medRxiv preprint doi: https://doi.org/10.1101/2021.03.05.21252520; this version posted March 9, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

Quantification of the spread of SARS-CoV-2 variant B.1.1.7 in Switzerland

Chaoran Chen[1,2], Sarah Nadeau[1,2,*], Ivan Topolsky[1,2,*], Marc Manceau[1,2,*], Jana S. Huisman[1,2,3*], Kim Philipp Jablonski[1,2], Lara Fuhrmann[1,2], David Dreifuss[1,2], Katharina Jahn[1,2], Christiane Beckmann[4], Maurice Redondo[4], Olivier Kobel[4], Christoph Noppen[4], Lorenz Risch[5], Martin Risch[5], Nadia Wohlwend[5], Sinem Kas[5], Thomas Bodmer[5] Tim Roloff[2,6,7], Madlen Stange[2,6,7],Adrian Egli[6,7], Isabella Eckerle[8] Rebecca Denes[9], Mirjam Feldkamp[9], Ina Nissen[9], Natascha Santacroce[9], Elodie Burcklen[9], Catharine Aquino[10], Andreia Cabral de Gouvea[10], Maria Domenica Moccia[10], Simon Grüter[10], Timothy Sykes[10], Lennart Opitz[10], Griffin White[10],Laura Neff[10], Doris Popovic[10], Andrea Patrignani[10], Jay Tracy[10], Ralph Schlapbach[10], Emmanouil T. Dermitzakis[11,12], Keith Harshman[11], Ioannis Xenarios[11], Henri Pegeot[11], Lorenzo Cerutti[11], Deborah Penet[11], Anthony Blin[11], Melyssa Elies[11], Christian Althaus[13], Christian Beisel[9,&], Niko Beerenwinkel[1,2,&], Martin Ackermann[3,14,&], Tanja Stadler[1,2,+]

[1] Department of Biosystems Science and Engineering, ETH Zürich, Basel, Switzerland

[2] Swiss Institute of Bioinformatics, Switzerland

[3] Department of Environmental Systems Science, ETH Zürich, Swiss Federal Institute of Technology, Zurich, Switzerland

[4] Viollier AG, Allschwil, Switzerland

[5] Dr Risch, Labormedizinisches Zentrum, Switzerland

[6] Clinical Bacteriology and Mycology, University Hospital Basel, Basel, Switzerland

[7] Applied Microbiology Research, Department of Biomedicine, University of Basel, Basel, Switzerland

[8] Geneva Centre for Emerging Viral Diseases, Infectious Disease Division, Geneva University Hospitals & Department of Microbiology and Molecular Medicine, University of Geneva

[9] Genomic Facility Basel, Department of Biosystems Science and Engineering, ETH Zürich, Basel, Switzerland

[10] Functional Genomics Center Zurich, ETH Zürich and University of Zurich, Zurich, Switzerland

[11] Health 2030 Genome Center, Geneva, Switzerland

[12] University of Geneva Medical School, Geneva, Switzerland

[13] Institute of Social and Preventive Medicine, University of Bern, Bern, Switzerland

[14] Department of Environmental Microbiology, Eawag, Dubendorf, Switzerland

[*,&] These authors contributed equally to this work

[+] Correpondence: tanja.stadler@bsse.ethz.ch

NOTE: This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.

February 26, 2021

medRxiv preprint doi: https://doi.org/10.1101/2021.03.05.21252520; this version posted March 9, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

Keywords: Pandemic, SARS-CoV-2, COVID-19, B.1.1.7, transmission advantage

Abstract

In December 2020, the United Kingdom (UK) reported a SARS-CoV-2 Variant of Concern (VoC) which is now coined B.1.1.7. Based on the UK data and later additional data from other countries, a transmission advantage of around 40-80% was estimated for this variant [1, 2, 3].

In Switzerland, since spring 2020, we perform whole genome sequencing of SARS-CoV-2 samples obtained from a large diagnostic lab (Viollier AG) on a weekly basis for genomic surveillance. The lab processes SARS-CoV-2 samples from across Switzerland. Based on a total of 7631 sequences obtained from samples collected between 14.12.2020 and 11.02.2021 at Viollier AG, we determine the relative proportion of the B.1.1.7 variant on a daily basis. In addition, we use data from a second lab (Dr Risch) screening all their samples for the B.1.1.7 variant. These two datasets represent 11.5 % of all SARS-CoV-2 confirmed cases across Switzerland during the considered time period. They allow us to quantify the transmission advantage of the B.1.1.7 variant on a national and a regional scale.

Taking all our data and estimates together, we propose a transmission advantage of 49-65% of B.1.1.7 compared to the other circulating variants. Further, we estimate the effective reproductive number through time for B.1.1.7 and the other variants, again pointing to a higher transmission rate of B.1.1.7. In particular, for the time period 01.01.2021-17.01.2021, we estimate an average reproductive number for B.1.1.7 of 1.28 [1.07-1.49] while the estimate for the other variants is 0.83 [0.63-1.03], based on the total number of confirmed cases and our Viollier sequencing data. Switzerland tightened measures on 18.01.2021. A comparison of the empirical confirmed case numbers up to 20.02.2021 to a very simple model using the estimates of the reproductive number from the first half of January provides indication that the rate of spread of all variants slowed down recently.

In summary, the dynamics of increase in frequency of B.1.1.7 is as expected based on the observations in the UK. Our plots are available online and constantly updated with new data to closely monitor the changes in absolute numbers.

medRxiv preprint doi: https://doi.org/10.1101/2021.03.05.21252520; this version posted March 9, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

1 Introduction

Reports of an increased transmissibility of SARS-CoV-2 variant B.1.1.7 (501Y.V1) were released in mid-December 2020 [4, 5, 6]. This variant carries the N501Y mutation in the spike protein which may increase the ACE2 receptor affinity [7]. Within only a few months, the variant was able to become the dominant lineage in the UK.

Since these first reports, great efforts were made in Switzerland to detect and trace B.1.1.7 [8]. The first cases of B.1.1.7 were confirmed on 24.12.2020 and retrospective analyses identified B.1.1.7 in samples dating back to October [8]. In total, 1370 infections with B.1.1.7 were confirmed up to 05.02.2021 [8].

An increase in prevalence does not necessarily imply a transmission fitness advantage of the virus. For example, a variant termed 20A.EU1 spread rapidly during summer 2020 across Europe. However, data suggests that extended travel and superspreading events rather than a viral transmission advantage caused that spread [9].

However, the variant B.1.1.7 rapidly increased in frequency in many high-prevalence regions across the UK. Such re-occuring patterns are hard to explain without a transmission advantage. Analyses of the B.1.1.7 variant in the UK suggest that the variant has a transmission fitness advantage anywhere between 40 and 80% [1, 2, 3]. Davies et al [3] further uses data from Denmark obtaining again similar estimates. Quantitative analysis of data on the spread of the 501Y mutation in Switzerland also suggests a transmission advantage in roughly that range [10].

In this paper, we collected whole genome sequencing data based on samples from Viollier AG to determine the proportion of B.1.1.7 through time. Further we used data from the diagnostic laboratory Dr Risch which screens all their samples for the B.1.1.7 variant. We then used this data to quantify the transmission fitness advantage of B.1.1.7 for Switzerland as well as for the seven Swiss economic regions (Grossregionen). We further calculated the reproductive number for B.1.1.7 and show how the B.1.1.7 case numbers developed through time.

