

Sequencing strategy and capacity building - HERA project 2021/2022

Comparison of the geographic sequencing coverage and the representativity of the sequenced samples

David Vostřák, Alena Fialová, Martina Leppová, Monika Liptáková,
Jan Moskalyk, Helena Jiřincová



1. Targeted surveillance:

- Vaccination failure
- Reinfection
- Hospitalisation < 60 YoA
- Imported cases
- Local outbreaks
- Unusual clinical course

2. Surveillance based on representative selection (geography, age, gender)

- Nine sequencing centers, plus laboratories that were not /cannot be funded from MoH

International support:

Cross-border Saxony sequencing support

ECDC sequencing support

Regional sequencing support (one laboratory funded from the regional financial sources)

FoSCU - one laboratory funded from the Hera project 2021/22

Does a surveillance system based on random sampling ensure the sequencing of a representative proportion of samples?

Find out the extent to which the sequenced data evenly cover: age distribution, geographic distribution, and gender requirements.

Method:

The dataset from GISAID was used only to verify the representativeness of the sample, regardless of other parameters. There are no necessary data in the GISAID database for the analysis of targeted surveillance. This analysis will be performed with the data set obtained from ÚZIS in April 2022.

The months selected for pilot verification were:

- September 2021 - dominance of the delta variant, school testing, increase in incidence
- November 2021 - introduction of omikron variant, mandatory testing of employees, high incidence
- February 2022 – dominance of omicron variant (BA.1/BA.2ú, downward trend in incidence
- April 2022 – dominance of omicron variant BA.5, decreasing trend of testing capacity

List of laboratories that participated in WGS SARS-CoV-2 in selected months

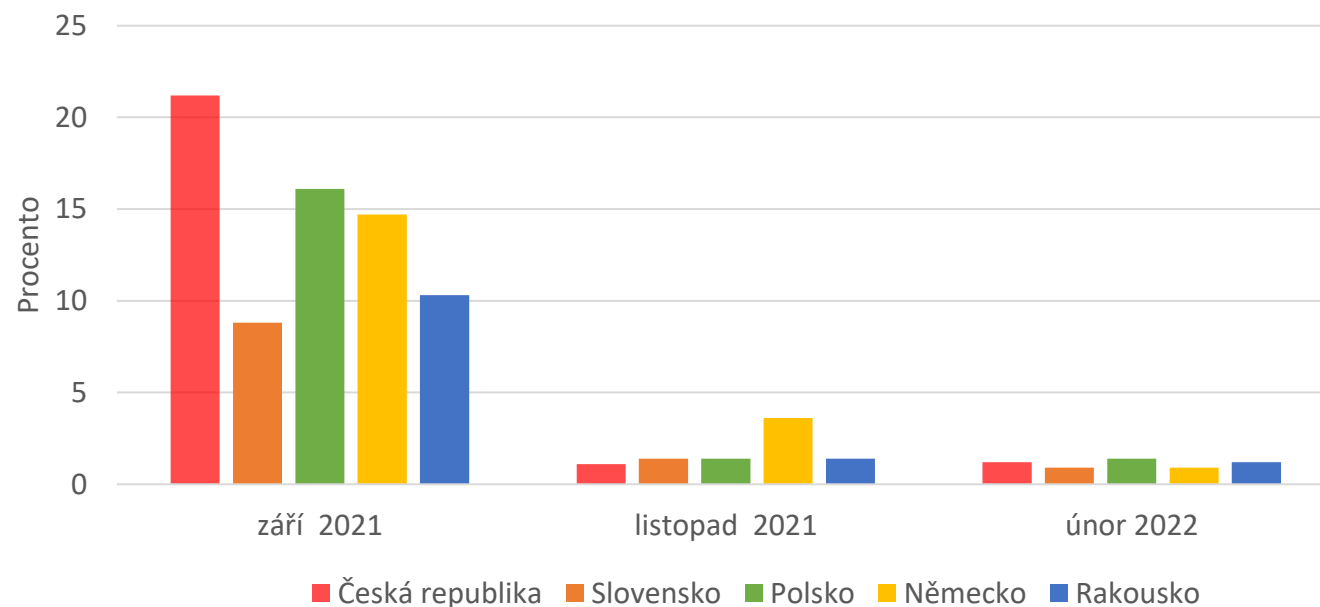
- CMBG FN Brno
- FN Ostrava
- FN Plzeň
- FN Sv. Anny Brno
- IMTM Olomouc
- KN Hradec Králové
- KN Liberec
- NRL NIPH , included ECDC sequenced samples
- Academy of Science CB
- FoSCu
- TU Dresden Dresden (Cross border Saxony cooperation, ukončena konec března 2022).

