TESSy - The European Surveillance System

STI Reporting Protocol 2019

Sexually Transmitted Infections Surveillance data for 2018

February 2019

Contents

Introduction

This reporting protocol is for the 2019 data call for sexually transmitted infections surveillance data for 2018.

ECDC's Reporting Protocols are data collection guidelines for reporting countries' data managers; the Reporting Protocol is intended to improve user-friendliness by providing a uniform structure to make it easier for data managers to find data collection information across different subjects.

Because reporting countries' data managers sometimes play multiple roles, it is sometimes relevant to distribute subject-specific material together with a Reporting Protocol. To maintain the uniform structure, this information is now included in *Annex 2*.

How to use this document

This Reporting Protocol provides information for reporting countries' data managers in three main sections:

- *Reporting to TESSy* – contains guidelines on how to prepare data for submission to TESSy, deadlines, subject-specific information (e.g. new changes to metadata), and links to further information.
- *Annex 1* – contains:
  - A history of metadata changes for the subject(s) covered by this Reporting Protocol.
  - The metadata set for the subject(s) covered by this Reporting Protocol.
- *Annex 2* – contains subject-specific material relevant for distribution with the Reporting Protocol, for example:
  - Guidelines for data collection.
  - Analysis plan.

Finding further information

 Paragraphs denoted by the information icon tell where you can find further information.

Updated links to all the schedules, documentation and training materials mentioned in this Reporting Protocol are included in the *TESSy document section*, including:

- Metadata sets and history.
- Tutorials for data transformation using respectively Excel and Access.
- TESSy user documentation.
- CSV and XML transport protocols.

Copyright

Reporting to TESSy

This section provides both an overview of the TESSy reporting process and tips on where you can find useful information.

The overall process is:

1. *Familiarise yourself with the data collection deadlines*.
2. *Prepare (export and transform) your data*.
3. *Check that your data comply with the metadata*.
4. *Check that your data source profile is up-to-date*.
5. *Submit your file(s) to TESSy*.
6. *Finalise your submission*.

Checking the data collection schedule

The latest data collection schedule is available in the *TESSy website*.

The deadline for reporting 2018 STI surveillance data is 30 September 2019.

Preparing data

After you have exported the data from your national database, you need to ensure that the data are in a format that TESSy can accept. This applies both to the type of file submitted to TESSy (only CSV and XML files can be submitted) and to the format of the data in certain fields.

Tutorials covering how you can transform your data to the correct TESSy format using Excel or Access are available on the TESSy documents website. Information on the file formats is available in the CSV Transport Protocol and XML Transport Protocol.

Checking metadata

The TESSy metadata define the fields and data formats that are valid as input to TESSy for a given subject.

As requirements to the data to be shared among TESSy users change, the data changes needed to support the new requirements are identified and agreed upon between the National Surveillance Focal Points, the surveillance networks and their coordination groups and ECDC's disease experts, and then implemented as changes to the TESSy metadata.

In order to ensure that your data can be saved correctly in TESSy, you therefore need to check that your data are correctly formatted according to the most recent metadata set.

Changes to the metadata for the subject of this Reporting Protocol are described in:

- *Changes to current STI metadata* – changes since the last Reporting Protocol.
- *Annex 1 STI metadata change history* – all preceding changes.

It is especially important to focus on:

- **Field formats**

  Many fields require that data are formatted in a specific way. For example, dates must be in the YYYY-MM-DD format; dates in the DD/MM/YYYY format will be rejected.

- **Coded values**

  Some fields only permit the use of specific values (coded values). For example, **M**, **F**, **UNK**, or **Other** are the coded values for *Gender* and any other value in a *Gender* field will be rejected.

A single metadata set file contains all the definitions and rules you need to comply with to format your data correctly for every subject (usually a disease). The file can be downloaded as an Excel file from the TESSy documents website.

By filtering the fields in the file by subject, you can see the fields required for your subject and the rules applying to these fields.

Checking your data source profile

Before submitting your file(s), please review the profile for your data source(s) in TESSy (go to **Data Sources**), and update the information, if necessary.

Complete and up-to-date data source information for each subject is important for improving interpretation of data - each surveillance system has different features that need to be taken into account when comparing data at an international level.

