

EPILABO: DATA COLLECTION VIA HD4DP 2.0 APPLICATION

DOCUMENTATION FOR PARTICIPATING LABORATORIES

Preamble

The purpose of this document is to provide general information on the Epilabo project, and more precisely on the new data collection put in place via the healthdata.be platform. The document is structured as followed:

1. Outline of the overall goals of the Epilabo project and improvements of the new data collection;
2. Description of the collected variables;
3. List of pathogens in the scope of Epilabo;
4. Technical information on how to submit data via the HD4DP 2.0 application.

Authors & Contact information

Pierre Hubin¹, Veerle Stouten¹, Dieter Vancauteran¹, Amber Van Laer¹, Marijke Pauwels²

¹ Epidemiology of Infectious Diseases, Sciensano

² Healthdata.be, Sciensano

Contact e-mail address: sentilab@sciensano.be

Links to relevant documentation

Epilabo project description:

<https://www.sciensano.be/en/network-sentinel-laboratories-epilabo>

Laboratory data networks project description:

<https://www.sciensano.be/en/projects/laboratory-data-networks-surveillance-infectious-diseases>

Be-HERA project description:

<https://www.sciensano.be/en/projects/consolidation-national-infrastructure-combined-microbiological-genomic-epidemiological-surveillance>

HD4DP 2.0 technical documentation:

<https://docs.healthdata.be/documentation/hd4dp-v2-health-data-data-providers/user-manual-application-hd4dp-v2>

FAIR documentation on Epilabo (including Data Collection Definition and other documents):

<https://fair.healthdata.be/dataset/4f4f0d95-921d-4247-917e-b9b91871d94d>

Information Security Committee deliberation:

<https://www.ehealth.fgov.be/ehealthplatform/file/view/AYU5CAbouwVJMAnc0FXx?filename=17-010-n552-EPILABO-gewijzigd%20op%206%20december%202022.pdf> (NL)

https://www.ehealth.fgov.be/ehealthplatform/file/view/AYU5B_PpuwVJMAnc0FXw?filename=17-010-f552-EPILABO-modifi%C3%A9e%20le%206%20d%C3%A9cembre%202022.pdf (FR)

1. Epilabo: Description of the project and objectives

The Epilabo project consists of a network of sentinel clinical microbiology laboratories coordinated by Sciensano. The system has been operating since 1983 and relies upon the voluntary participation of the laboratories in the weekly submission of diagnostic data for more than 40 pathogens. Epilabo contributes to improving public health by:

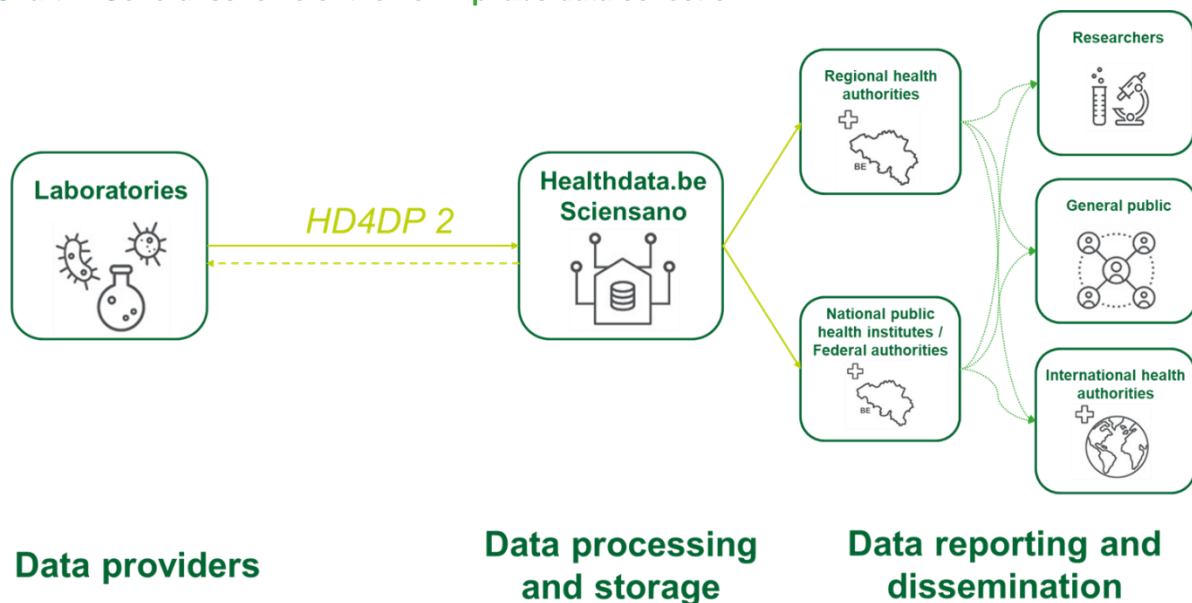
- monitoring epidemiological trends, and detecting potential outbreaks;
- assessing the impact of prevention programs;
- predicting the potential spread of infectious diseases;
- reporting data to international health authorities (ECDC, WHO).