Zdroj dat: GISAID (<https://www.gisaid.org/>)

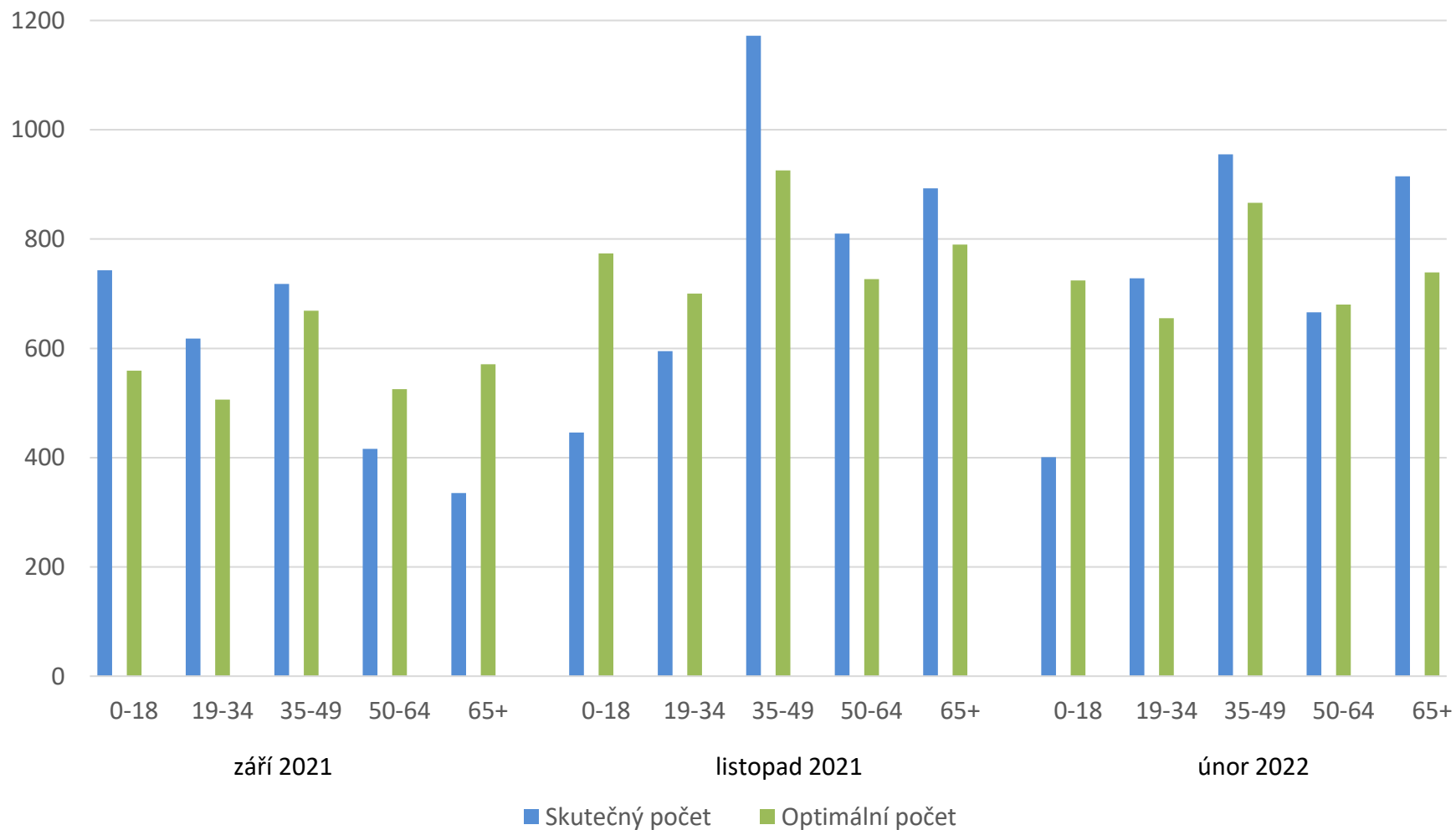
Number of WGS uploaded to GISAID in selected

