity due to surveillance system structures and operations when pooling data to estimate burden of disease

Source of heterogeneity		Impact of and approaches to accounting for heterogeneous surveillance system structures and operations
1.1.1 System design	+	Pooling data from different surveillance system designs to estimate overall burden of disease requires statistical transformation including: 1. rates per population using denominator values for catchment population of each reporting source [25]. 2. estimates stratified for each type of surveillance system. 3. meta-regression modelling to account for differences in reporting by system design (e.g. Global Burden Disease DisMod-MR tool) [4].
1.1.2 Reporting mode	+	Pooling data from active and passive surveillance systems to estimate overall burden of disease is statistically feasible and methodologically valid. If levels of completeness are thought to differ greatly between the different reporting modes, a threshold of a minimum number of reported cases can be established for national data to be included in the estimation of disease burden [30].
1.1.3 Data format	+	Pooling case-based and aggregate surveillance data is possible. Nonetheless, some studies estimating burden of disease have relied on case data and excluded aggregate data (e.g. Burden of Communicable Disease in Europe [30]).
1.1.4 Legal status	+	Pooling data from surveillance systems with differing legal status to estimate burden of disease is statistically feasible and methodologically valid.
1.1.5 Permanence of systems	+	Pooling data from surveillance systems of differing permanence to estimate disease burden requires one of the following steps: 1. limiting data from permanent surveillance systems to the shortest common reporting period. 2. imputation of data for the missing time period in the temporary surveillance system. 3. exclusion of data from Member States with temporary surveillance systems and their separate presentation.
1.1.6 Data sources	+++	Estimates of disease burden often pool data from surveillance systems which employ similar data sources. To account for variations in severity of cases reported by different sources, the measure of outcome can be standardised using metrics such as Disability-Adjusted Life Years (DALYs) [4].
1.2.1 Geographical coverage	+	Estimates of disease burden often pool data from surveillance systems which cover different national and sub-national populations. This requires that estimates of disease burden from each surveillance system are reported as rates per population covered by the respective surveillance system [25].
1.2.2 Population under surveillance	+++	Surveillance systems established in specific populations may not be representative of the general population, be it by socio-demography (e.g. age), risk behaviours (e.g. people who inject drugs) or severity of disease. Thus, combining such data with that from general population surveillance systems can bias estimates. Thus, data from different populations should either be stratified by type of subpopulation or excluded from pooled analysis.
1.3 Case definition	+++	Inconsistent case definition in different surveillance systems will lead to high levels of heterogeneity and pooling is, therefore, not advised.
2.3 Case ascertainment	+++	If national surveillance systems have different case ascertainment, pooling data to estimate burden of disease is not advised. Data from countries with major differences in case ascertainment (e.g. the existence of a screening programme) should be excluded from pooling.
1.5 Timeliness of information flows	++	Pooling of national data without reviewing and, if appropriate, adjusting for reporting delays is not advised. If reporting delays exceed an agreed threshold, statistical adjustments should be taken into consideration. (see Table 2.1).

** Definitions of grading: +Limited impact and bias unlikely; ++Limited impact but bias possible; +++Major impact and bias likely*

2.3.2 Impact of heterogeneity in disease determinants on burden of disease estimates

Preconditions for pooling national surveillance data: Pooling data from systems with heterogeneity in disease determinants has been undertaken extensively to obtain global and regional estimates for a specific disease. Although the impact of this source of heterogeneity on the burden estimates is limited, it might be of interest, depending on the purpose, to pool only a representative number of cases within a particular population profile or within a specific setting (e.g. similar weather patterns), to represent the burden within a population with specific vulnerability and/or exposure characteristics. As a consequence, the preconditions may differ depending on the research purpose.

Approaches to enable pooling data: If a global burden is envisioned and in order to facilitate further interpretation and comparative analyses of national and international estimates of disease burden, relevant additional information should also be collected (e.g. the existence and coverage of key public health interventions, health service configuration).

Table 2.8. Impact of and approaches to accounting for heterogeneity in disease determinants when pooling data to estimate burden of disease

Source of heterogeneity		Impact and approaches to accounting for heterogeneity in disease determinants
2.1 Heterogeneity of intervention effectiveness	+	Estimation of disease burden by pooling data of countries with major heterogeneity of service effectiveness is statistically feasible and methodologically valid. However, the collection of extra information from each country on the type and configuration of services may allow a comparative analysis to assess the effectiveness of interventions.
2.2 Heterogeneity of populations characteristics	+	Estimation of burden of disease by pooling data from countries with different population structures is undertaken extensively [31]. The estimates should preferably be reported according to socio-demographic characteristics if they present different disease risks.
2.3 Heterogeneity time period	+	Pooling data from countries with different peak seasons of disease incidence requires that data are either pooled, so that the peaks occur in the same time units of analysis (e.g. year), or are presented separately.
2.4 Heterogeneity of environment	+	Data from countries with different environments that may impact disease incidence may be pooled. The estimates should preferably be reported according to environmental characteristics if they present different disease risks.

** Definitions of grading: +Limited impact and bias unlikely; ++Limited impact but bias possible; +++Major impact and bias likely*

2.3.3 Impact of heterogeneity of data quality on burden of disease E=estimates

Preconditions for pooling national surveillance data: Data may be pooled to estimate the burden of disease unless information is missing for the relevant time period (e.g. year) or country. If data are only available from one or few countries, the robustness and representativeness of any estimates for the burden of disease in Europe, or sub-regions of Europe, will be reduced. The absence of covariate data does not affect the estimation of a global burden of disease, but may not allow estimates for important subpopulations (e.g. by age).

Approaches to enable pooling of data: If data are missing for key covariates (e.g. age), values may be imputed if the purpose of pooling is to draw valid and robust conclusions for specific subpopulations. Validated statistical approaches are available to impute the missing values for covariates of interest. In similar fashion, correction factors can be estimated to account for underreporting to eliminate the underestimation of the observed outcomes.