If your data source information is out-of-date and you do not have access rights to update it, please request your National Focal Point for Surveillance or National Coordinator to do so.

In-depth information on the data source variables is available in the TESSy user documentation.

Submitting your data

Data are submitted through the TESSy web interface (go to **Upload**).

Finalising your submission

The compliance of your data with the validation rules in the metadata is checked automatically during the data upload process.

The result of your upload – i.e. rejected or validated – is displayed immediately after the conclusion of the check in the **Validation details** webpage. Please review the result carefully:

- If your file has been rejected, there will be a message explaining each instance of non-compliance with the metadata that you need to correct.
- If your file has been validated, there might be warnings and remarks relating to possible data quality issues or to potential overwriting of existing records that you should consider.

When you file has been validated and you are satisfied that all corrections have been made, please ensure prompt approval – unapproved uploads can block for the approval of other uploads.

The TESSy user documentation provides information on reviewing validation results and adjusting reporting periods to avoid overwriting existing records.

Following submission, further validation checks are performed by ECDC disease experts. If issues are identified, experts may contact you for clarification and/or to upload corrected data.

Changes to current STI metadata

The coded value list for the variable SiteOfInfection has been changed to include NA. This is because some validation rules require the value NA.

The previous metadata changes are described in *Annex 1*.

Information on changes to the metadata for other subjects is available on the TESSy documentation website.

Annex 1 STI metadata

This section describes:

- *The STI metadata set*
- *Changes to the STI metadata since 2015*

Current record type versions

Table 1 shows the record type versions to be used when reporting 2018 STI surveillance data to TESSy.

Table 1: STI record type versions for 2018 data

Disease	Case-based record type version	Aggregated record type version
Chlamydia	CHLAM.3	STIAGGR.1
Gonorrhoea	GONO.3	STIAGGR.1
LGV	LGV.2	STIAGGR.1
Congenital Syphilis	CONSYPH.2	AGGR.1
Syphilis	SYPH.3	STIAGGR.1

STI metadata set

Table 2 lists the metadata variables for STI.

The common variables apply to reporting case-based data for all STI and consist of 20 variables. These variables can be divided into system-related and epidemiological variables.

The variables used for reporting aggregated data include age class, gender, case classification, date used for statistics, reporting country and the number of cases. The hierarchy for aggregated reporting is as follows:

1. Gender
2. Age class
3. Transmission

Table 2: STI metadata variables

Variable Name	Syphilis	Gonorrhoea	Chlamydia	LGV	Congenital Syphilis
Common variables					
1. RecordID	√	√	√	√	√
2. RecordType	√	√	√	√	√
3. RecordTypeVersion	√	√	√	√	√
4. Subject	√	√	√	√	√
5. Status	√	√	√	√	√

6. DataSource	√	√	√	√	√
7. Age	√	√	√	√	√
8. Gender	√	√	√	√	√
9. Outcome	NA	NA	NA	NA	√
10. DateofOnset	√	√	√	√	√
11. DateOfDiagnosis	√	√	√	√	√
12. DateOfNotification	√	√	√	√	√
13. DateUsedForStatistics	√	√	√	√	√
14. ReportingCountry	√	√	√	√	√
15. Classification	√	√	√	√	√
16. ClinicalCriteria	NA	NA	NA	NA	NA
17. LaboratoryResult	√	√	√	√	√
18. EpiLinked	NA	NA	NA	NA	NA
19. PlaceOfNotification	√	√	√	√	√
20. PlaceOfResidence	√	√	√	√	√
Disease-specific variables					
21. ClinicalServiceType	√	√	√	√	
22. CountryOfBirth	√	√		√	√
23. CountryOfNationality	√	√		√	
24. ProbableCountryOfInfection	√	√		√	
25. Transmission	√	√	√	√	
26. HIVStatus	√	√	√	√	
27. SexWorker	√	√		√	
28. ContactSW	√	√		√	
29. SiteOfInfection	√	√	√	√	
30. StageSYPH	√				
31. StageSYPHdetailed	√				
32. CountryOfBirthOfMother					√
33. CountryOfNationalityOfMother					√
34. AgeMonth					√