měsíc/rok	počet sekvenací
září 2021	2845
listopad 2021	4025
únor 2022	3672

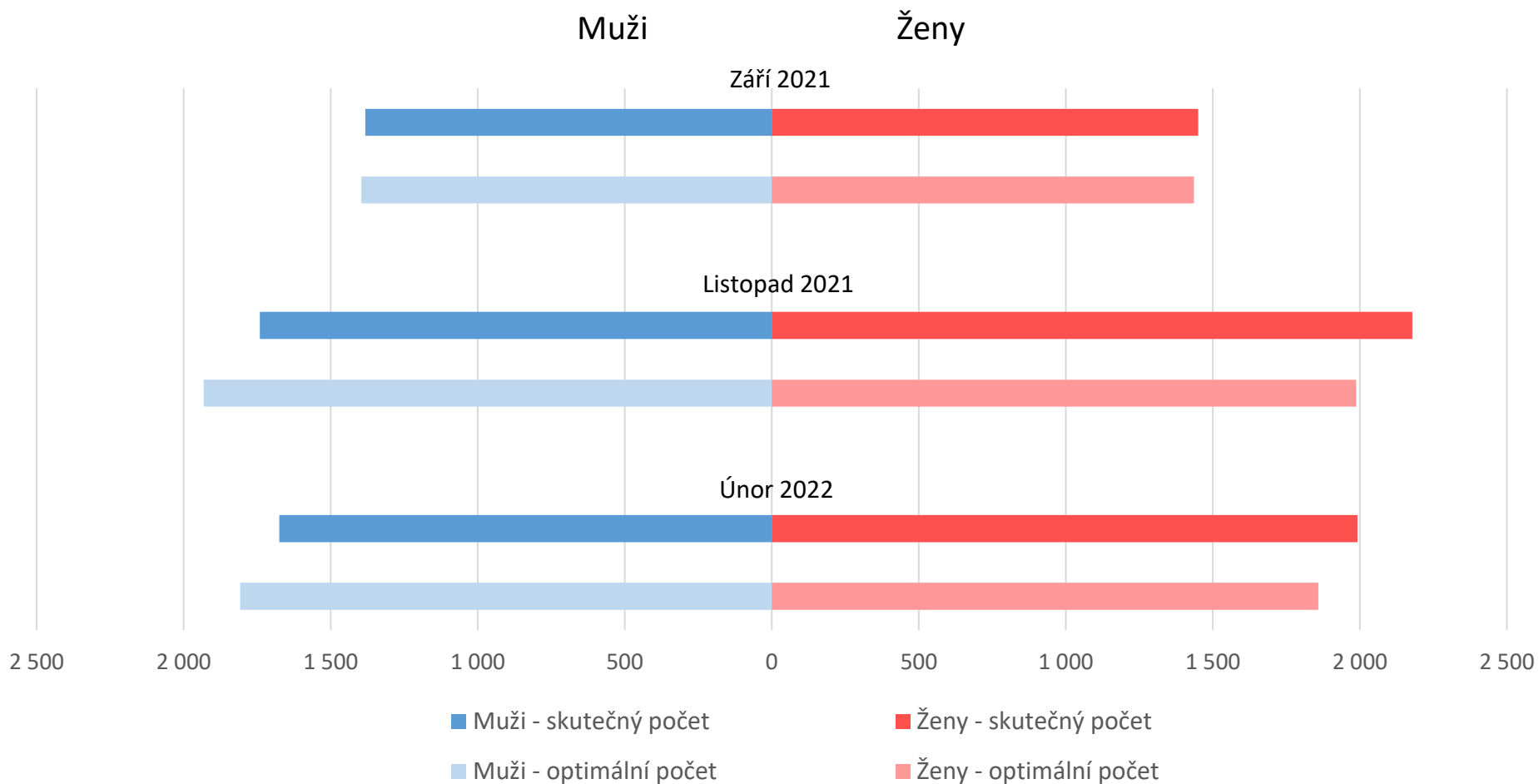
Ratio (in %) of sequenced samples from all samples positive for SARS-CoV-2 in selected months versus incidence



Graph 1: Comparison of the real and optimal number of WGS by age groups in selected months. The optimal number is related to the actual proportion of the age group in the population.