Table 2.9. Impact of and approaches to accounting for heterogeneity of data quality when pooling data to estimate burden of disease

Source of heterogeneity		Impact of and approaches to accounting for heterogeneity of data quality
3.1 Missing all reports from one or more countries	++	Pooling data to which one or more countries have not contributed may give rise to heterogeneity if these countries share a common factor (e.g. political structure or environment). Otherwise, missing country data will impact on the amount of information available (and thus the power of the analysis) and the representativeness of any estimates. Missing country data have been imputed by comparison of statistical and geographic neighbours [3].
3.2 Missing time period	++	Pooling data for a missing time period is possible although pooling should be performed for those surveillance systems who report consistently over the envisioned time period. Imputation of missing time data for national systems has been described. For example, the Global Health Estimates prepared by WHO allow for three years data to be imputed using a two-step process where mortality rates for each country and cause were interpolated and then extrapolated [3].
3.3 Missing covariate data	+++	Pooling national data with missing covariates will not invalidate estimates of the burden of disease but will limit estimates for important subgroups with high levels of missing data. Missing covariate data can be imputed and further details are available in case study 5.
3.4 Under-reporting of cases	+++	Under-reporting of cases will underestimate the true burden of disease with important consequences for the development of appropriate health policy and prioritisation. Correction factors, that may be calculated from all or a group of countries, may be applied to the estimates [20, 32].

** Definitions of grading: +Limited impact and bias unlikely; ++Limited impact but bias possible; +++Major impact and bias likely*

Section 3. Case studies of impact of specific sources of heterogeneity on surveillance objectives

Case study 1. Impact of system design, data sources and population under surveillance on trend analysis

Number in checklist:

1.1.1 System design - impact on trend analysis: +++

Pooling data from different surveillance systems is possible if their design remains stable over the analysis period. However, any alteration to system design during the time period analysed should be regarded as a major change.

- **Example:** In the United States of America, a study to compare the effects of the seven-valent pneumococcal conjugate vaccine on the hospitalisation of children aged under five years with invasive pneumococcal disease demonstrated good comparability between data recorded by sentinel and comprehensive surveillance systems [33].

1.1.6 Data sources - impact on trend analysis: +++

Pooling data for trend analysis from stable surveillance data that record information from different sources is statistically feasible and methodologically valid. Nonetheless, changes in data sources during the study period should be regarded as a major change and will require an assessment of the impact of the modification.

- **Example:** A wide variety of different data sources (hospital, physician, laboratory, other services) report data on *Haemophilus influenzae* to the different EU/EEA national surveillance systems [34].

1.2.2 Population under surveillance - impact on trend analysis: +++

Pooling data from different populations for trend analysis is statistically feasible. However, changes in the population covered by a surveillance system should be regarded as a major change and will require an assessment of the impact of the modification before pooling. This source of heterogeneity is distinct from geographical coverage (1.2.1), but similar to population heterogeneity (2.2).

- **Example:** Although the majority of countries that report gonorrhoea cases indicate that most of their data are obtained from dedicated specialist services, the number of cases reported will be dependent on the level of services that are dedicated to vulnerable and high-risk populations, such as men who have sex with men and young people [35, 36].

Issues

- **Fewer health facilities under surveillance (system design):** This may result in a lower proportion of the true number of cases in the country being recorded by the surveillance system. However, if good service catchment population denominators are available, and the trend analysis is based on rates, it can minimise or eliminate this source of heterogeneity.
- **Different health facilities reporting data (population covered):** This is very likely to result in differences in the types of cases (e.g. by disease severity, age, other risk factors) being reported. However, if data can be stratified by type of health facility, it minimises the impact on any analysis of trends.
- **Sentinel systems may be more exhaustive than comprehensive systems:** This may result in less under-reporting than previously in the sentinel health facilities. If there is a considerable change in levels of under-reporting, even a trend analysis using rates will be biased.

Assessment

This should include a qualitative assessment (steps 1 and 2) and further quantitative assessment, if deemed necessary (step 3). Note that the quantitative assessment in step 3 provides indications about the impact of the heterogeneity, but not evidence, as the assessment is ecological. The following steps are recommended:

1. **Expert opinion:** This should be the first step in assessing the impact of changes on the surveillance system and whether these are major sources of heterogeneity. The assessment is qualitative, based on existing knowledge of the system and system changes, and should include discussions with key stakeholders. Depending on whether the outcome of these discussions means the changes are major or minor, the following should occur:
   - **Minor:** The discussion should be recorded and the changes in the surveillance systems should be noted in any presentation of the results. The discussion should be underpinned by at least a visual assessment of the pre and post-change data as described in **Step 2**.
   - **Major:** The discussion should be recorded and the changes assessed by procedures described in **Steps 2 and 3**.

2. **Visual assessment of numbers/rates pre- and post-change:** Plot the total numbers of cases/rates by time. Do the post-change values visually follow the pre-change trends?
   - **If no**: Check if other differences may explain the change in trends: plot the numbers/rates by time for EU-EEA countries excluding the country of interest. Is there a European-wide trend that may explain differences between pre- and post-change numbers/rates for the country of interest? Check with a disease expert if there are any disease-specific interventions/changes that may explain this change in trend. If differences can be explained by external factors, you have some indication that the surveillance system change may have had a minor impact. If not, you have some indication that the surveillance system change may have had a major impact. If you would like to statistically assess the potential change, **go to 3**
   - **If yes:** you have some indication that the surveillance system change may have had a minor impact. If you would like to statistically assess the potential change, **go to 3.**

3. **Statistically assess numbers/rates pre- and post-change:** An interrupted time series analysis (I-TSA) could be undertaken to assess if there was a level or a slope change after the surveillance system change. At its simplest, this could be done by including an indicator variable (0=pre-change, 1=post-change) in a regression model that includes time and counts/rates of disease. With an interaction between the indicator variable and time, you can check if the change is significant and if a slope changes. I-TSA can be done at several levels of complexity. When performing an I-TSA analysis, in consultation with appropriate statistical support, consider: a) the best regression model to use (check for overdispersion and violation of regression assumptions) b) different functional forms of trends pre- and post-change c) introduction of lags after change d) a transition period between pre- and post-change. Is the amount of change considerable and/or statistically significant?