As part of the continuation of the Epilabo project, a new data collection has been set up via the healthdata.be platform. **Chart 1** provides an overview of the different stakeholders involved from data collection to final use of the data. These new developments will lead to significant improvements in terms of data security, integrity, and quality. Data security will be guaranteed by relying on the healthdata.be platform as the processing and storage of health-related data is the main mission of this service¹.

Data collection takes place through the “HealthData 4 data providers” (HD4DP) 2.0 application², which is developed and maintained by healthdata.be. Relying on this platform for data transfer also contributes to improved data security and integrity. Data can be uploaded via the application in three ways:

1. HD4DP 2.0 allows a System-to-System (S2S) transfer of JSON files via an API.
2. Records can be submitted in batch by uploading a CSV file to a dedicated sFTP server.
3. A form, accessible from a web browser, can be filled in for a single record entry.

Chart 1: General scheme of the new Epilabo data collection.



More information on the technical details of data submission is provided in section 4 of this document. Next to the data upload process, the well-established data management and validation process put in place from the receipt of the raw data to its arrival in the analysis environment also contributes to the

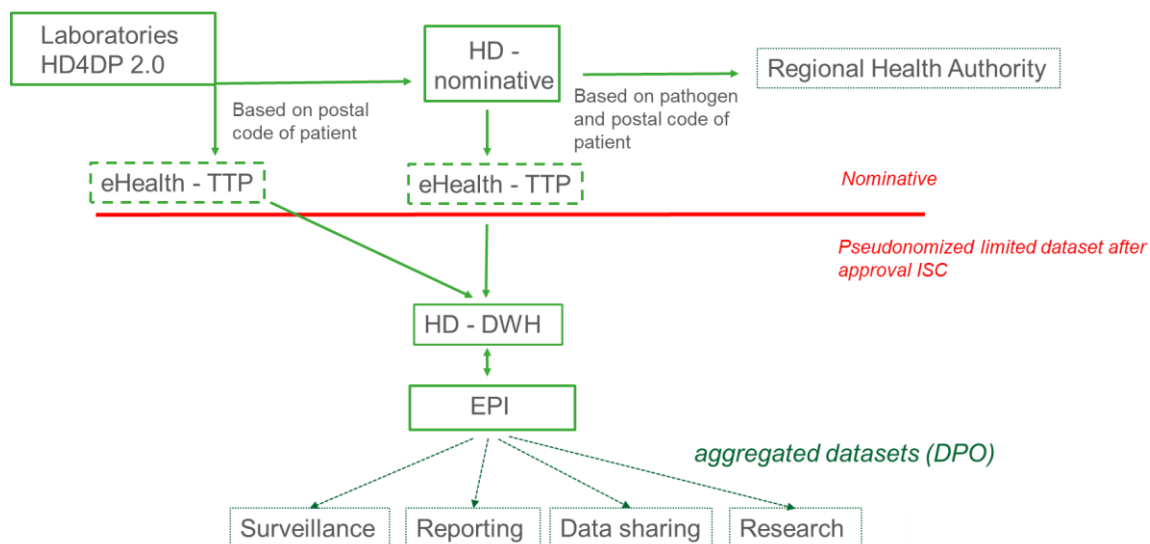
¹ For more information on objectives of healthdata.be: <https://healthdata.sciensano.be/fr/%C3%A1-propos-healthdatabe>
² Health Data For Data Providers version 2 (HD4DP 2.0) is an electronic data capture system, designed for the collection of clinical data in electronic format: <https://healthdata.sciensano.be/fr/hd4dp>

improvement of data integrity and quality. Finally, the use of coding systems, relying on international standards when available, to standardize the information collected is a major improvement towards data quality and interoperability. Specifically, LOINC³ and SNOMED CT⁴ codes are used to code sample types, test methods, and pathogens. The coding system was developed in collaboration with ReTaM⁵ and microbiology laboratories.

Additionally, feedback will be provided to the participating laboratories. First, the HD4DP 2.0 application itself offers functionalities for data providers to monitor the submission of the data files in real-time. Secondly, automated reports summarizing key information about the content of the data uploaded by each participating laboratory will be sent on a weekly basis. Lastly, an interactive application will be developed enabling exploration of the aggregated datasets and displaying epidemiological indicators to inform healthcare professionals and the general public.

