Report

Analyses for surveillance of SARS-CoV-2: variant monitoring and nowcasting of hospitalisations with additional ad-hoc analyses

April 1, 2022 - April 30, 2023

Auftraggeberin
Bundesamt für Gesundheit BAG / Federal Office of Public Health FOPH (Anna Fesser)

Auftragnehmerin
ETH Eidgenössische Technische Hochschule Zürich, Department of Biosystems Science and Engineering (Prof. Dr. Tanja Stadler, Computational Evolution Group)

Introduction

In the context of several research projects, the Computational Evolution group led by Prof. Dr. Tanja Stadler (D-BSSE, ETH Zurich) has developed methods for monitoring the COVID-19 pandemic. These include nowcasting of new hospital admissions and a genomic monitoring dashboard to track SARS-CoV-2 variants (https://cov-spectrum.org). As of April 2022, the SARS-CoV-2 genomic surveillance programme did not include funding for detailed genomic analyses. The aim of the contract was to optimally support COVID-19 surveillance by providing analyses using the methods developed by the Computational Evolution group and making them available to the Federal Office of Public Health (FOPH). For this purpose, FOPH temporarily provided the group with line list data including the necessary information for nowcasting of hospitalisations and for monitoring variants.

In the Computational Evolution group, the following staff were involved in the project:

- Stadler, Tanja (project lead)
- Chen, Chaoran (co-lead variant monitoring)
- Lison, Adrian (co-lead nowcasting)
- du Plessis, Louis (coordination)

Provisioning of data

FOPH provided the involved staff with de-identified line list data containing the following information per case:

- Eingangsdatum (eingang_dt)
- Fall Datum (fall_dt)
- Symptombeginn (manifestation_dt)
- Hospitalisierungsdatum (hospdatin)

- Datum Tod (pttoddat, pro Woche)
- Alter (0-9;10-19;20-29...;80+)
- Kanton (ktn)
- Impfstatus (none, 1 dosis, 2 doses, 3 doses, 4 doses, ….., geimpft ja/nein)
- Datum aller Impfungen
- Zeitraum letzte Impfdosis
- Meldedatum der Hospitalisierung
- Meldedatum des Todesfalls (pro Woche)
- GISAID ID der Sequenz falls Sequenz vorhanden
- Sex (f, m)

The data were transferred as asymmetrically encrypted RDS files via a shared cloud folder (polybox, hosted by ETH Zurich). They were only decrypted in-memory during processing and not stored as decrypted files. The private keys for description were only shared among the involved staff.

The first data were provided on February 21, 2023. They were then provided on a weekly basis for 9 consecutive weeks until April 25, 2023.

Variant Monitoring

The CoV-Spectrum dashboard was initially developed in 2021 to support the tracking of SARS-CoV-2 variants and identification of new variants of concern. Available at https://cov-spectrum.org (using data from GISAID) and https://open.cov-spectrum.org (using public domain data from INSDC), it is a global resource for scientists and public health agencies to analyse the national and international spread of SARS-CoV-2 variants. Throughout the period of the contract, CoV-Spectrum recorded over 1600 visitors per day on average and was cited in over 900 designation proposals for Pango lineages[1]. Its features include:

- Filtering by country, date range and sampling strategy
- Filtering by lineage and combinations of mutations
- Filtering by the sequence quality
- Relative and absolute number of sequences over time
- Estimated numer of cases over time
- Geographic distribution
- Relative growth advantage estimation
- Reproduction number estimation
- Mutation distribution
- Wastewater prevalence (in Switzerland only)

[1] https://github.com/cov-lineages/pango-designation/issues?q=cov-spectrum+OR+covspectrum+

After receiving the line list data from FOPH within this contracted project, an internal and password-protected instance was created and shared at https://swissinternal.cov-spectrum.org (Figures 1-4). It contained the following additional features:

- Age distribution
- Hospitalization probabilities

The swiss-internal instance was removed after the end of the contract.