- **If no:** This does not mean that there are no real differences. The statistics are influenced by sample size. Is the amount of change large? If yes, you still may want to exclude this country from the analysis.
- **If yes:** What is the amount of change? Is it an important change? If sample size is large, you may get p-values <0.05 despite the amount of change not impacting much.
- **Limitations:** The change in system design may not be very defined in time, thus making it difficult to carry out an I-TSA I. Design considerations could include using a lag in the I-TSA, or excluding from the model time periods during which the changes are taking place [37]. Limitations include factors other than system design may mask changes in trend as well as the recommended number of at least 9 data points pre- and post-change [38].

Actions

- **Employ robust population denominators:** this will allow rates to be calculated which may be less biased than counts in any trend analysis. Some limitations to be considered include:
  - **Robust population denominators.** Examples include the difficulty in obtaining robust hospital catchment population (especially if the hospital is a centre of excellence which attracts out-of-area cases) [25].
  - **Difficulty in obtaining population denominators by specific population profiles.** Normally robust denominators of local population by age and sex are readily available. However, this may become more difficult for certain socio-demographic (e.g. social class or ethnic group) or behavioural characteristics (e.g. sexual orientation, drug misuse) [39].
- **Exclusion or inclusion of country data:** If this procedural approach is taken, reasons for inclusion/exclusion need to be discussed/justified as well as implications and limitations of results. The decision may be made *a priori* based on knowledge and/or expert opinion, or may also be made based on statistical indicators for which it will be important to discuss limitations of statistical methods in the report. The main considerations of this approach are:
  - **If excluding countries:** the reduced data points will impact on the amount of information available (and thus the power of the analysis) and the representativeness of any estimates, especially if one or more excluded countries share common characteristics.
  - **If including the country:** consider whether to account for country in the trend analysis, and if so, whether to use a fixed effects or random effects analysis.

Case study 2. Impact of case definition on trend analysis

Number in checklist:

1.3 Case definition: impact on trend analysis: +++

Pooling data from surveillance systems with inconsistent case definitions is not advised. However, before deciding against pooling of data, the different case definitions should be reviewed to ensure that sources of heterogeneity are large and cannot be controlled for.

- **Example:** In the United States, between 1992 and 1993 the case definition for AIDS was expanded to include patients with a CD4 count less than 200 per microlitre, regardless of symptoms. The reported increase in the number of AIDS cases was attributable to the expanded AIDS case definition [40].

Issues

- **Analyses of pooled data employing different case definitions:** The inclusion of cases with different status (e.g. confirmed, probable or possible) will include diagnostic criteria with differing levels of sensitivity and specificity. This differential misclassification can result in an ascertainment bias.

Assessment To determine if case definitions differ markedly:

- Review of different case definitions employed. Are the differences between case definitions minor?
  - **If no, go to 3**
  - **If yes, go to 2**
- Have case definitions changed over time?
  - **If no**: pooling cases is methodological valid. Note that all analyses should highlight the different case definitions and the steps undertaken to satisfy the investigators that the impact of heterogeneity was limited.
  - **If yes, go to 3**
- Assess the numbers of cases in the national surveillance system whose definitions is consistent with other countries or over time. Are there adequate data to apply a common case definition to national surveillance data to exclude those cases not consistent with the definition?
  - **If no**: to avoid erroneous analyses and conclusions, the recommendation is to exclude country data.
  - **If yes:** exclude cases that do not comply with the common case definition, analyse the remaining dataset and highlight the exclusion of cases in the report.

Actions

- **Exclusion or inclusion of country data:** If this procedural approach is taken, reasons for inclusion/exclusion need to be discussed/justified as well as the implications and limitations of results. The decision may be made *a priori* based on knowledge and/or expert opinion, or may also be made based on statistical indicators for which it will be important to discuss limitations of statistical methods in the report. The main considerations of this approach are:
  - **If excluding countries:** the reduced data points will impact on the amount of information available (and thus the power of the analysis) and the representativeness of any estimates, especially if one or more excluded countries share common characteristics.
  - **If including the country:** consider whether in the trend analysis to account for country, and if so, whether to use a fixed effect or random effects.
- **Stratification of data** so that trends are analysed only in each strata of data using data from those countries which use a consistent case definition.

