



Co-funded by the
European Union

The project „Genomic surveillance of selected infectious diseases in the Czech Republic“ (HERA2CZ) is co-funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the European Health and Digital Executive Agency (HADEA). Neither the European Union nor the European Health and Digital Executive Agency (HADEA) can be held responsible for them.

D1.9 – Sequencing Data Sharing Report

<i>Project Title:</i>	<i>Genomic surveillance of selected infectious diseases in the Czech Republic</i>
<i>Project Acronym:</i>	<i>HERA2CZ</i>
<i>Project ID:</i>	<i>Grant Agreement - Project 101113387</i>
<i>Type of action:</i>	<i>EU4H Project Grants</i>
<i>Call:</i>	<i>EU4H-2022-DGA-MS-IBA-1</i>
<i>Topic:</i>	<i>EU4H-2022-DGA-MS-IBA-01-02</i>
<i>Beneficiary:</i>	<i>National Institute of Public Health, the Czech Republic</i>
<i>Coordinator:</i>	<i>Jana Kozáková, MD</i>
<i>Authors of the Deliverable:</i>	<i>Jana Kozáková (WP1), Helena Jiřincová (WP2), Martina Bielaszewska (WP3), Jana Kozáková (WP4), Helena Žemličková (WP5), Markéta Šimková</i>
<i>Date of the Deliverable:</i>	<i>17th December 2024</i>

National Institute of Public Health

Genomic surveillance of selected infectious diseases in the Czech Republic, HERA2CZ

Project: 101113387 — HERA2CZ — EU4H-2022-DGA-MS-IBA-1

D1.9 – Sequencing Data Sharing Report

Disclaimer



Co-funded by the
European Union

Co-Funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or of the European Health and Digital Executive Agency (HADEA). Neither the European Union nor the European Health and Digital Executive Agency (HADEA) can be held responsible for them.

Dissemination Level

PU	Public	<input type="checkbox"/>
SEN	Sensitive	<input checked="" type="checkbox"/>

History of Changes

Version	Date	Name	Chapters edited	Changes tracked
v1		<i>J. Kozáková, H. Jiřincová, M. Bielaszewska, O. Daniel, P. Křížová, Z. Okonji H. Žemličková, V. Jakubů, M. Šimková</i>	All	Document prepared

National Institute of Public Health

Genomic surveillance of selected infectious diseases in the Czech Republic, HERA2CZ

Project: 101113387 — HERA2CZ — EU4H-2022-DGA-MS-IBA-1

D1.9 – Sequencing Data Sharing Report

Table of Contents

Table of Contents.....	3
List of Abbreviations	4
List of Figures	4
1. Executive Summary	5
2. Project's Baseline.....	7
2.1. Description of the Current State of Legislation	8
2.2. Current State of the Laboratory Equipment	11
2.3. State of Technological Infrastructure for Data Sharing	12
2.4. Security, Data Management and Data Storage	12
2.5. Personnel Situation.....	12
2.6. Data Processing.....	12
2.7. Expected Development in Sequencing Field.....	13
2.8. Data Standardization, Procedures, Formats, and Data Quality Control	15
2.9. State-of-the-art Definition of Data Sharing	18
3. Project Phase	20
3.1. Data Handling in the Project, Changes During the Project, and Progress Towards Process Automation	20
3.2. Automation, Applications, New Data Visualization Tools.....	21
4. Future Outlook	22
4.1. How the Project Will Contribute to Improved Data Handling	22
4.2. How Data Generated in the Project Will be Included in ECDC's Regular Reporting and How and Where Else They Will be Published	23
5. Project's Context - Genomic Surveillance of Selected Infectious Diseases in the Czech Republic (HERA2CZ)	24

National Institute of Public Health

Genomic surveillance of selected infectious diseases in the Czech Republic, HERA2CZ