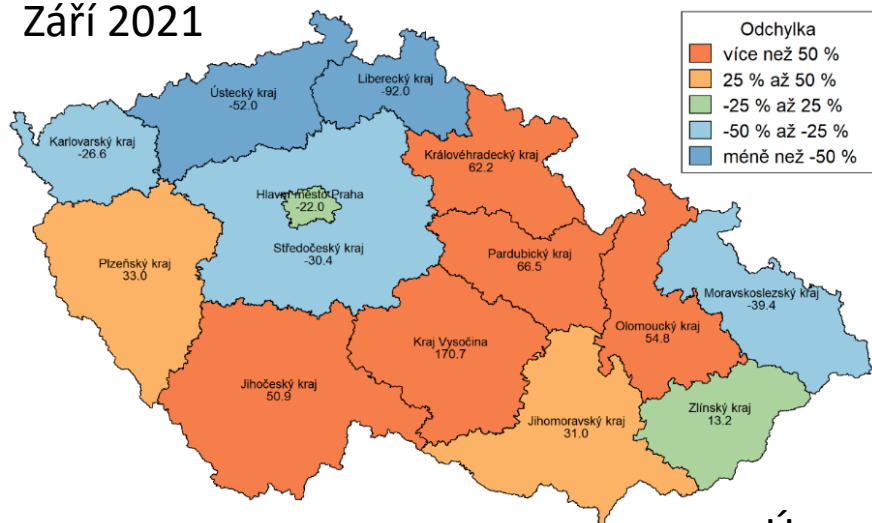


Graph 2: Comparison of the real and optimal number of WGS by gender in selected months
The optimal number is related to the actual proportion of men and women in the population

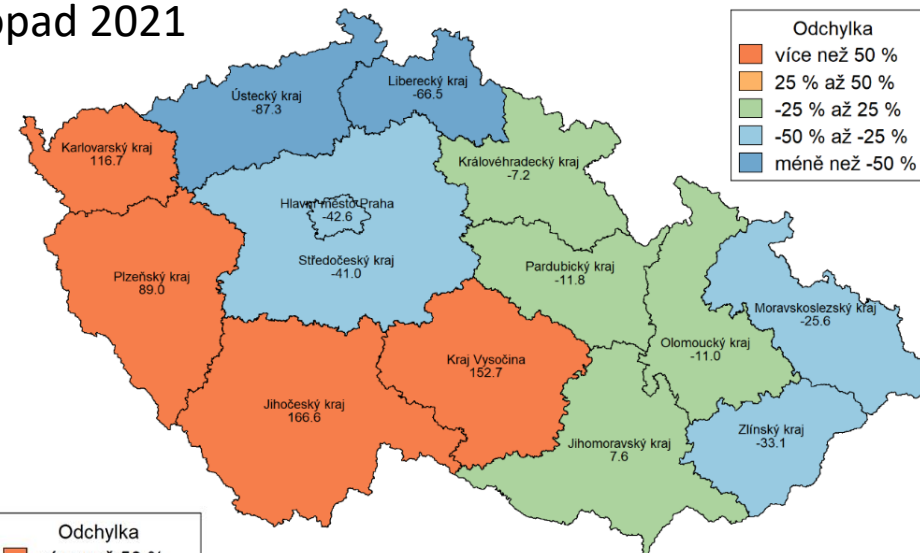


**Obrázek 3: Porovnání skutečného a optimálního počtu sekvenací dle kraje za vybrané měsíce.
Optimální počet je vztážen ke skutečnému podílu všech pozitivních vzorků v kraji za vybraný měsíc.**