The core plots presented here will be updated regularly on [11] and part of it is currently incorporated into the Swiss National COVID-19 Science Task Force website [12]. The code and data used for this paper is publically available on Github [13]. All sequences based on samples fro Viollier are available on GISAID.

medRxiv preprint doi: https://doi.org/10.1101/2021.03.05.21252520; this version posted March 9, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

2 Method

2.1 Data

The main source of data are whole genome sequences based on samples from Viollier AG. Viollier AG is a large diagnostic lab processing SARS-CoV-2 samples from across Switzerland. Every week, SARS-CoV-2 samples were randomly selected for sequencing among all positively tested samples at the Viollier AG. For each sample, we know the date of the test and the canton in which the test was performed.

We perform whole genome sequencing in three facilities. The sequencing protocol for the samples sequenced at the Genomics Facility Basel and the Functional Genomics Center Zurich is described in [14]. Samples processed at the Health 2030 Genome Center used the Illumina COVIDSeq library preparation reagents following the protocol provided by the supplier [15]. These reagents are based on the ARTIC v3 multiplex PCR amplicon protocol [16]. When sufficient volume was available, 8.5ul of RNA extracted from patient nasopharyngeal samples were used in the cDNA synthesis step; if 8.5ul were not available, the maximum volume possible was used. Pooled libraries were sequenced on the Illumina NovaSeq 6000 using a 50-nucleotide pair-end run configuration. Post-sequencing library read de-multiplexing was done using an in-house developed processing pipeline [17].

The downstream bioinformatics procedure to obtain consensus sequences is described in [14, 18]. In the time frame of interest – from 14.12.2020 to 11.02.2021, a total of 7631 sequences were generated and checked for the B.1.1.7 variant. Thus, per week we produced around 900 sequences. In the same time period, 154'241 people were confirmed with a SARS-CoV-2 infection, thus we provide sequences for 4.9% of all cases. We define a sequence to be a B.1.1.7 sample if at least 80% of the lineage-defining, non-synonymous nucleotide changes according to [4] are present. The dataset used downstream is the number of sequences per day and the number of identified B.1.1.7 variants per day.

Additionally, we have data from the Dr Risch medical laboratories. Dr Risch medical laboratories used the Taqpath assay from Thermofisher for their diagnostic and recorded the S gene target failure (SGTF). SGTF samples are potential B.1.1.7 variants, as the B.1.1.7 variant causes a SGTF due to its deletion at position 69-70 in the spike protein. Further, Dr Risch medical laboratories screen their samples for the 501Y mutation by a variant-specific PCR test. If a sample is identified as a potential VoC by these procedures, it is sent for whole genome sequencing to the University Hospital

medRxiv preprint doi: https://doi.org/10.1101/2021.03.05.21252520; this version posted March 9, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

Grossregion	Total confirmed cases	Sequenced	Proportion
Central Switzerland	15344	454	2.96%
Espace Mittelland	29901	2292	7.67%
Nordwestschweiz	20795	2063	9.92%
Tessin	8094	74	0.91%
Zurich	29110	1502	5.16%
Lake Geneva	27651	539	1.95%
Ostschweiz	23345	707	3.03%
Total	154240	7631	4.95%

Table 1: The proportion of sequenced cases out of all cases for the Viollier dataset.

Basel in order to confirm the B.1.1.7 variant. The sequencing protocol is described in [19]. For the recent samples, the confirmation may still be outstanding. However, since typically a SGTF plus a 501Y mutation corresponds indeed to a B.1.1.7 variant, we consider these samples as B.1.1.7 variants already prior to the results of the whole genome sequencing confirmation.

When the initial screening for B.1.1.7 in Switzerland was started just before Christmas 2020, the first approach to identify the variants was by specific amplification of the indicative Spike region followed by in house Sanger sequencing at the national reference center for emerging viral infections (CRIVE) in Geneva.

Overall, in our considered time period, 10175 samples were screened from Dr Risch. In the final dataset, we record the daily number of positive tests at Dr Risch and the number of B.1.1.7 variants identified.

Taking both datasets together, 11.5% of all SARS-CoV-2 cases during our considered time period (14.12.2020 to 11.02.2021) were screened for the B.1.1.7 variant. While Viollier processes samples from all over Switzerland, the intensity varies across regions. The set of sequenced samples inherits this uneven geographical distribution with the result that, relatively to the number of all cases, over ten times more samples were sequenced from the region Nordwestschweiz than from the region Ticino (Table 1). Dr Risch's data has a different geographic distribution than Viollier and, for example, a much better coverage of Eastern Switzerland. In summary, both datasets differ in their geographic biases. We present results on the different datasets and compare these results.

medRxiv preprint doi: https://doi.org/10.1101/2021.03.05.21252520; this version posted March 9, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

Grossregion	Total confirmed cases	Sequenced	Proportion
Central Switzerland	15344	279	1.82%
Espace Mittelland	29901	3238	10.83%
Nordwestschweiz	20795	979	4.71%
Tessin	8094	111	1.37%
Zurich	29110	1013	3.48%
Lake Geneva	27651	326	1.18%
Ostschweiz	23345	4229	18.12%
Total	154240	10175	6.60%

Table 2: The proportion of characterized cases out of all cases for the Risch dataset.

2.2 Estimating the transmission fitness advantage of a new variant

In what follows, we define two models describing the dynamics with which a new variant with a transmission fitness advantage spreads in a population. The first model is based on the assumption of discrete time, while the second model is based on the assumption of continuous time. Both models have been considered extensively in the literature to estimate fitness advantage (see e.g. [20, 3]). While in an epidemic, a generation does not end after a fixed time span (discrete time model), generations are typically less variable than modelled under an exponential distribution (continuous time model). Thus we view the two models as two extremes with the actual dynamics being described by a process in between. We provide estimates based on both models and suggest that the true parameter may be anywhere within the ranges spanned by the two models. In the next sections, we provide details of how we estimate the transmission fitness advantage of B.1.1.7 based on daily data of the total number of samples and B.1.1.7 samples under these two models.

2.2.1 Discrete time model

We call X the common (non-B.1.1.7) variants and Y the B.1.1.7 variant. The process starts in generation 0 with x_0 cases caused by variant X and y_0 cases caused by variant Y. Let the number of cases in generation n be $x(n)$ and $y(n)$ for variants X and Y.

Let the reproductive number R_d of variant X in generation n be $R_d(n)$. Let the transmission advantage of variant Y be f_d. Then the reproductive number of Y in generation n is $(1 + f_d)R_d(n)$. Thus, we assume a multiplicative fitness advantage.