Project: 101113387 — HERA2CZ — EU4H-2022-DGA-MS-IBA-1

D1.9 – Sequencing Data Sharing Report

List of Abbreviations

AMR	<i>Antimicrobial Resistance</i>
ARI	<i>Acute Respiratory Infections</i>
ARI/ILI surveillance	<i>Acute Respiratory Infection and Influenza-Like Illness Surveillance National System</i>
CCRE	<i>Carbapenem and/or Colistin Resistant Enterobacterales</i>
CNV	<i>Copy Number Variants</i>
EQA	<i>External Quality Assessment</i>
FWD	<i>Food- And Water-Borne Diseases</i>
GISRS	<i>Global Influenza Surveillance and Response System</i>
GSD	<i>Genetic Sequence Data</i>
ILI	<i>Influenza-Like Illnesses</i>
IVPP	<i>Influenza Viruses with Human Pandemic Potential</i>
MDR	<i>Multidrug Resistant Bacteria</i>
MLST	<i>Multilocus Sequence Typing</i>
MPXV	<i>Monkeypox Viruses</i>
NIPH	<i>National Institute of Public Health</i>
NRLs	<i>National Reference Laboratories</i>
PC	<i>Partnership Contribution</i>
PIP	<i>Pandemic Influenza Preparedness</i>
QC	<i>Quality Control</i>
RSV	<i>Respiratory Syncytial Viruses</i>
RT-PCR	<i>Reverse Transcription Polymerase Chain Reaction</i>
SARI	<i>Severe Acute Respiratory Infections</i>
SMTA2	<i>Standard Material Transfer Agreements</i>
SNP	<i>Single Nucleotide Polymorphism</i>
SV	<i>Structural Variants</i>
WGS	<i>Whole Genome Sequencing</i>
WHE	<i>WHO Health Emergencies Programme</i>

List of Figures

Table 1: Criteria for assessing assembly-level quality, Mbp – Mega base pairs, kb - kilobases 17

1. Executive Summary

The **Sequencing Data Sharing Report (D1.9)** builds upon the original **Complementarity Strategy (D1.6)**. This updated document delineates progress in sharing sequencing data, reflecting the evolving practices in data collection, management, and sharing in the field of genomics at the National Institute of Public Health (NIPH) in the Czech Republic. The NIPH's primary mission is to prevent and promote public health, while secondary objectives include fostering research, enhancing patient care, and maintaining security and ethical standards.

The D1.9 strategy is integral to the **HERA2CZ project**, which aims to strengthen the technological capabilities of National Reference Laboratories (NRLs). It promotes data sharing and the responsible, transparent utilization of genomic information, while adhering to privacy regulations that safeguard patient rights.

In conjunction with the Sustainability Strategy (D1.5), the original Complementarity Strategy (D1.6) laid the groundwork for ensuring that the HERA2CZ project's outcomes are enduring, accessible, and operational in the long term. The D1.9 report provides an updated summary of progress made during the project and anticipates future developments in this area.

To enhance clarity and usability, the D1.9 document is organized into three primary sections:

1. **Projects' Baseline** – Detailing the initial conditions and challenges before project initiation.
2. **Project Phase** – Highlighting the transformations underway during the project's execution.
3. **Future Outlook** – Anticipating advancements and implications beyond the project timeline.

Projects' Baseline

- **Legislative Background:** Overview of national and international legal frameworks pertinent to the HERA2CZ project.
- **Laboratory Equipment:** Assessment of available laboratory resources at the project's inception.
- **Technological Infrastructure for Data Sharing:** Evaluation of IT equipment and capabilities at the start of the project.
- **Security and Data Storage:** Initial techniques and protocols for data storage and security.
- **Personnel:** Assessment of human resources and specialization requirements.
- **Data Processing Procedures:** Review of sequencing data handling, processing, and storage practices.
- **Expected Development:** Insights into anticipated advances in sequencing technologies, legislation, and data management.
- **Data Standardization, Procedures, Formats, and Quality Control:** Framework for ensuring data consistency and reliability.
- **State-of-the-Art of Data Sharing:** Current practices in data sharing and storage at national and international levels.

Project Phase

- **Handling of Data Generated in the Project:** Overview of improvements in data processes, including potential automation.
- **Automation, Applications, and New Tools for Data Visualization:** Discuss application of innovative technical tools within the HERA2CZ project.

National Institute of Public Health

Genomic surveillance of selected infectious diseases in the Czech Republic, HERA2CZ

Project: 101113387 — HERA2CZ — EU4H-2022-DGA-MS-IBA-1

D1.9 – Sequencing Data Sharing Report

Future Outlook

- **Better Data Management:** Project's impact on enhancing data management practices.
- **Inclusion of Data in ECDC Reporting:** Processes for integrating project-generated data into ECDC reporting and publication channels.

The **Sequencing Data Sharing Report (D1.9)** will be complemented by **Sustainability report (D1.8)** upon the project's conclusion. D1.8 will serve as a definitive summary of progress achieved during the HERA2CZ project. It will document advancements in genomics data sharing and provide a foundational resource for future initiatives in the healthcare sector. The report's insights and frameworks will guide efforts to ensure responsible and effective use of genomic data, ultimately supporting public health and research objectives.