STI epidemiological variables

- In the Annual Epidemiological Report (AER), the date of diagnosis (as a proxy for date of consultation) is used for all analyses. If possible, please report this date as the Date used for Statistics.
- For chlamydia, gonorrhoea, syphilis, congenital syphilis and LGV, only confirmed cases should be reported at EU level. Cases must be classified according to EU case definitions. If different case definitions are used, please specify this in the Data Source data.
- When reporting the age-class for aggregated data, please use the following categories: 0-4, 5-14, 15-19, 20-24, 25-34, 35-44, 45-64, ≥ 65. This is coded in TESSy as AGECLASS2 (coded value list name), subset of codes 5_ (codes for age categories are preceded by 5_). If this age grouping is not applied, it might not be possible to use the

country data for parts of the analysis. If this coding is not possible, please contact the TESSy team to identify possible alternative options.

STI metadata change history

Metadata changes prior to 2015 can be found on the TESSy documents website.

Table 3: Summary of implemented changes to STI metadata

Year	Subject	Description
2015	**STI**	There were no STI metadata changes in 2015.
2016	**STI**	There were no STI metadata changes in 2016.
2017	**STI**	There were no STI metadata changes in 2017.
2018	**SYPH**	Coded value list for "SiteOfInfection" was changed to include the value NA. The value NA is required for a validation rule.

Annex 2 Subject-specific material

STI data collection

The surveillance data on five STIs are collected as part of the enhanced surveillance for STIs and concern cases that were diagnosed in 2018. **Case-based data are preferred**. If case-based data are not available, the aggregate format (RecordType: STIAGGR) broken down by 1) gender, 2) age class and 3) transmission may be used. Please note that reporting in aggregated format means that your country's data might not be used for some of the more detailed analyses.

Historical data (1990-2017)

Data for 1990-2017 has been collected in previous data collections and have been published in previous STI surveillance reports and the ECDC surveillance atlas. Historical data can be corrected by re-uploading to TESSy. These updates are taken into account for the AER and reflected in updates of the surveillance atlas if done before the deadline.

Historical data should be uploaded in the following ways:

- 1990-1999: annual data in **aggregated data** format broken down by 1) gender – 2) age class (preferred category) – 3) transmission. This means that 'Date of Statistics' is YYYY (1990-1999) only for the inclusion of historical data.
- 2000 onwards: **case-based data is preferred**. If case-based data are not available the aggregate format broken down by 1) gender, 2) age class and 3) transmission may be used.

STI data analysis plan

This data analysis plan describes the analyses performed for the STI chapters of the Annual Epidemiological Report. All analyses described below will be performed, however a selection will be used in the report. Conversely, additional analyses which are data-driven will be performed. Analyses that are considered to be useful for future inclusion in the AER will be included in future versions of this document. All analyses except where otherwise specified are based on the variable Date of Diagnosis as the date field.

The STI surveillance data will be presented as a chapter of the AER in an online report (*https://ecdc.europa.eu/en/annual-epidemiological-reports*) linked to the ECDC surveillance atlas of infectious diseases (*http://ecdc.europa.eu/en/data-tools/atlas/Pages/atlas.aspx*). A selection of analyses will be presented in the report and atlas.

Completeness and origin of reporting

The completeness of data will be analysed overall. Completeness by country will also be analysed and fed back to countries to show how they fare compared to the overall EU/EEA average. An analysis of the 'Datasource' variables will give an overview by country of the availability of data in 2018 and provides an overview of the origin of the data. The information is needed to be able to interpret the actual data on STIs. The main characteristics which will be assessed include: type of system (comprehensive, sentinel), type of reporting (clinical, laboratory), legal framework (compulsory, voluntary).

Data will be presented in the following in tables:

- Overview of reported (aggregated or case-based) data per country and time period by STI. Aggregate and case-based datasets need to be merged for the analyses.
- Overview of country data sources per country by STI.

- Overview of completeness (value versus unknown or not reported) per variable overall and by reporting country per STI.