We have $x(n) = x(0) \times R_d(0)R_d(1) \dots R_d(n-1)$ and $y(n) = y(0) \times R_d(0)R_d(1) \dots R_d(n-1)(1+f_d)^n$. If R_d is constant through time, we have

medRxiv preprint doi: https://doi.org/10.1101/2021.03.05.21252520; this version posted March 9, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

$$x(n) = x(0)R_d^n \text{ and } y(n) = y(0)((1+f_d)R_d)^n.$$

Let the proportion of variant Y at generation n be $p(n)$. We have,

$$\begin{aligned} p(n) &= \frac{y(n)}{x(n)+y(n)} = \frac{y(0)(1+f)^n}{x(0)+y(0)(1+f)^n} \\ &= \frac{p(0)(1+f_d)^n}{1-p(0)+p(0)(1+f_d)^n} \\ &= \frac{1}{1+(1+f_d)^{-n}\left(p(0)^{-1}-1\right)}. \end{aligned}$$

Thus, $p(n)$ is the logistic function. It does not depend on R_d.

If we write time in days t rather than generations n and assume a generation time of g days, we get

$$p(t) = \frac{1}{1+(1+f_d)^{-t/g}\left(p(0)^{-1}-1\right)}. \tag{1}$$

We now switch our parameterization to the more common

$$p(t) = \frac{1}{1+e^{-a(t-t_0)}} \tag{2}$$

for parameter estimation from daily data. Parameter a is the logistic growth rate and parameter t_0 the sigmoid's midpoint.

The two free parameters, a and t_0, are related to the two free parameters in Equation 1, f_d and $p(0)$, as follows:

$$a = \frac{\ln(1+f_d)}{g}, \quad t_0 = g\,\frac{\ln(p(0)^{-1}-1)}{\ln(1+f_d)}.$$

In particular, we get $f_d = e^{ag} - 1$.

medRxiv preprint doi: https://doi.org/10.1101/2021.03.05.21252520; this version posted March 9, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

2.2.2 Continuous time model

In continuous time, instead of R_d and generation time g, we define the transmission rate β and the recovery rate μ. Under this model, the reproductive number is $R_c = \beta/\mu$. Further, since an individual in the discrete model recovers after a generation of duration g (during which they left R_d offspring), we note that g is related to the expected time to recovery $1/\mu$ in the continuous model, and in fact assume $g = 1/\mu$ in what follows. Again our initial numbers of the variants X and Y are $x(0)$ and $y(0)$. Calendar time is denoted by a continuous parameter t. We then have in expectation,

$$x(t) = x(0)e^{(\beta-\mu)t}. \tag{3}$$

We note that $\beta - \mu$ is coined the Malthusian growth parameter [3].

Further we again assume that variant Y has a transmission fitness advantage of f_c, with transmission rate $\beta(1 + f_c)$ and recovery rate μ. The population size of the variant at time t is thus

$$y(t) = y(0)e^{(\beta(1+f_c)-\mu)t}. \tag{4}$$

The proportion of the variant in the population at time t is

$$\begin{aligned} p(t) &= \frac{y(0)e^{(\beta(1+f_c)-\mu)t}}{x(0)e^{(\beta-\mu)t} + y(0)e^{(\beta(1+f_c)-\mu)t}} \\ &= \frac{p(0)e^{\beta f_c t}}{1 - p(0) + p(0)e^{\beta f_c t}} \\ &= \frac{1}{1 + e^{-\beta f_c t}(p(0)^{-1} - 1)} \end{aligned}$$

where we again recognize the logistic function. We turn again to the more common parameterization,

$$p(t) = \frac{1}{1 + e^{-a(t-t_0)}} \tag{5}$$

where we thus have $f_c = \frac{a}{\beta}$.

The reproductive number is $R_c = \beta/\mu$ and the mean time to recovery, $1/\mu$, is equaled to g. Then, $\beta = R_c/g$. Thus,

medRxiv preprint doi: https://doi.org/10.1101/2021.03.05.21252520; this version posted March 9, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

$$a = \frac{f_c R_c}{g}, \quad t_0 = g\frac{\ln(p(0)^{-1} - 1)}{f_c R_c}.$$

In particular, we have $f_c = \frac{ag}{R_c}$. Note that the estimated fitness advantage under this model depends on the reproductive number R_c and is thus changing if R_c is changing through time.

2.2.3 Connection between discrete and continuous time

The discrete and continuous models are very similar. Both have the intitial conditions $x(0)$ and $y(0)$. For the dynamics, the discrete model has parameters R_d and g while the continuous model has parameters β and μ. We have $R_c = \beta/\mu$ and we further assumed that $g = 1/\mu$ ($1/\mu$ is the expected time until recovery in the continuous setting while g is the time to recovery in the discrete setting). The different parameterizations of fitness advantage are coined f_c and f_d. We now determine how f_c and f_d are related.

To compare the two models, we now assume that their overall dynamics for the variant X are the same. After n generations of duration g, we have for variant X,

$$\begin{aligned} & x(0)e^{(\beta-\mu)ng} = x(0)R_d^n \\ \iff & e^{(\beta-\mu)g} = R_d \\ \iff & R_d = e^{R_c - 1} \end{aligned}$$

Using a Taylor expansion for β/μ close to 1, we obtain that indeed $R_d = R_c$.

medRxiv preprint doi: https://doi.org/10.1101/2021.03.05.21252520; this version posted March 9, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

For the two models to produce the same growth also for variant Y, we require,

$$\begin{aligned}
& y_0 e^{(\beta(1+f_c)-\mu)ng} = y_0((1+f_d)R_d)^n \\
\iff & (\beta(1+f_c)-\mu)g = \ln((1+f_d)R_d) \\
\iff & (\beta-\mu)g + f_c\beta g = \ln(R_d) + \ln(1+f_d) \\
\iff & f_c\beta g = \ln(1+f_d)) \\
\iff & f_c = \frac{\ln(1+f_d)}{\beta g} \\
\iff & f_c = \frac{\ln(1+f_d)}{R_c}.
\end{aligned}$$

In the last step, we make use of $R_c = \beta/\mu$ and $g = 1/\mu$.

This is equivalent to $f_d = e^{R_c f_c} - 1$. Using a Taylor expansion we get $f_d = 1 + R_c f_c + O((R_c f_c)^2) - 1$ and thus $f_d = R_c f_c$ for small $R_c f_c$.

2.2.4 Maximum likelihood parameter estimation

Next we explain how we estimate a and t_0 of the logistic functions (Eqn. 2 and 5) from our data using maximum likelihood. We consider that we have data at times $t_1, ..., t_d$. At time t_i, we obtained n_i samples, where n_i is fixed, non-random.

We assume that the true number of B.1.1.7 variants at time t_i is a random variable, K_i which is binomially distributed with parameter $p(t_i)$, i.e.

$$K_i \sim \mathcal{B}\left(n_i, p(t_i)\right), \qquad \text{where } p(t_i) = \frac{e^{a(t_i - t_0)}}{1 + e^{a(t_i - t_0)}} \quad .$$

In particular we assume here a deterministic logistic growth model for the increase in the proportion of variant Y (Eqn. 2 and 5), on top of which only the drawing process is random. This model simplifies naturally to a very popular logistic regression. This is an instance of a Generalized Linear Model, where the natural parameter of the binomial distribution is a linear function of predictors, the only predictor considered here being the time t.