Case study 3. Impact of missing time period on trend analysis

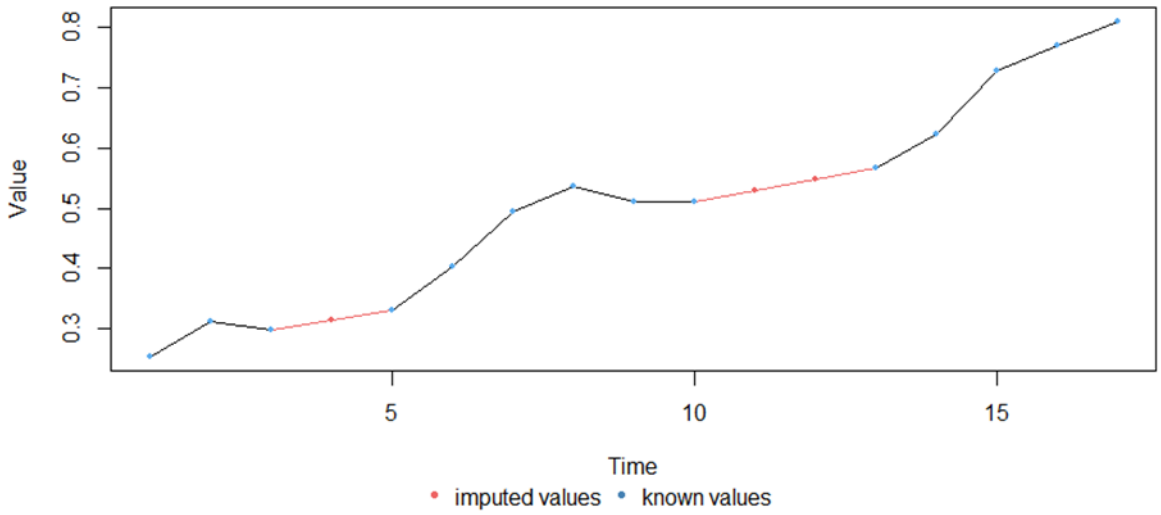
Number in checklist

3.2 Missing time period: impact on trend analysis: +++

Pooling data from surveillance systems with missing time period data is not advised without appropriate statistical approaches to impute missing data.

- **Example:** The surveillance of antimicrobial consumption undertaken by ECDC has examples of data being imputed for countries with missing time period data. In the example below, data were not reported by country C on the use of antibiotics in the hospital sector (expressed as Defined Daily Dose/1000 inhabitants/day) for a number of years. However, imputation of the missing years using interpolation from existing R packages has allowed a trend to be estimated for Country C and its inclusion in EU estimates[i].

Figure 3.1. Reported and imputed antibiotic use (Defined Daily Dose/1000 inhabitants/ day) in the hospital sector for country C, 2000-2016.



Issues

- **Analyses undertaken on data from countries providing complete time series:** The results can lead to a loss of power and precision for the proposed trend analysis.

Assessment

Before imputing missing time period data, it is advisable to consider the following assessments:

- **Review missing time period data:** If the time unit for your analysis has many records with missing values, you could consider using a less precise time unit for your analysis that may have fewer records with missing values (e.g. moving from a weekly analysis to a monthly analysis). This will depend on the objectives of your analysis and the acceptable amount of missing data. By changing time unit, you will lose information.
- **Review data from other countries for missing time period:** This is particularly so for infections that can be characterised by sudden and large outbreaks (e.g. influenza). If a large outbreak has occurred in neighbouring countries during the time period for which a country's data are missing, imputing missing values using national data before and after the period may be biased, as they will not reflect the epidemic that may have occurred.

[i] In this example, missing years have been imputed by applying the R Amelia package [49] on antimicrobial data submitted by EU/EEA Member States to ECDC. Access to the data can be obtained through the ECDC surveillance atlas for diseases https://www.ecdc.europa.eu/en/surveillance-atlas-infectious-diseases

Actions

- **Exclusion or inclusion of country data:** If this procedural approach is taken, reasons for inclusion/exclusion need to be discussed/justified as well as implications and limitations of the results. The decision may be made *a priori* based on knowledge and/or expert opinion or based on statistical indicators for which it will be important to discuss the limitations of statistical methods in the report. The main considerations of this approach are:
  - **If excluding countries:** the reduced data points will impact on the amount of information available (and thus the power of the analysis) and the representativeness of any estimates, especially if one or more excluded countries share common characteristics.
  - **If including the country:** ensure that imputation methodology is clearly presented; consider using existing guidelines for the reporting of missing data and the imputation processes performed [41].
- **Complete-case, available-case, single imputation and missing indicator methods:** although these methods are generally simpler to understand and perform, all have been extensively criticised as being inadequate for dealing with the missing data patterns and time series [41-45].
- **Imputation of missing data:** There exist many different approaches to imputing missing data, but these will reduce the robustness of any estimates. Numerous guidelines describe steps and procedures if data are to be imputed. The key recommendation is that sensitivity analyses should be performed [41,46-48]. The key methods are highlighted below:
  - **WHO Global Health Estimates:** WHO have described a method to impute missing year data using a two-step process [3]:
    - To interpolate, a logistic regression was fitted for each missing country-sex-cause group, using death rates six years prior and six years after the missing data year as the dependent variable and year as the independent variable. If the logistic regression did not converge due to small numbers, the death rate was estimated as the average rate in the three years prior and three years following the missing data year.
    - To extrapolate for up to three years, a logistic regression was fitted to the first or the final six years of data (including interpolated estimates) for each country-sex-cause. Again, if the logistic regression did not converge due to the small number of deaths recorded, the death rate was estimated as the average of the first or last three years.
  - **Multivariate methods:** There are a number of methods to impute missing time series data. These have often been cited as superior to other methods for replacing missing time series data as they use multiple variables [49, 50]. A number of statistical packages using multiple imputation are available (e.g. AMELIA, iVAR, mice and VIM).
  - **Univariate imputation:** These methods have been put forward as more appropriate for time-series data since univariate time series do not possess more than one attribute, hence algorithms need to employ inter-time correlations [51]. Statistical packages to impute univariate missing time series data are available in R.

Case study 4. Impact of data sources and case ascertainment on risk factor analysis

Number in checklist

1.1.6 Data sources: impact on risk factor analysis: +++

Pooling data from surveillance systems with differing reporting sources to identify risk factors is statistically possible, as a statistical risk factor analysis is able to account for potential differences in risk characteristics that might occur when some data sources are more likely to include cases with different risk profiles.

- **Example:** A report of influenza vaccine effectiveness against influenza A(H3N2) in Europe using multicentre case-control studies in primary care and hospital sectors found much older cases in hospitals [5].