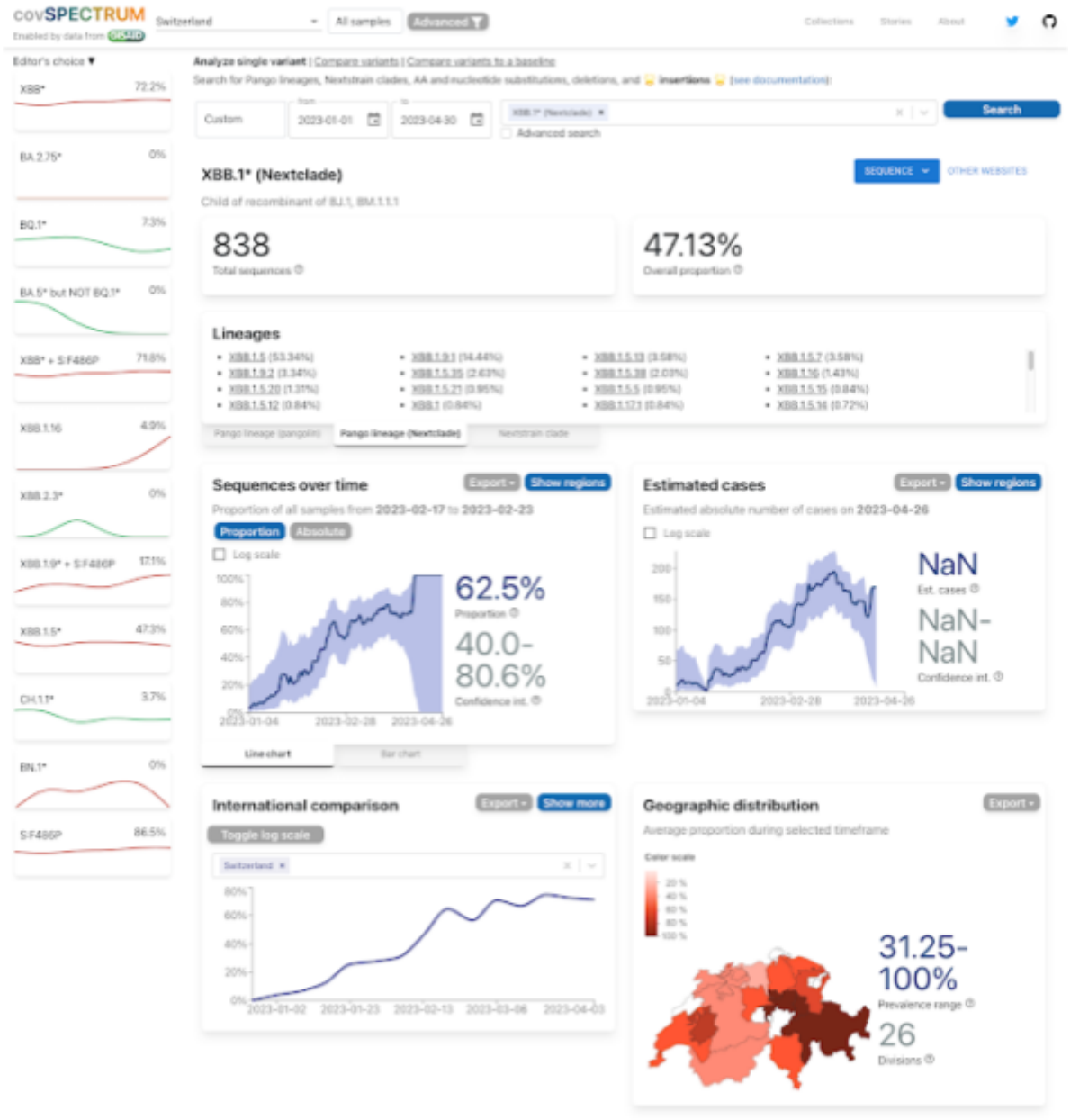

Figure 1. Screenshot from swissinternal.cov-spectrum.org (part 1)