We use the *Python* library *statsmodel* [21] to recover maximum likelihood estimates (MLEs) and confidence intervals based on an asymptotic Gaussian distribution for the parameters of the logistic regression based on our data, i.e. the fixed values $t_1, \ldots, t_d$, $n_1, \ldots, n_d$ as well as the numbers of

medRxiv preprint doi: https://doi.org/10.1101/2021.03.05.21252520; this version posted March 9, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

samples at each time point being the variant B.1.1.7, $k_1, \ldots, k_d$. Parameters a, t_0, f_d, f_c as well as the proportions of variant B.1.1.7 $p(t)$ through time are simple transformations of the parameters of the logistic regression. Their MLEs are the same transformations applied to the MLEs of the logistic regression parameters. The difference between the MLEs and the true parameters are again Gaussian, with a covariance matrix found by applying the delta method. This is used to construct confidence intervals for all these quantities.

2.3 Estimation of the effective reproductive number

We use the number of confirmed cases per day from the Federal Office of Public Health, Switzerland, for 14.12.2020 to 11.02.2021. Then, for each day, we estimate the number of B.1.1.7 variants by multiplying the total number of confirmed cases by the proportion of B.1.1.7 in our dataset (Viollier or Risch). We then estimate an effective reproductive number of the B.1.1.7 variant and of the non-B.1.1.7 variants using these data. For this estimation, we use the method developed in [22]. This method consists of two main parts: first, the observed case data is related to the corresponding time series of infections. We smooth the observations using LOESS smoothing to remove weekend effects. Then, we deconvolve with the delay of infection to symptom onset (gamma distributed with mean 5.3 and sd 3.2) and the delay from symptom onset to case confirmation (gamma distributed with mean 5.5 and sd 3.8). Second, we estimate the effective reproductive number from the time series of infection incidence using EpiEstim [23]. The reported point estimate is the estimate on the original case data. To account for uncertainty in the observation process, the observed daily case incidences are additionally bootstrapped 1000 times, resulting in an ensemble of alternative case incidence time series and corresponding estimated effective reproductive numbers. These are used to construct the 95% confidence interval around the effective reproductive number, and to calculate the standard deviation of the ratios of effective reproductive number estimates (see below).

We perform the estimation of the reproductive number in two different ways. First, we estimate smooth changes in the reproductive number, by estimating it across the entire time series using a 3-day sliding window. Second, we assume the reproductive number was constant during time intervals in which the non-pharmaceutical interventions did not change. Since 18.01.2021, Switzerland has implemented a set of tighter measures (in particular, shops are closed and the size of gatherings is restricted to five people [24]). Thus we fix the reproductive number to be constant between 01.01.2021

medRxiv preprint doi: https://doi.org/10.1101/2021.03.05.21252520; this version posted March 9, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

and 17.01.2021. Then the reproductive number is allowed to change and again fixed to be constant from 18.01.2021 onwards.

To compare the effective reproductive number R of the B.1.1.7 variant (Y) to that of non-B.1.1.7 variants (X), we take the ratio $\rho = \frac{R_Y}{R_X}$ at every time point. The standard deviation of this ratio σ_ρ was found through Gaussian error propagation of the standard deviation of the individual R estimates (σ_X, σ_Y):

$$\sigma_\rho = \sqrt{\frac{1}{(R_X)^2}\sigma_Y^2 + \frac{(R_Y)^2}{(R_X)^4}\sigma_X^2}.$$

2.4 Projections under a simple epidemiological model

We finally display the expected number of confirmed cases in the future under the continuous model; thus case numbers of variant X and Y are changing according to Equations 3 and 4. We initialize the model ($x(0)$ and $y(0)$) on 01.01.2021 with the estimated number of B.1.1.7 and non-B.1.1.7 cases on that day. We assume a reproductive number $R = \beta/\mu$ for the non-B.1.1.7 as estimated on the national level for 01.01.2021-17.01.2021. Finally, we assume that the expected generation time $1/\mu$ is 4.8 days and the fitness advantage is the estimated f_c for the region and dataset of interest (Table 3).

3 Results

We estimate the logistic growth rate a and the sigmoid's midpoint t_0 based on the Viollier and Risch data (Table 3). Taking the estimates of both datasets together, we obtain a growth rate a of around 0.09-0.10 per day for Switzerland. For each economic region, the estimated uncertainty interval of a overlaps with the Swiss-wide uncertainty interval. We have little data for two out of seven regions (Ticino and Central Switzerland; <550 sequences in total) resulting in very wide uncertainty intervals. From the t_0 estimates, we observe that the Geneva region is about 1-2 weeks ahead of the rest of Switzerland with respect to the spread of B.1.1.7, which confirms estimates from [10]. We estimate that B.1.1.7 will become dominant in Switzerland in March 2021.

In Fig. 1 and 2, we graphically illustrate the logistic growth in frequency of B.1.1.7 and show the daily data together with an estimate of the proportion of B.1.1.7 under the logistic growth model

medRxiv preprint doi: https://doi.org/10.1101/2021.03.05.21252520; this version posted March 9, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

(parameter $p(t_i)$ on each day t_i).

Next, we estimate the reproductive number for B.1.1.7 and non-B.1.1.7 on a national scale (Fig. 3). During 01.01.2021-17.01.2021, the reproductive number for B.1.1.7 was significantly above 1 (Viollier: 1.28 [1.07-1.49], Risch: 1.46 [1.23,1.68]) while the reproductive number of non-B.1.1.7 tended to be below 1 (Viollier: 0.83 [0.63-1.03] , Risch: 0.82 [0.61,1.03]).

Based on the Viollier data, the reproductive number did not change much after 18.01.2021, however the confidence intervals are wide. As expected, the ratio of the reproductive numbers for B.1.1.7 and non-B.1.1.7 stays roughly constant throughout January. Our data does not allow us to estimate the reproductive number for February yet.

For the Risch data, the reproductive number of non-B.1.1.7 remained roughly constant throughout January. Based on our discrete and continuous model described above, we would then expect also a roughly constant reproduction number for B.1.1.7. However, the reproduction number for B.1.1.7 dropped throughout January, which means that the ratio of the reproductive numbers for B.1.1.7 and non-B.1.1.7 dropped. This is not expected under our models. We will monitor this pattern over the next weeks to see if there is bias in the recent Risch data or if this phenomena in the data is not covered by our model.

We use the average reproductive number during 01.01.2021-17.01.2021 to calculate the transmission fitness advantage f_c of B.1.1.7 under our continuous time model. Further, we calculate the transmission fitness advantage f_d under a discrete time model. In both cases, we assume a generation time g of 4.8 days (which is the mean generation time assumed in the estimation of the reproductive number). The fitness values are shown in Table 3. On a national level, we estimate a fitness advantage of 49-65% across methods and datasets. The regional estimates have an overlap with this interval. We note that we use the national reproductive number for the regional f_c estimate. Since Ticino had a lower and the Lake Geneva region a higher reproductive number averaged over all variants [25], the f_c for Ticino may be an underestimate and the f_c for Geneva an overestimate.