1.4. Case ascertainment: impact on risk factor analysis: +++

If national surveillance systems have differing case ascertainment, pooling data for risk factor analysis may introduce selection or differential misclassification biases. Differential ascertainment may arise from a variety of factors including differing operation of the screening programmes (see example below), or an increased awareness of populations at risk.

- **Example:** As hepatitis C is mostly asymptomatic, national testing policies will determine the number of diagnoses and account for the variation in notification rates between countries. Countries with extensive testing programmes that target at-risk and vulnerable populations, may report high notification rates even though the prevalence estimated from sero-surveys may be low [29].

Issues

- **Selection bias:** The assumption is that inclusion in each data source is not dependent on the risk factor under investigation. However, if reporting of cases with certain characteristics is more or less likely in one surveillance system than another, it may bias the pooled estimates and lead to erroneous interpretation of these results.
- **Ascertainment bias:** Differential misclassification bias can be introduced when one country's surveillance system preferentially ascertains one group of cases compared to others, whether that be for example, by improved services, alternative screening strategies or differentially sensitive diagnostic tests.

Assessment If a selection bias is suspected, the existence and size of the bias can be estimated.

- **Existence and size of selection bias:** The calculation of a relative odds ratio (ROR) of the odds ratio among participants (ORSub) to the corresponding estimate in the source population (ORtot) [26, 27]. Causal diagrams such as the directed acyclic graphs can be also used to understand the structural relation between variables and to distinguish causal effects from biases [26, 27]. If there is evidence of a selection bias, a number of approaches (described below) may be undertaken.

Actions

- **Exclusion of country data:** If this procedural approach is taken, reasons for exclusion as well as implications and limitations of results need to be discussed and justified. The decision may be made *a priori* based on knowledge and/or expert opinion, or may also be made based on statistical indicators for which it will be important to discuss limitations of statistical methods in the report. The main limitation of this approach is that the reduced data points will impact on the amount of information available (and thus the power of the analysis), and the representativeness of any estimates, especially if one or more excluded countries share common characteristics.
- **Stratification of data:** Calculate estimates of risk stratified by type of surveillance system although this will limit the generalisability of the estimates.
- **Statistical adjustment for selection bias:** A number of different approaches to adjusting for selection bias:
  - **Adjust for the covariates linked to the bias:** this is a common method to control for selection bias which adjusts for factors that link the exposure and the outcomes [26, 27].
  - **Inverse probability weighting:** in this two-stage process, the probability of being selected in the study is calculated for each individual, and each person in the study is weighted with an inverse of the probability of selection [28].
  - **Sensitivity analyses** should be performed if the methods proposed above cannot control for the selection bias [26] so that by changing different parameters and assumptions, the impact on the effect estimate can be assessed.

Case study 5. Impact of missing covariate data on risk factor analysis

Number in checklist

3.3 Missing covariate data: impact on risk factor analysis: +++

Pooling data with missing covariate data is statistically feasible and methodologically valid, although the reduced information available may lessen the power of the risk factor analysis and introduce biases. Multiple approaches to assessing the nature of missing covariate data and to imputing missing values are available and a brief overview is provided below.

- **Example:** The surveillance of HIV diagnoses in Europe is hampered by high levels of missing data for transmission category (heterosexual, injecting drug users (IDU), men who have sex with men (MSM)), especially for more recent years (Figure A.5.1). However, the missing transmission category data have been imputed to provide a more complete dataset (Figure 3.2)[ii]

Figure 3.2. Number of new HIV diagnoses by transmission category including missing data, EU/EEA, 2009-2018

[ii] In this example, missing data have been imputed using the ECDC HIV Estimates Accuracy Tool https://www.ecdc.europa.eu/en/publications-data/hiv-estimates-accuracy-tool and applied on HIV data submitted by EU/EEA Member States to ECDC through the European Surveillance System. Access to data can be obtained from https://www.ecdc.europa.eu/en/publications-data/european-surveillance-system-tessy