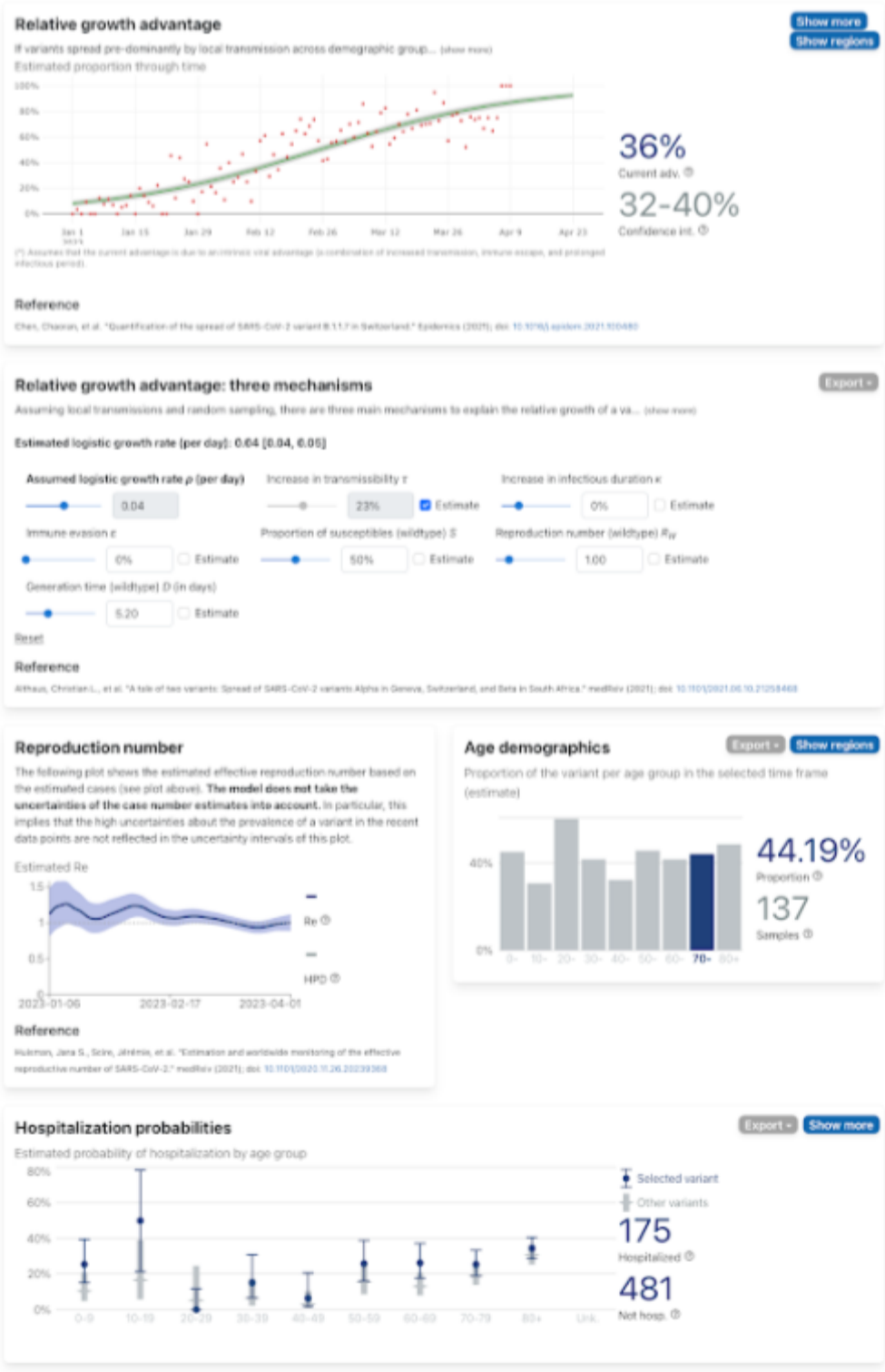

Figure 2. Screenshot from swissinternal.cov-spectrum.org (part 2)