Next, we show the expected dynamics of the epidemic under the continuous model using parameter values based on epidemic conditions in the first half of January (Fig. 4 and 5). We show the development of confirmed case numbers based on this model in the blue (B.1.1.7) and green (non-B.1.1.7) areas. In particular, in January, the model projects a decline in overall case numbers due to a decline in non-B.1.1.7 variants. However, the B.1.1.7 variant is increasing. Under this model, once

medRxiv preprint doi: https://doi.org/10.1101/2021.03.05.21252520; this version posted March 9, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

B.1.1.7 becomes dominant, the total number of cases will be increasing again. It is important to note that this simple model assumes that the transmission dynamics do not change and in particular does not include the effects of vaccination, immunity after infection, and population heterogeneity.

We investigate if our empirical data follows this trend. The dark blue line shows the estimated number of B.1.1.7 cases. Again, this number is the product of the total number of confirmed cases for a day by the proportion of the B.1.1.7 variant for that day. In dark green, we plot the total number of confirmed cases. If the empirical data develops as the model, the dark blue line follows the upper end of the blue area and the dark green line follows the upper end of the green area.

Throughout January, the model and the empirical data follow the same trends across datasets and regions with the exception of Ticino and the Lage Geneva region. For the national level and the five regions with a good match in January, we observe that the empirical total case numbers drop faster compared to the model starting early February. Data over the next week about the B.1.1.7 variant will reveal if this pattern will be confirmed for the variant and if the reproductive number for both the B.1.1.7 variant and the non-B.1.1.7 variants is declining in February.

The discrepancy for Ticino and Lake Geneva is not surprising: they had a reproductive number which was different from the national reproductive number in the first half of January. For Ticino, the empirical case numbers drop faster than the model, which is in line with a lower reported reproductive number compared to the national level. For the Lake Geneva region, the empirical case numbers drop slower than the model, which is in line with a higher reported reproductive number compared to the national level[25]. For all regions but Ticino, we have enough data to estimate a reproductive number for the non-B.1.1.7 variants for 01.01.2021-17.01.2021. While for Switzerland, we obtained a point estimate of 0.83, the point estimates for all regions but Lake Geneva are between 0.81-0.83. Thus using the point estimate for all of Switzerland for the regional plots in Fig. 5 is justified. For Geneva, we obtain a point estimate of 0.88. We use this point estimate estimate in a Lake Geneva specific model (Fig. 6). Again we observe that the total number of confirmed cases dropped recently faster compared to the model.

medRxiv preprint doi: https://doi.org/10.1101/2021.03.05.21252520; this version posted March 9, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

Region	a	t_0	$f_d = exp(ag) - 1$	$f_c = ag/R_c$
Switzerland (Viollier)	0.09 [0.09; 0.10]	62 [60; 64]	0.56 [0.50; 0.62]	0.54 [0.49; 0.58]
Switzerland (Risch)	0.10 [0.09; 0.10]	63 [62; 65]	0.59 [0.54; 0.65]	0.57 [0.53; 0.61]
Central Switzerland	0.09 [0.05; 0.14]	72 [57; 88]	0.55 [0.22; 0.88]	0.53 [0.27; 0.79]
Espace Mittelland	0.08 [0.07; 0.09]	62 [58; 66]	0.46 [0.38; 0.54]	0.45 [0.39; 0.52]
Nordwestschweiz	0.12 [0.10; 0.14]	59 [57; 62]	0.79 [0.63; 0.94]	0.70 [0.60; 0.81]
Tessin	0.06 [-0.01; 0.12]	81 [34; 129]	0.31 [-0.11; 0.74]	0.33 [-0.06; 0.72]
Zurich	0.09 [0.07; 0.11]	69 [63; 75]	0.52 [0.38; 0.66]	0.50 [0.39; 0.62]
Lake Geneva	0.11 [0.09; 0.14]	52 [49; 55]	0.71 [0.50; 0.91]	0.64 [0.50; 0.79]
Ostschweiz	0.10 [0.08; 0.13]	63 [57; 69]	0.66 [0.42; 0.89]	0.61 [0.44; 0.78]

Table 3: Estimates for the growth rate a and the sigmoid's midpoint t_0 (measured in days after Dec. 14) as well as the transmission fitness advantages f_d and f_c are reported. In the f_c calculation, the Swiss-wide estimate of the reproductive number for the time period 01.01.2021-17.01.2021 is assumed for the R_c.

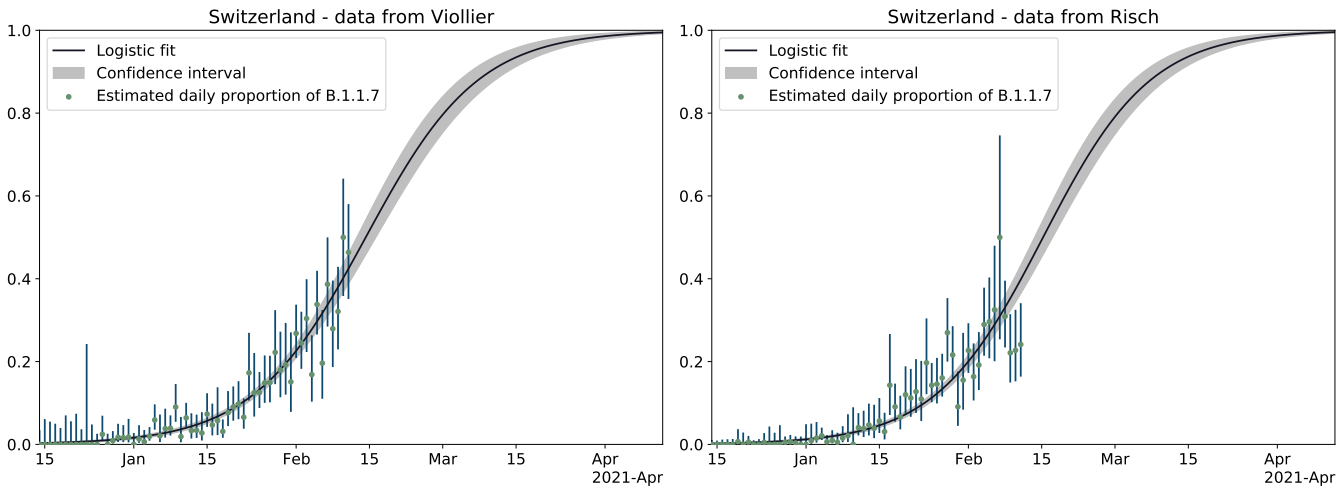

Figure 1: Logistic growth of frequency of B.1.1.7 in Switzerland. Green points are the empirical proportions of B.1.1.7 for each day (i.e. number of B.1.1.7 samples divided by total number of samples). Blue vertical lines are the estimated 95 % uncertainty of this proportion for each day, assuming a simple binomial sampling and Wilson uncertainty intervals. A logistic growth function fit to the data from all of Switzerland is shown in black with the 95 % uncertainty interval of the proportions in gray (i.e. $p(t)$ from Eqn. 2 and 5).

medRxiv preprint doi: https://doi.org/10.1101/2021.03.05.21252520; this version posted March 9, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