A particular feature of the new data collection system is that it allows to collect either pseudonymized or nominative data. Chart 2 displays the overall dataflow showing both options. In case of the pseudonymized data flow, pseudonymization occurs via eHealth before reaching the data warehouse of healthdata.be. On the other hand, the nominative data collection, which is based on the national social security number (NISS), goes through the Be-HERA platform⁶. Via the latter flow, relevant data can be transferred in real-time to the regional health authorities in order to support the mandatory notification systems for several infectious diseases. This nominative data collection can also facilitate the process of data linkage of the laboratory test results with other health and/or administrative databases. Such a linked data infrastructure is an important objective to improve surveillance of infectious diseases for public health. It should be pointed out that nominative data is retained for a limited period of time on the Be-HERA platform and can only be accessed by a specific team of healthdata.be.

Chart 2: Schematic representation of the nominative/pseudonymized dataflows within the new Epilabo data collection.



³ Logical Observation Identifiers Names and Code (LOINC) is an international standard for identifying health measurements, observations, and documents: <https://loinc.org/>

⁴ Systematized Nomenclature of Medicine Clinical Terms (SNOMED CT) determines international standard for medical terms : <https://www.snomed.org/>

⁵ Reference Table Management (ReTaM) is visual representation of the authentic source of laboratory codes: <https://www.ehealth.fgov.be/standards/kmehr/en/page/retam-exports>

⁶ HERA-BE-WGS, defined under the Health Emergency Preparedness and Response Authority (HERA) supports the central collection, processing and linking of clinical/epidemiological data with microbial genomic indicator data derived from whole genome sequencing (WGS) analysis: <https://www.sciensano.be/en/projects/consolidation-national-infrastructure-combined-microbiological-genomic-epidemiological-surveillance>

The COVID-19 pandemic highlighted the value of having a robust, digital system for timely data collection of laboratory test results at the national level, on which surveillance can be based. It is in this context of pandemic preparedness that the Epilabo developments continue, with the improved data collection being set up to ensure reliability, scalability (capability to handle large amounts of data), and flexibility (possibility to add newly emerging pathogens and/or required data collection fields depending on the situation). Furthermore, the COVID-19 crisis has also proven the need to move towards more interoperable systems at the supra-national level. The efforts regarding the new data collection system of Epilabo also aim at enhancing its interoperability, in particular by the use of international standards for the coding system of laboratory test results. This is in line with the objectives of EU-HIP⁷, a European-funded project aiming at the development of IT health systems of the Member States to improve interoperability and allowing better health threat assessment at the European level.

⁷ EU-HIP, EU interoperability with HERA's IT platform develops new IT systems and strengthens existing national IT systems for improving health threat assessments in European countries: <https://www.sciensano.be/en/projects/eu-interoperability-heras-it-platform-0>

2. Description of variables to collect

The list of required variables to collect is provided in Table 1. Links to templates in CSV and JSON formats that can be used to submit data are provided in the **Annex**.

The Data Collection Definition (DCD) provided in **Links to relevant documentation** includes the list of LOINC and SNOMED-CT codes that can be used to code test results. The DCD and templates also include additional non-mandatory variables which can be collected. For example additional contact information on the health professional prescribing the test can be provided and might be useful for the regional health authorities in the context of mandatory notifiable infectious diseases. List of codes for other variables such as patient gender are also defined in the DCD. Negative test results can also be reported, they represent an important complementary piece of information for several pathogens.