Preparing datasets for analyses

- Case-based and aggregate datasets are merged for each STI.
- Country population denominators used to calculate rates are based on data from the Eurostat database. Rates are not calculated for countries with sentinel surveillance systems unless the countries have provided ECDC with denominator data and coverage for the specific surveillance system.
- Recode age into age class: <15, 15-19, 20-24, 25-34, 35-44, 45+; if data on age are unavailable or coded in the wrong age-class, the country will be excluded from the analysis;
- Total numbers in tables include unknowns; total numbers in graphs exclude unknowns except where specified.
- Country of birth, country of nationality and reporting country: these variables are explored by country/region and to be recoded after this exercise to describe the origin of the reported cases. Country of nationality is the most important variable, if not available the country of birth should be used instead.
- Probable country of infection will be recoded into 'acquired abroad' and 'acquired in home country'.
- HIV status: tabulate HIV status including the unknowns; recode into HIV positive, HIV negative, HIV unknown/ NA.
- Site of infection: tabulate site of infection per STI including the unknowns.

Descriptive epidemiology

The analyses will describe the occurrence of STI by time, place and population across countries and group countries and variables if possible.

Overall numbers and rates

- Present number of STI cases for the last 10 years, overall and by country.
- Rates are calculated per 100 000 population for chlamydia, gonorrhoea and syphilis; rates for congenital syphilis are calculated per 100 000 live births.
- Present STI rates for the last 10 years, overall and by country (only for countries with comprehensive surveillance systems or where an adjustment can be made).
- Calculate the change in STI rates (and number of cases) over the previous five years and compared to the previous year. Other comparisons will be done based on the epidemiology of the infection.

Gender

- Present number of STI cases by gender for the last 10 years, overall and by country.
- Present STI rates by gender for the last 10 years, overall and by country.
- Calculate the increase by gender by STI over the last five years and compared to the previous year.
- Gender distribution: male-to-female ratio by reporting country per STI for the report year; compare the trend in male-to-female ratio by STI.

Age

- Age distribution: present the age distribution by STI for the report year.
- Age-specific rates: present age-specific rates for the report year and compare with age-specific rates for the previous nine years. Describe trends in age-specific rates.

Age-Gender

- Present the age distribution by gender and by STI in the report year. Describe recent trends in age and gender specific rates.

Transmission

- Present proportion of cases reported in the report year by transmission category. Describe trends, changes in distribution over the last year, five years, ten years.
- For men who have sex with men (MSM) (Transmission=MSM): trends in STI by year, for age groups ≤ and > 35 years of age; calculate rates based on number of men with MSM transmission with denominator being male population; explore grouping of countries; for case-based data only (possibly include aggregate dataset if 'transmission' is provided).
- Transmission: present percentage of MSM by country per STI for the report year; compare with the trend in percentage MSM by STI over last five and ten years; (except for congenital syphilis). (Case-based data only; possibly include aggregate dataset if 'transmission' is provided).

Geographical information

- Present the country of nationality (home country versus other) by country per STI in the report year; present the probable country of infection (as percentage acquired abroad) per STI in the report year.
- For congenital syphilis: present the country of birth/nationality of mother by country (home country versus other).
- Examine the completeness of subnational data – provide examples for countries where completeness is higher than 80%.

HIV Status

- Present the proportion of HIV positive by country per STI for the report year; present percentage HIV positive by gender and transmission category per STI over the last ten years.
- Compare proportion of HIV positive cases with overall HIV rates taken from HIV report, by country.

Sex work

- Percentages of SW/CSW by country per STI by gender (and in total numbers) (except for chlamydia), if completeness allows.

Site of infection

- Proportion of cases by site of infection, by transmission category for chlamydia, gonorrhoea and primary syphilis for the report year.

Clinical Service

- Clinical service: present the distribution of clinical services in percentage by country per STI for the report year (except for congenital syphilis). Present by grouping similar services. Analyse any changes over time.

Chlamydia – analysis by level of chlamydia control activities

- Present overall rates of chlamydia by country grouped according to level of chlamydia control (as presented in "Chlamydia Control in Europe - A survey in the Member States 2012").
- Analyse overall rates of gonorrhoea and syphilis by level of chlamydia control to identify any associations.

Syphilis – infectiousness

- Syphilis: present the percentage in infectious syphilis and by stage of infection for the report year and investigate any changes over last five and ten years.

Depending on the interpretation of the above results more detailed analyses may be suggested regarding specific STI, risk group or countries.