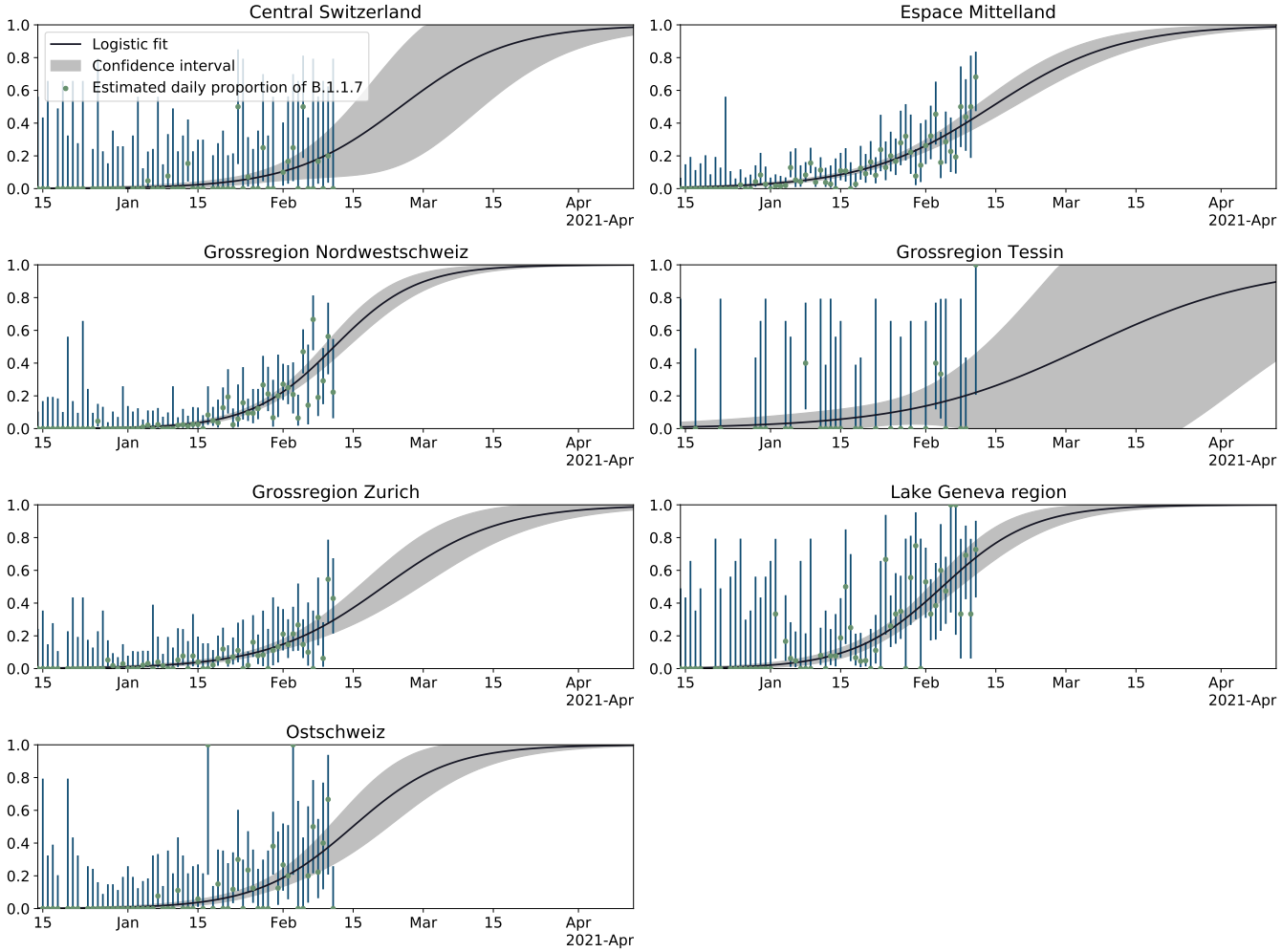

Figure 2: Logistic growth of frequency of B.1.1.7 in the seven economic regions of Switzerland. For details see legend of Fig. 1.

medRxiv preprint doi: https://doi.org/10.1101/2021.03.05.21252520; this version posted March 9, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

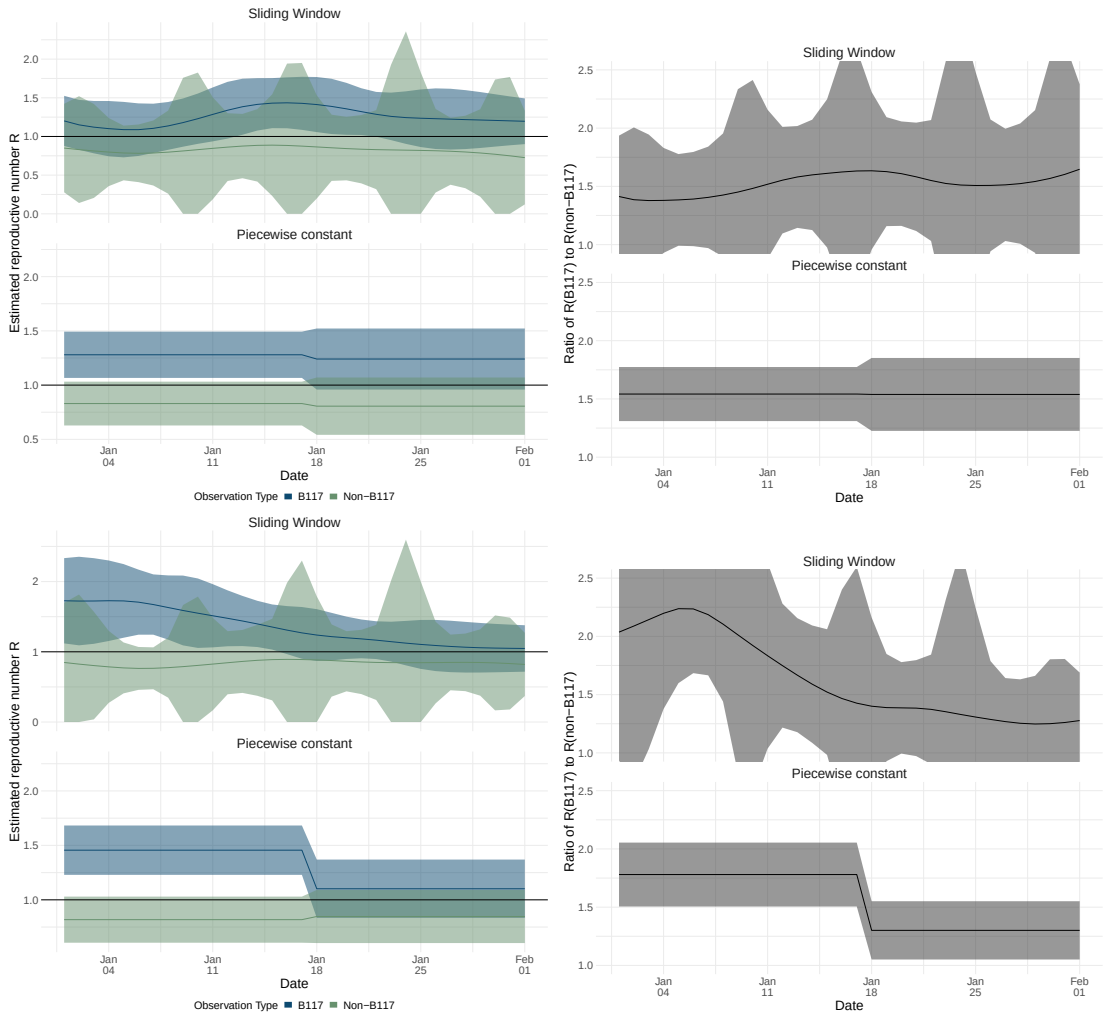

Figure 3: Estimates of the effective reproductive number R of the B.1.1.7 variant and non-B.1.1.7 variants. Results in the top row are based on Viollier data, and in the bottom row based on Risch data. Within each panel, the top row shows the results of the continuously varying R estimation, and the bottom of the piecewise constant R estimation. The left column shows the R estimates, whereas the right shows the ratio between R estimated for B.1.1.7 and R estimated for all non-B.1.1.7 variants. The confidence intervals for the R of non-B117 variants show a 7-day periodicity due to lower case reporting on weekends.