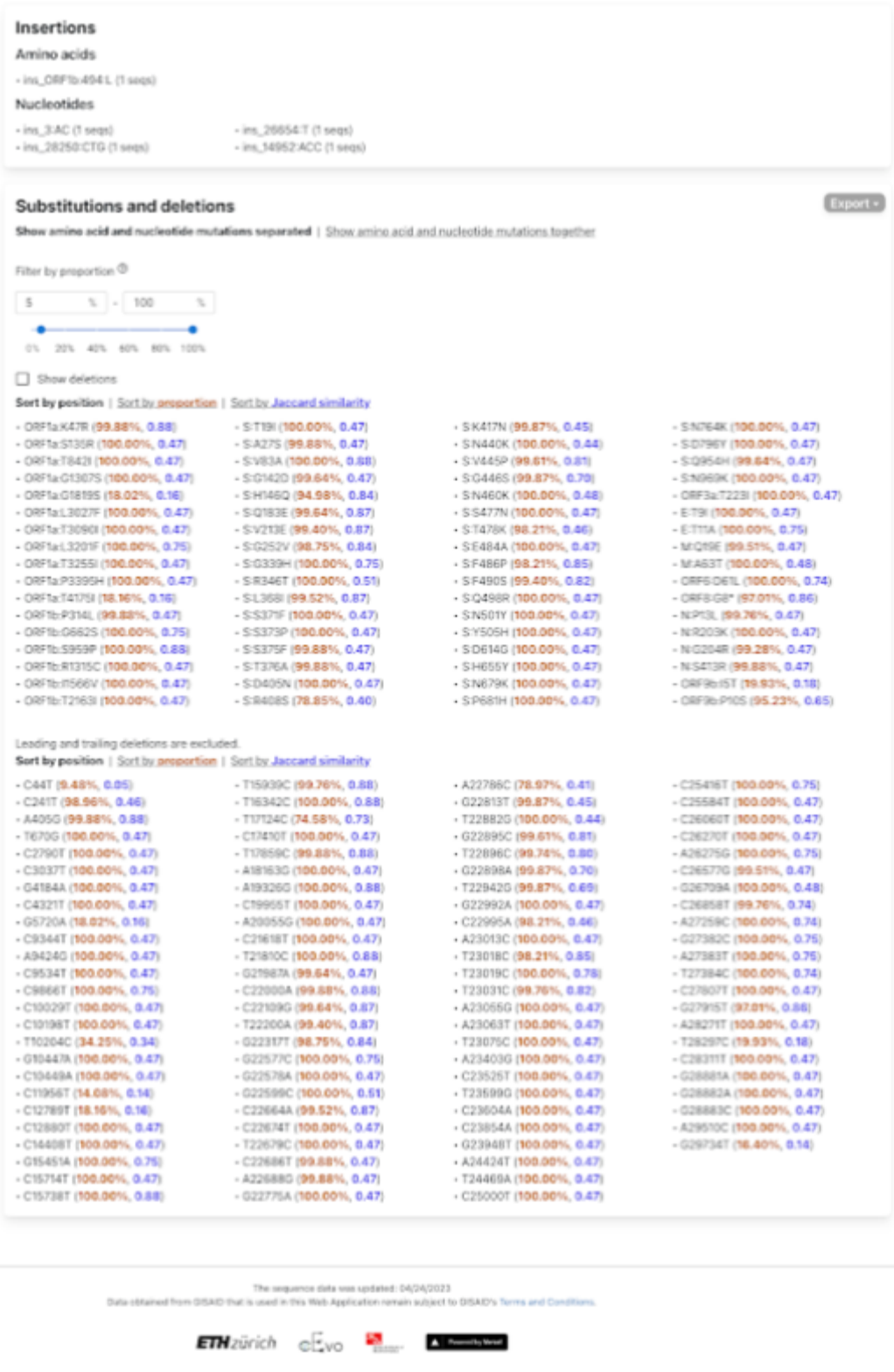

Figure 3. Screenshot from swissinternal.cov-spectrum.org (part 3)