Table 1: List of variables to collect in the new Epilabo data collection system

Logical Name	Description	Clinical Building Block	Required	Type
PatientIdentificationNumber_System	Type of identification system (eg NISS)	HdBe-Patient	Y	From list
PatientIdentificationNumber_Value	Patient identification number	HdBe-Patient	Y	Text
DateOfBirth	Patient date of birth	HdBe-Patient	Y	DD/MM/YYYY
Gender	Patient gender	HdBe-Patient	Y	From list
AddressInformation_Postcode	Patient postcode	HdBe-Patient	Y	From list
Requester_HealthProfessional_HealthProfessionalIdentificationNumber_System	Health professional identification system (requester)	HdBe-LaboratoryTestResult	Y	From list
Requester_HealthProfessional_HealthProfessionalIdentificationNumber_Value	Health professional identification value (requester)	HdBe-LaboratoryTestResult	Y	From list
Performer_HealthProfessional_HealthProfessionalIdentificationNumber_System	Health professional identification system (performer)	HdBe-LaboratoryTestResult	Y	From list
Performer_HealthProfessional_HealthProfessionalIdentificationNumber_Value	Health professional identification value (performer)	HdBe-LaboratoryTestResult	Y	From list
Performer_HealthProfessional_HealthcareOrganization_HealthcareOrganizationIdentificationNumber_System	Healthcare organization identification system (performer)	HdBe-LaboratoryTestResult	Y	From list
Performer_HealthProfessional_HealthcareOrganization_HealthcareOrganizationIdentificationNumber_Value	Healthcare organization identification value (performer)	HdBe-LaboratoryTestResult	Y	From list
Specimen_SpecimenId	Specimen ID	HdBe-LaboratoryTestResult	Y	Text
Specimen_CollectionDateTime	Collection date/time	HdBe-LaboratoryTestResult	Y	DD/MM/YYYY: Time
LaboratoryTest_TestCode	Test code	HdBe-LaboratoryTestResult	Y	From list (LOINC)
LaboratoryTest_TestDateTime	Test date/time	HdBe-LaboratoryTestResult	Y	DD/MM/YYYY: Time
LaboratoryTest_TestResultID	Test result ID	HdBe-LaboratoryTestResult	Y	Text
LaboratoryTest_TestResultCodeableConcept	Test result (qualitative)	HdBe-LaboratoryTestResult	Y	From list (SNOMED CT if culture)

3. List of pathogens

The list of pathogens in the scope of the Epilabo project is provided in Table 2. This list will be evaluated on a yearly basis and may evolve over time.

Table 2: List of pathogens in the scope of Epilabo

Adenovirus	Hantavirus ^{a,b}	Norovirus
Bartonella	Hepatitis A Virus (HAV) ^{a,b}	Parainfluenza
<i>B. pertussis</i> ^{a,b}	Hepatitis B Virus (HBV)	Plasmodium ^{a,b}
<i>B. burgdorferi</i>	Hepatitis C Virus (HCV)	Respiratory Syncytial Virus (RSV)
Campylobacter	Hepatitis E Virus (HEV)	Rotavirus
<i>C. psittaci</i> ^{a,b}	Herpes simplex virus 1 en 2	Rubivirus ^{a,b}
<i>C. trachomatis</i>	Influenza A	Salmonella
Cryptococcus	Influenza B	Shigella ^b
Cryptosporidium	Legionella pneumoniae ^{a,b}	<i>S. pneumoniae</i>
Cyclospora	Listeria ^{a,b}	<i>S. pyogenes</i> ^{a,b}
<i>E. coli</i> (VTEC-EHEC) ^{a,b}	Morbillivirus ^{a,b}	<i>T. pallidum</i> ^a
<i>E. histolytica</i>	Measles virus	Varicella zoster virus in cerebrospinal fluid
Enterovirus	<i>M. pneumoniae</i>	<i>Y. enterocolitica</i>
Giardia	<i>N. gonorrhoeae</i>	Mpox ^{a,b}
<i>H. influenzae</i> ^{a,b}	<i>N. meningitidis</i> ^{a,b}	SARS-CoV-2 ^{a,b}

^aPathogen subject to mandatory notifiable disease for AViQ (Wallonia)

^bPathogen subject to mandatory notifiable disease for COCOM (Brussels)

4. Technical information to submit data via HD4DP2.0

General information about the HD4DP 2.0 application as well as detailed technical guides describing how to submit data is provided on the first page of this document in **Links to relevant documentation**. More specifically the technical documentation detailing S2S implementation and CSV upload via sFTP is available in the link below:

<https://docs.healthdata.be/documentation/hd4dp-v2-health-data-data-providers/technical-manual-application-hd4dp-v2>.

A recording of a live demo of data uploads to the HD4DP 2.0 application via S2S (JSON) and sFTP (CSV) is available in the following link:

<https://docs.healthdata.be/documentation/hd-ict-information-sessions/ict-information-session-20230515>

5. Annex

Templates in CSV (HD_DCD_submcsv_HDBP0037_EPILABO_All_Repeatables_Full_v1_20230927_HEADER.csv) and JSON (HD_DCD_submjson_HDBP0037_EPILABO_All_Repeatables_Full_v1_20230927.json) formats as well as the DCD documentation for Epilabo project are available on the FAIR portal, under the Resources tab. Link to the fair portal is given in **Links to relevant documentation**. The templates include all variables that can be submitted to the Epilabo platform, however not all variables should be completed in order to be able to submit the file through HD4DP2.0. The list of required variables is given in Table 1.