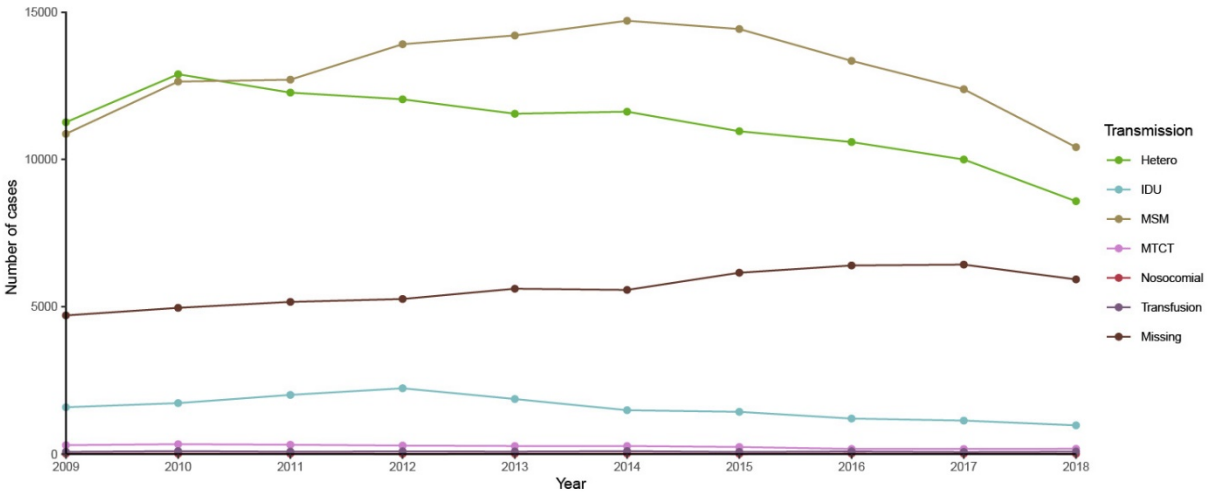
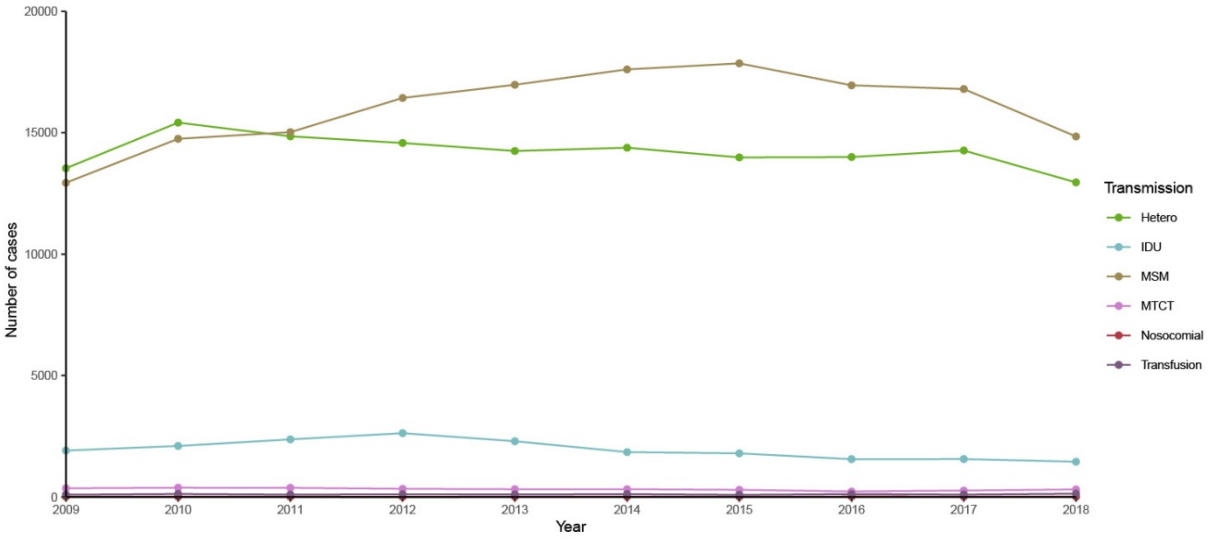


Figure 3.3. Number of new HIV diagnoses by transmission category in which missing data has been imputed, EU/EEA, 2009-2018



Issues

- **Analyses undertaken only on data available for complete time series:** The results of such analyses can be biased and also lead to a loss of power and precision for the proposed trend analysis.

Assessment

Before imputing missing data, it is advisable to consider the following assessments:

- **Nature of missing data:** An initial assessment must be performed to ascertain the nature of the missing data [42]. Missing data can be classified as one of the following three categories:
  - **Missing Completely At Random (MCAR)** in which missingness is completely random and not linked to any other patient characteristics. Most simple techniques for handling missing data, including complete and available case analyses, can give unbiased results although the analysis will be less efficient [52, 53].
  - **Missing At Random (MAR)** in which missing data commonly depend on available patient information and for which a number of multiple imputation methods are available.
  - **Missing Not At Random (MNAR)** in which missing data depend on usually unobserved patient characteristics. There is no consensus of how to manage such missing data [41].