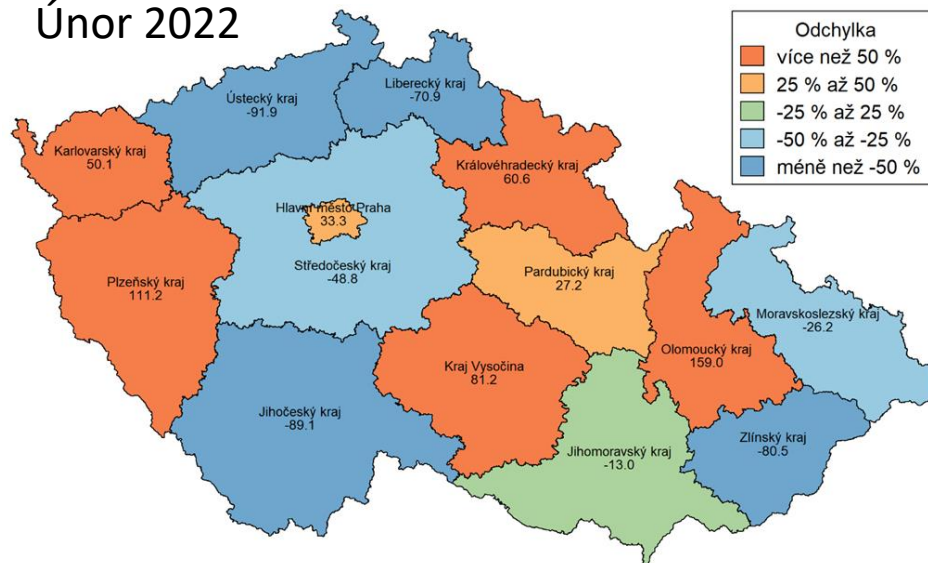
Září 2021



Listopad 2021



Únor 2022



Komentář ke grafům 1, 2 a obrázku 3:

- Odchytky od optimálního stavu z hlediska věkových skupin a pohlaví nejsou příliš velké s výjimkou věkové skupiny 0 – 18 let.
- V případě věkových skupin se do sekvenovaných vzorků přelévá počet vyšetřených vzorků v dané věkové skupině, který je spíše závislý na testovací strategii než na incidenci (testování ve školách – září 2021, testování v zaměstnání – listopad 2021).
- Poměrně významná neprosekvenovatelnost v kategorii 0 – 18 let byla v listopadu 2021 a v únoru 2022.
- Velká disproporce je v distribuci mezi kraji.
- Zavedení národního surveillančního algoritmu by mělo vyřešit vyvážení dle **všech** parametrů.

When analyzing data from GISAID, where the type of sampling (representative/targeted sampling) cannot be separated, it was found:

- Within the age distribution, the 0-18 years category is permanently underestimated, except for the period targeted for testing in schools.
- The distribution of samples according to the gender of persons is represented equally.
- The regional distribution does not meet the requirements for equal representation and does not take into account relative incidence.
- The Ústecký and Liberecký regions with a negative deviation of more than 50% from the required number of examinations are the permanently under-scanned regions, and the Zlín region may also fall among these regions.
- The regions of Středočeský and Moravskoslezský are the permanently under-served regions with a 25-50% negative deviation. The inclusion of the FoSCU in frame of HERA grant among sequencing laboratories improved sequencing in Prague in February 2022.
- The Vysočina region, is permanently sequenced with more than 50% positive deviation.
- The Pilsen region shows up to + 80% deviation.
- **Porovnání s požadavky na cílenou surveillance bude součástí podrobné zprávy dodané do konce května 2022 po zpracování dat z ÚZIS.**

- To ensure uniform sequencing, it would be appropriate to use an algorithm to define a significant sample, i.e. a sample determined for sequencing. **Porovnání s požadavky na cílenou surveillance bude součástí podrobné zprávy dodané do konce května 2022 po zpracování dat z ÚZIS.**
- **Task for the future:**
- Develop the pipeline that enable to streamline sample selection, facilitate logistics and realize the production of detailed sequencing data.
- The pipeline implementation promises a big shift in the field of sequencing when appropriate samples are selected according to actual sequencing capacity
- In summary, it can be said that the implementation of the pipeline will ensure better supervision of the epidemiological situation in the Czech Republic and provide the basis for more effective decision-making by the government based on data and records/evidence. It will provide the basis for further scientific work and enable the creation of balanced reports for the public.

Project „Enhancing Whole Genome Sequencing (WGS) and/or Reverse Transcription Polymerase Chain Reaction (RT-PCR) national infrastructures and capacities to respond to the Covid-19 pandemic in the European Union and European Economic Area“ had received funding from the European Centre for Disease Prevention and Control under the Grant Agreement number ECDC/HERA/2021/004 ECD.12218.



More information about the project: <http://www.szu.cz/ecdc-1>

NOTE:

Information included reflects only the authors' view and the Centre (ECDC) is not responsible for any use that may be made of the information it contains.