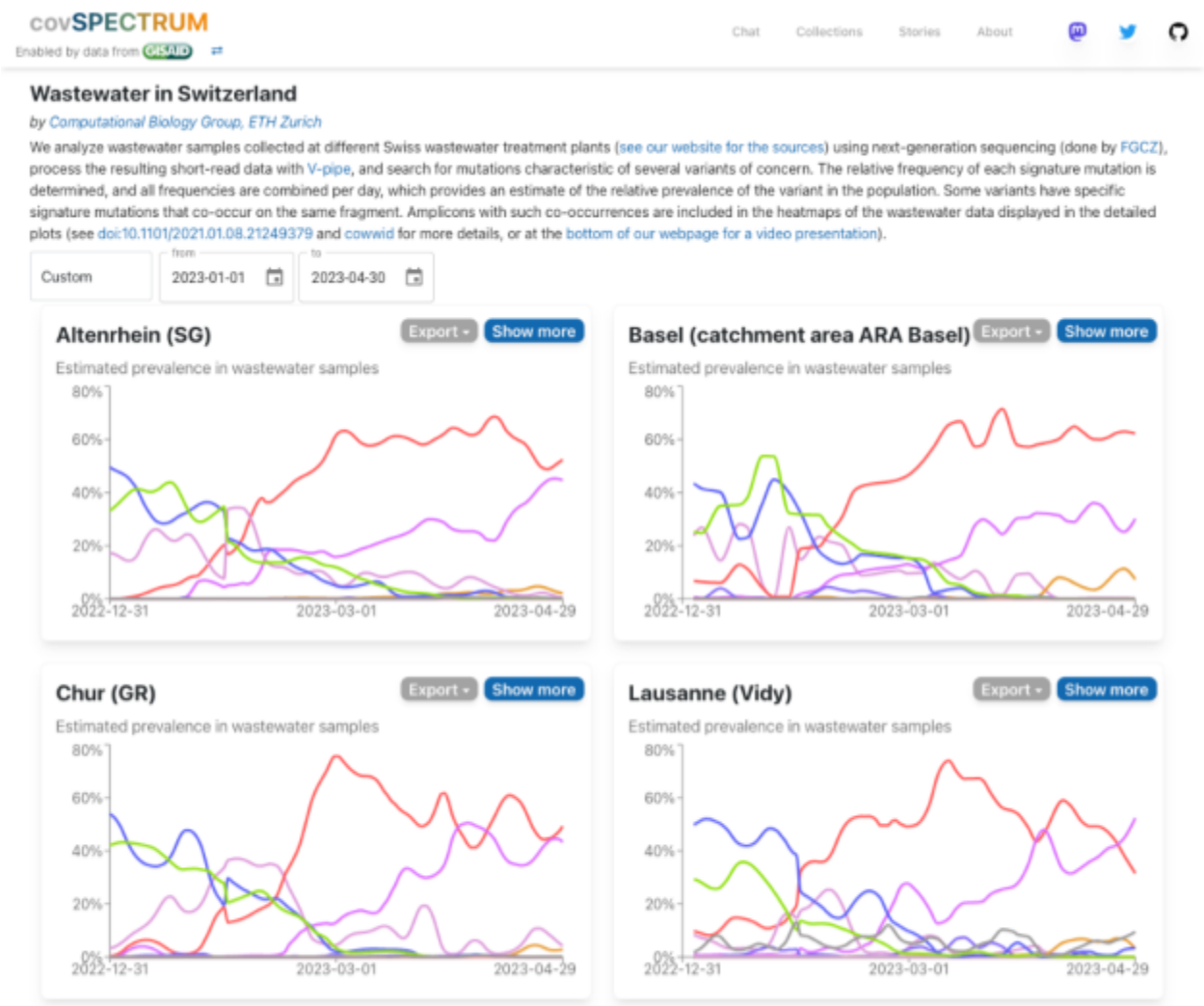

Figure 4. Screenshot from swissinternal.cov-spectrum.org (part 4)