Actions

- **Exclusion or inclusion of country**

References

1. European Centre for Disease Prevention and Control. Annual Epidemiological Report 2016 – Infectious disease surveillance summary, EU/EEA, 2014 data. [Internet]. Stockholm: ECDC; 2016. 2014 data. [Online] 2016. https://ecdc.europa.eu/en/infectious-diseases-public-health/surveillance-and-disease-data/annual-epidemiological-reports. [Accessed 18 July 2018]
2. World Health Organization. WHO Recommended Surveillance Standards. WHO. Geneva: https://www.who.int/csr/resources/publications/surveillance/whocdscsrisr992.pdf [Accessed 18 July 2018]
3. World Health Organization. Methods and data sources for global causes of death 2000-2011. [Online] June 2013. http://www.who.int/healthinfo/statistics/GHE_TR2013-3_COD_MethodsFinal.pdf [Accessed 18 July 2018.]
4. Global Burden of Disease Study Group. Global, regional, and national incidence, prevalence, and years lived with disability for 310 diseases and injuries, 1990–2015: a systematic analysis for the Global Burden of Disease Study 2015. Lancet 2016 388:1545-1602.
5. Kissling E, Rondy M and I-MOVE/I-MOVE+ study team. Early 2016/17 vaccine effectiveness estimates against influenza A(H3N2): I-MOVE multicentre case control studies at primary care and hospital levels in Europe. Eurosurveillance 2017 22:30464. 10.2807/1560-7917.ES.2017.22.7.30464.
6. Trifirò G, Coloma PM, Rijnbeek PR, *et al*. Combining multiple healthcare databases for postmarketing drug and vaccine safety surveillance: why and how? 6, 6 2014, Journal of Internal Medicine 2014 275:551-561. 10.1111/joim.12159.
7. Ragin C & Taioli E. 5 Meta-Analysis and Pooled Analysis-Genetic and Environmental Data. http://www.ph.ucla.edu/epi/faculty/zhang/courses/epi244/readings/ref%205-2.pdf [Accessed 20 September 2018]
8. Spirito A, Abebe KZ, Iyengar S *et al*. Sources of site differences in the efficacy of a multisite clinical trial: the Treatment of SSRI-Resistant Depression in Adolescents. Journal Consulting Clinical Psychology 2009 77:439-50. 10.1037/a0014834.
9. Kraemer HC. Pitfalls of multisite randomized clinical trials of efficacy and effectiveness. 3, 2000, Schizophrenia Bulletin 2000 26:533-41. 0586-7614.
10. Council of the European Union, European Parliament. Decision No 1082/2013/EU of the European Parliament and of the Council of 22 October 2013 on serious cross-border threats to health and repealing Decision No 2119/98/EC Text with EEA relevance - EU Law and Publications. [Online] 2013. https://ec.europa.eu/health/communicable_diseases/overview_en. [Accessed 28 September 2018]
11. European Centre for Disease Prevention and Control. EU protocol for harmonised monitoring of antimicrobial resistance in human Salmonella and Campylobacter isolates. ECDC; Stockholm 2016. https://ecdc.europa.eu/sites/portal/files/media/en/publications/Publications/antimicrobial-resistance-Salmonella-Campylobacter-harmonised-monitoring.pdf [Accessed 28 September 2018]
12. Paquet C, Coulombier D, Kaiser R, Ciotti M. Epidemic intelligence: a new framework for strengthening disease surveillance in Europe. Eurosurveillance 2006 11 :212-4. 1560-7917.
13. European Centre for Disease Prevention and Control. Annual Epidemiological Reports (AERs). [Online] https://ecdc.europa.eu/en/annual-epidemiological-reports [Accessed 28 September 2018].
14. European Centre for Disease Prevention and Control. Surveillance Atlas of Infectious Diseases. [Online] ECDC; Stockholm https://ecdc.europa.eu/en/surveillance-atlas-infectious-diseases [Accessed 28 September 2018].
15. European Centre for Disease Prevention and Control. Long-term surveillance strategy. ECDC; Stockholm 2013 https://ecdc.europa.eu/sites/portal/files/media/en/publications/Publications/long-term-surveillance-strategy-2014-2020.pdf [Accessed 28 July 2018]
16. Kretzschmar M, Mangen MJ, Pinheiro P *et al*. New Methodology for Estimating the Burden of Infectious Diseases in Europe. PLoS Medicine 2012 9:e1001205. 10.1371/journal.pmed.1001205.
17. The Cochrane Collaboration. 9.5.1 What is heterogeneity? [Online] 2011. https://handbook-5-1.cochrane.org/chapter_9/9_5_1_what_is_heterogeneity.htm [Accessed 28 July 2018].
18. Pearson A, Chronias A & Murray M. Voluntary and mandatory surveillance for methicillin-resistant Staphylococcus aureus (MRSA) and methicillin-susceptible S. aureus (MSSA) bacteraemia in England. Journal Antimicrobial Chemotherapy 2009 64(Supplement 1):i11-i17
19. Schwab F & Gastmeier P. The step from a voluntary to a mandatory national nosocomial infection surveillance system: the influence on infection rates and surveillance effect. Antimicrobial Resistance Infection Control 2012 1:24
20. Nardone A, Decludt B, Jarraud S *et al*. Repeat capture-recapture studies as part of the evaluation of the surveillance of Legionnaires' disease in France. Epidemiology Infection 2003 131:647-54. 0950-2688.
21. Magdalena R, Pantazis N, Janiec J *et al*. Potential adjustment methodology for missing data and reporting delay in the HIV Surveillance System, European Union/European Economic Area, 2015. Eurosurveillance 2018 23:10.2807/1560-7917.ES.2018.23.23.1700359.