medRxiv preprint doi: https://doi.org/10.1101/2021.03.05.21252520; this version posted March 9, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

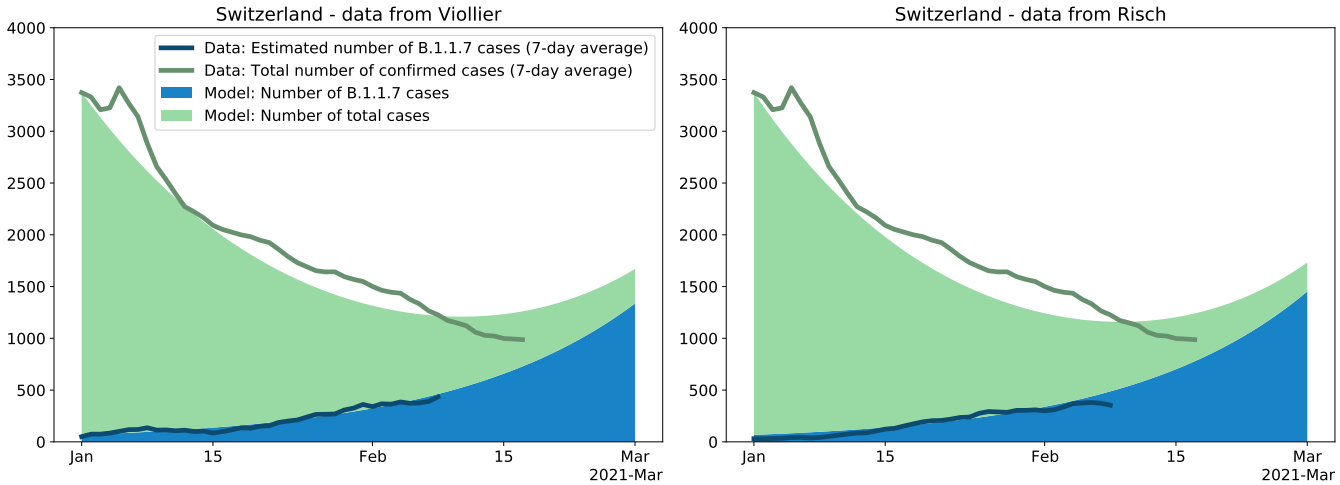

Figure 4: Change in the number of B.1.1.7 variants and in the number of all cases through time. Based on the average reproductive number R_c for Switzerland estimated for the time period 01.01.2021-17.01.2021 (i.e. prior to the tightening of measures on 18.01.2021) and the transmission fitness advantage f_c for the same time period, we plot the expected number of B.1.1.7 variants (blue) and the expected number of non-B.1.1.7 variants (green) under the continuous model. The model is initialized on Jan. 1 with the total number of cases and the estimated number of B.1.1.7 cases on that day. This model is compared to data: The dark green line is the total number of confirmed cases (7-day average). The dark blue line is the estimated number of confirmed B.1.1.7 cases.

medRxiv preprint doi: https://doi.org/10.1101/2021.03.05.21252520; this version posted March 9, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

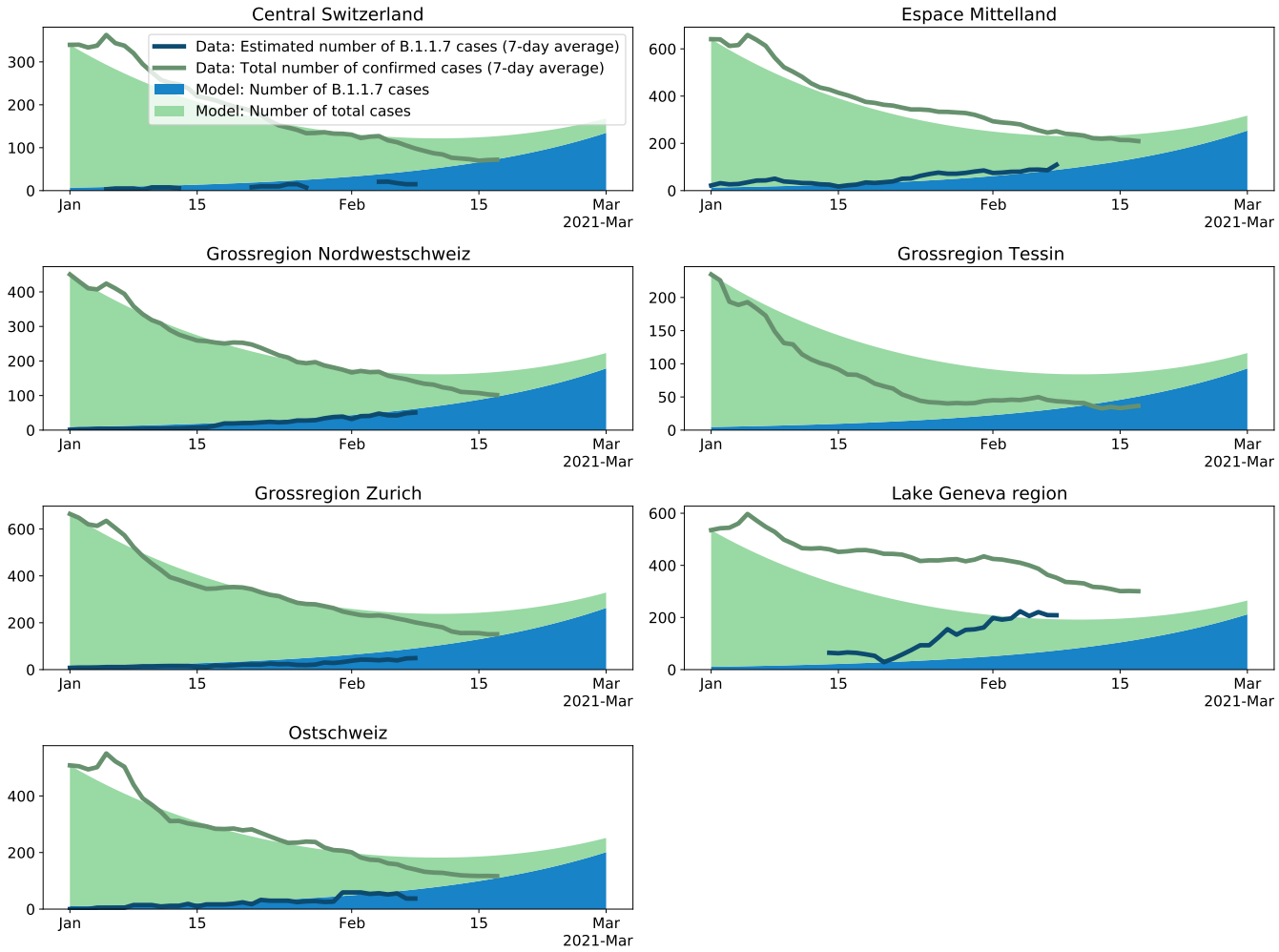

Figure 5: Change in the number of B.1.1.7 variants and in the number of all cases through time. For details see legend of Fig. 4. We use the reproductive number estimated for the whole of Switzerland for the continuous time model such that we can compare to what extend regions differ from the national dynamic. The regional transmission fitness advantage is taken from Table 3.