Nowcasting

The objective of the nowcasting was to provide up-to-date statistics of new hospital admissions of patients who tested positive for COVID-19. Due to relevant delays between hospital admission and the report of hospitalisation to FOPH, the count time series of hospital admissions by date of hospitalisation is downward-biased: it is missing hospitalisations that have already occurred but are not yet reported. Using a statistical nowcasting model, this downward bias should be corrected by estimating the number of occurred-but-not-yet-reported hospital admissions. Moreover, the date of hospital admission can be missing for some cases. The nowcasting model developed in the Computational Evolution group is able to also account for these missing hospitalisation dates.
The statistical model that was employed for nowcasting used hierarchical Bayesian modeling to estimate the quantities of interest while quantifying the uncertainty of the estimates. The delay between hospitalisation and report was jointly estimated from the line list data, using the last three months of data. Thereby, the model accounted for time-varying reporting delays and differences in reporting between weekdays. Missing dates of hospitalisation were jointly inferred (under a missing-at-random assumption, i.e. assuming that incomplete cases have reporting delays comparable to complete cases). Uncertainty resulting from noise in the reporting, uncertain reporting delays, and the missing data was accounted for. Two different versions of the nowcasting model were applied to the line list data. The first version was a non-parametric model, which used an exponential smoothing prior (with log-linear trend) on the hospitalisation time series. The second version was a semi-mechanistic transmission model, i.e. with an explicit model of the infection process leading to hospitalisations. It was smoothed using an exponential smoothing prior (with log-linear trend) on the effective reproduction number. In addition to the two model versions, a so-called “ensemble” nowcast was calculated. The ensemble nowcast is simply the arithmetic mean of the nowcast from the non-parametric and from the semi-mechanistic model at each date.

During preprocessing, entries with negative delays (date of hospitalization later than the date of report of the hospitalization) were handled by treating the date of hospitalization as erroneous and setting it to missing/unknown. Moreover, cases with a reporting delay longer than a maximum of 2 months were discarded from the analysis, which applied to strictly less than 5% of cases. This was also done to exclude potentially erroneous, very large delays observed in the data (up to a year or longer). For the semi-mechanistic transmission model, we used information on the symptom onset dates of patients to model the delay between infection and symptom onset. We further assumed an incubation period of the Omicron variant, modeled as gamma distributed with shape 8.5 and scale 0.41 (Manica et al.), and a generation interval of the Omicron variant, modeled as log-normally distributed with a log mean of 0.98 and a log standard deviation of 0.47 (Park et al.).

Preprocessing of the data was done using R version 4.1.0. The nowcasting models were specified as probabilistic programs in Stan and fitted using the package *cmdstanr*. The nowcasting results for each week were provided to FOPH in tabular format, showing for each model version and each date of hospitalisation the number of already reported, the estimated

number of not-yet-reported, and the resulting estimated total number of cases. For the not-yet-reported cases, the median and the lower and upper bound of the 95% uncertainty interval were reported. In addition to the tabular results, a weekly HTML report was produced, showing an interactive plot of the nowcasts from the different models (see Figure 5).

Nowcasts were produced for each of the 10 weeks in which data was provided (February 21 - April 25, 2023). During this time period, hospitalisation counts were consistently low (below 35 cases per day) in Switzerland, such that nowcasts could only be computed at the national level without further stratification by canton or age group. In retrospective, the nonparametric and the semi-mechanistic models showed very high agreement for all nowcasts. Moreover, the 95% uncertainty intervals covered to true hospitalisation count very reliably. In part, these results were to be expected, as the period covered by the nowcasts showed very stable hospitalisation dynamics without relevant changes in trends. At the same time, because of the low case numbers, the nowcast uncertainty was, in relative terms, rather large compared to the overall count.