22. Noufaily A, Enki DG, Farrington P *et al*. Modelling reporting delays for outbreak detection in infectious disease data. Journal Royal Statistical Society: Series A 2015 178:205-222. 10.1111/rssa.12055.
23. Brookmeyer R & Liao JG. The analysis of delays in disease reporting: methods and results for the acquired immunodeficiency syndrome. American journal of epidemiology 1990 132:355-65.
24. Freidl G, Sonder G Bovée L *et al*. Hepatitis A outbreak among men who have sex with men (MSM) predominantly linked with the EuroPride, the Netherlands, July 2016 to February 2017. Eurosurveillance 2017 22: 30468. doi: 10.2807/1560-7917.ES.2017.22.8.30468.
25. European Centre for Disease Prevention and Control. AER for 2016 - Antimicrobial consumption. 2018. ECDC; Stockholm 2018 https://ecdc.europa.eu/sites/portal/files/documents/AER_for_2016-AMC.pdf [Accessed 28 July 2018]
26. Nohr EA & Liew Z. How to investigate and adjust for selection bias in cohort studies. Acta Obstetricia et Gynecologica Scandinavica 2018 97:407-416. 10.1111/aogs.13319.
27. Hernán MA, Hernández-Díaz S & Robins JM. A structural approach to selection bias. Epidemiology 2004 15:615-25. 1044-3983.
28. Robins JM & Finkelstein DM. Correcting for noncompliance and dependent censoring in an AIDS Clinical Trial with inverse probability of censoring weighted (IPCW) log-rank tests. Biometrics 2000 56:779-88. 0006-341.
29. European Centre for Disease Prevention and Control. Hepatitis C. Annual epidemiological report for 2016. ECDC. Stockholm 2018 https://ecdc.europa.eu/sites/portal/files/documents/AER_for_2016-hepatitis-C.PDF [Accessed 12 September 2018].
30. Cassini A, Colzani E, Pini A *et al*. Impact of infectious diseases on population health using incidence-based disability-adjusted life years (DALYs): results from the Burden of Communicable Diseases in Europe study, European Union and European Economic Area countries, 2009 to 2013. Eurosurveillance 2018, Vol. 16. doi: 10.2807/1560-7917.ES.2018.23.16.17-00454.
31. Degenhardt L, Peacock S, Colledge S, Leung J *et al*. Global prevalence of injecting drug use and sociodemographic characteristics and prevalence of HIV, HBV, and HCV in people who inject drugs: a multistage systematic review. Lancet Glob Health 2017 5:e1192–e1207. doi: 10.1016/S2214-109X(17)30375-3
32. World Health Organization. Assessing tuberculosis under-reporting through Inventory studies. WHO Geneva 2012 https://www.who.int/tb/publications/inventory_studies/en/
33. Hampton LM, Zell ER, Schrag S, Cohen AL. Sentinel versus population-based surveillance of pneumococcal conjugate vaccine effectiveness. Bulletin World Health Organization 2012. 90:568-77. 10.2471/BLT.11.098178.
34. European Centre for Disease Prevention and Control. Haemophilus influenzae. In: ECDC. Annual epidemiological report for 2016. ECDC. Stockholm 2018.
35. European Centre for Disease Prevention and Control. Gonorrhoea. In: ECDC. Annual epidemiological report for 2016. ECDC; Stockholm; 2018. https://ecdc.europa.eu/sites/portal/files/documents/AER_for_2016-gonorrhoea.pdf
36. C M Lowndes, K A Fenton, the ESSTI (European Surveillance of STIs) Network*. Surveillance systems for STIs in the European Union: facing a changing epidemiology Sexually Transmitted Infections 2004 80:264-271. 10.1136/sti.2004.010389.
37. Shadish WR, Cook TD, Campbell DT. Experimental and Quasi-Experimental Designs for Generalized Causal Inference. Boston: Houghton Mifflin; 2002
38. Briesacher BA, Soumerai SB, Zhang F, Toh S *et al*. A critical review of methods to evaluate the impact of FDA regulatory actions. Pharmacoepidemiol Drug Safety 2013 22:986-994
39. Gobin M, Verlander N, Maurici C, Bone A *et al*. Do sexual health campaigns work? An outcome evaluation of a media campaign to increase chlamydia testing among young people aged 15–24 in England. BMC Public Health 2013 13:484. 10.1186/1471-2458-13-484.
40. Current Trends Update: Impact of the Expanded AIDS Surveillance Case Definition for Adolescents and Adults on Case Reporting -- United States, 1993. Centers for Disease Control and Prevention. 1994, MMWR Weekly, Vol. 43, pp. 167-170.
41. Sterne JAC, White IR, Carlin JB, Spratt M *et al*. Multiple imputation for missing data in epidemiological and clinical research: potential and pitfalls. BMJ 2009 338:1756-1833.
42. Donders A, van der Heijden GJ, Stijnen T & Moons KG. Review: A gentle introduction to imputation of missing values. Journal Clinical Epidemiology 2006 59:1087-1091.
43. Groenwold RHH, White IR, Donders ART, Carpenter JR *et al*. Missing covariate data in clinical research: when and when not to use the missing-indicator method for analysis. Canadian Medical Association Journal 184:1265-1269. 10.1503/cmaj.110977.
44. Knol MJ, Janssen KJ, Donders AR, Egberts AC *et al*. Unpredictable bias when using the missing indicator method or complete case analysis for missing confounder values: an empirical example. Journal Clinical Epidemiology 2010 63:728-36. 10.1016/j.jclinepi.2009.08.028.
45. van der Heijden GJ, Donders AR, Stijnen T, Moons KG. Imputation of missing values is superior to complete case analysis and the missing-indicator method in multivariable diagnostic research: A clinical example. Journal of Clinical Epidemiology 2006 59:1102-1109. 10.1016/j.jclinepi.2006.01.015.
46. Kneipp SM & McIntosh M. Handling missing data in nursing research with multiple imputation. Nursing

Research 2001 50:384-9. 0029-6562.
47. Wood AM, White IR & Thompson SG. Are missing outcome data adequately handled? A review of published randomized controlled trials in major medical journals. Clinical Trials 2004 1:368-376. 10.1191/1740774504cn032oa.
48. von Elm E, Altman DG, Egger M *et al*. Strengthening the reporting of observational studies in epidemiology (STROBE) statement: guidelines for reporting observational studies. BMJ 2007 335: 806-808. 10.1136/bmj.39335.541782.AD.
49. Zhang Z. Multiple imputation for time series data with Amelia package. Annals Translational Medicine 2016 4:56. 10.3978/j.issn.2305-5839.2015.12.60.
50. Liu S & Molenaar PCM. 45. iVAR: A program for imputing missing data in multivariate time series using vector autoregressive models. 4, 11 12 2014, Behavior Research Methods 2014 46:1138-1148. 10.3758/s13428-014-0444-4.
51. Moritz S, Sardá A, Bartz-Beielstein T *et al*. Comparison of different Methods for Univariate Time Series Imputation in R. https://arxiv.org/abs/1510.03924 [Accessed 28 September 2018]
52. Greenland S & Finkle WD. A critical look at methods for handling missing covariates in epidemiologic regression analyses. American Journal Epidemiology 1995 142:1255-64.
53. White IR & Carlin JB. Bias and efficiency of multiple imputation compared with complete-case analysis for missing covariate values. Statistics Medicine 2010 29:2920-2931. 10.1002/sim.3944.
54. Zhang Z. Missing data imputation: focusing on single imputation. Annals Translational Medicine 2016 4:9. 10.3978/j.issn.2305-5839.2015.12.38.
55. Lachin JM. Fallacies of Last Observation Carried Forward. Clinical Trials 2016 13:161–168. doi:10.1177/1740774515602688
56. Seaman SR & White IR. Review of inverse probability weighting for dealing with missing data. Statistical Methods Medical Research 2013 22:278-295. 10.1177/0962280210395740.

PDF TQ-03-19-759-EN-N

European Centre for Disease Prevention and Control (ECDC)

Gustav III:s Boulevard 40, 16973 Solna, Sweden

Tel. +46 858601000
Fax +46 858601001
www.ecdc.europa.eu

An agency of the European Union
www.europa.eu

Subscribe to our publications
www.ecdc.europa.eu/en/publications

Contact us
publications@ecdc.europa.eu

Follow us on Twitter
@ECDC_EU

Like our Facebook page
www.facebook.com/ECDC.EU

Publications Office of the European Union

PDF ISBN 978-92-9498-383-1