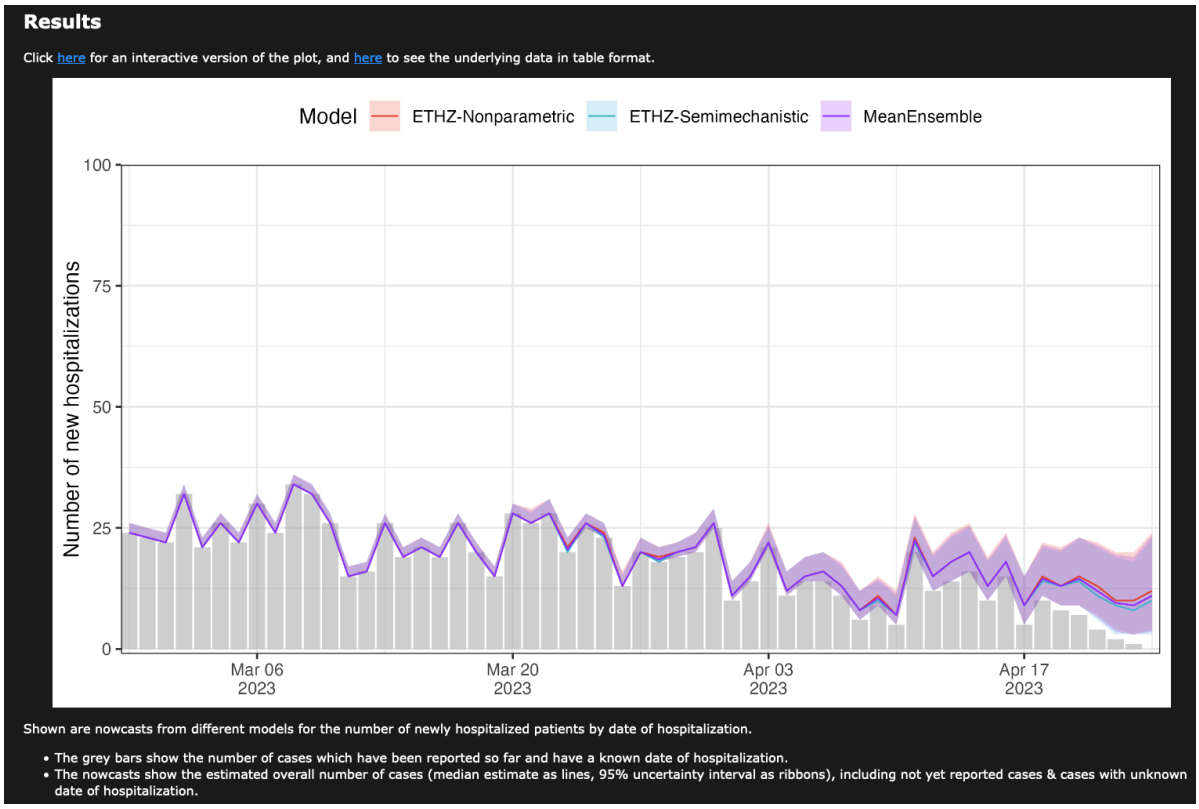

Figure 5. Screenshot showing an excerpt of the weekly HTML report for the hospitalisation nowcast.

Additional analyses

No additional ad-hoc analyses were requested.

Conclusion

Within the present service contract, the Computational Evolution group provided results of epidemiological analyses to FOPH with the goal of supporting COVID-19 surveillance. In particular, nowcasting of hospitalisations and monitoring of variants was performed. This was made possible through the regular provision of line list data by FOPH.

The Computational Evolution group provides a publicly accessible variant dashboard CoV-Spectrum since 2021. As part of this service contract, an additional internal instance, which facilitated the analysis of the severity of different variants, was provided. The public dashboard is widely used and remains an essential tool for continued surveillance. In the event of a new Variant of Concern (VOC) emerging, the public dashboard stands ready to monitor its spread. Based on the work within this service contract, there is the infrastructure to again quickly establish a local instance for FOPH to examine the VOC's severity pattern and age distribution.

The nowcasting models employed within this service contract to estimate the number of new COVID-19 hospitalizations produced reliable nowcasts that accounted for estimation uncertainty. The results were displayed in an easily accessible HTML report, equipped with interactive plots. During the analyzed period, nowcasts were generally accurate, but the low case counts observed during this period prevented further disaggregation by canton or age group. Should the pandemic situation become increasingly challenging for hospitals in the future, the nowcasting models and data processing pipelines developed can be readily applied to new FOPH data upon request.

The nowcasts conducted under the present contract used the non-public line list data provided by FOPH. Essentially, however, the only requirement for nowcasting is a two-dimensional count of hospitalized cases, categorized by the date of hospitalisation and the date of the report. If such data were made publicly accessible, for instance through the website www.covid19.admin.ch, it would permit various research groups to produce nowcasts without further service contracts. These could then be collated on a public dashboard, similar to the approach taken in Germany, as demonstrated on https://covid19nowcasthub.de.

In essence, the service contract between the Computational Evolution group and FOPH enabled the exchange of critical data for real-time analysis in a safe way and established data pipelines and reporting formats that could be deployed again in the future if required by the pandemic situation.