medRxiv preprint doi: https://doi.org/10.1101/2021.03.05.21252520; this version posted March 9, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

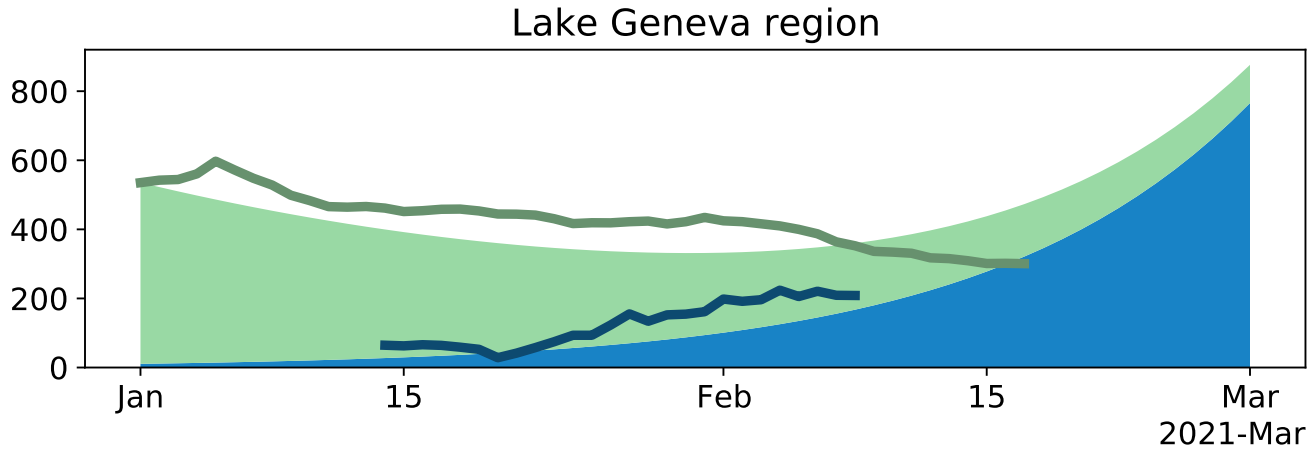

Figure 6: Change in the number of B.1.1.7 variants and in the number of all cases through time. For details see legend of Fig. 4. Compared to 4, we here use the average reproductive number estimated for non-B.1.1.7 in Geneva for the time period 01.01.2021-17.01.2021. The transmission fitness advantage is calculated based on this reproductive number and the estimate of the growth rate a for the Lake Geneva region.

medRxiv preprint doi: https://doi.org/10.1101/2021.03.05.21252520; this version posted March 9, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

4 Discussion

We have quantified the transmission fitness advantage of B.1.1.7 for two Swiss datasets. One dataset also allows us to obtain estimates for the seven Swiss economic regions. Swiss-wide estimates point towards a transmission fitness advantage of 49-65%. Our estimates predict that in March, B.1.1.7 will be dominant in Switzerland. This is in line with predictions by [10] based on Swiss data tracking 501Y mutations. Cases are confirmed 8-11 days after the time of infection in Switzerland, thus actually the majority of new infections may already be caused by B.1.1.7 as of today – the 24.02.2021.

The increase in frequency of B.1.1.7 in Switzerland occurred as expected given the large transmission advantage. In fact, unless measures such as contact tracing target in particular B.1.1.7, this increase in relative frequency will occur given a transmission advantage. However, variables such as implemented measures, the adherence to measures, or levels of immunity in the population determine how fast the variants increase in absolute numbers.

We show that in the first half of January, the absolute numbers of B.1.1.7 increased (R-value 1.28 [1.07-1.49] for 01.01.2021-17.01.2021) based on our sequence dataset) while the absolute numbers of all other cases decreased (R-value 0.83 [0.63-1.03]). For the rest of January, the reproductive number for B.1.1.7 and non-B.1.1.7 did not change much according to the Viollier data. The Risch data suggests a drop in only the B.1.1.7 reproductive number. However, based on our models on the transmission advantage, we do not expect a change in the ratio of the R-values of the two variants. Future data will reveal if this drop reflects a bias in the data or if this phenomena in the data is not covered by our model.

We observe that the number of confirmed cases in February is lower compared to our model scenario using parameter estimates from the first half of January. This may reflect an overall slow-down of transmission. Data on B.1.1.7 over the next week will reveal if the same slow-down is also observed for the B.1.1.7 variant.

Our data reveals that different regions from Switzerland were sampled with different intensity. Thus, we also performed analyses for the seven economic regions in Switzerland. Here, we expect that the homogeneous sampling intensity is met better compared to the national level. We may observe cases though which are imports from other economic regions compared to local transmission (the same problem of course exists on a national level, though the smaller the scale, the more imports

medRxiv preprint doi: https://doi.org/10.1101/2021.03.05.21252520; this version posted March 9, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

we expect). However, the economic regions represent well-defined regions where we expect a lot of mixing within and less mixing across regions. The estimates of the transmission fitness advantage for the economic regions are largely in line with the national estimates, though of course with larger uncertainty. Two of the regions (Ticino, Central Switzerland) have too little data to make precise statements. Geneva is estimated to be around 1-2 weeks ahead in the B.1.1.7 dynamics compared to Switzerland as a whole. One explanation may be a large number of UK travelers in ski resorts in the Valais in December 2020.

We characterized 4.9% of all confirmed cases over our study period through whole genome sequencing. Our sequencing efforts will continue on a weekly basis with all data being submitted to GISAID. These efforts, where we aim to sequence a random subset of cases throughout Switzerland, are envisioned to facilitate the global tracking of SARS-CoV-2, to allow monitoring the spread of known VoCs, and to enable rapid identification of new VoCs. Such whole genome sequencing efforts are complemented by Swiss-wide VoC-specific tracing efforts [8]. In particular, in the Dr Risch medical laboratories, samples are checked for S dropouts and screened for the 501Y mutation. If a VoC is suspected based on these analyses, whole genome sequencing is performed. In that way, we enrich our whole genome dataset, yielding to an overall characterization of 11.5% of all confirmed cases.

Overall we see consistent signal for a large transmission advantage of B.1.1.7 in Switzerland. This confirms estimates from studies from the UK [1, 2, 3], Denmark [3], and Switzerland ([10]; looking at 501Y mutations). In February, the epidemic appears to have slowed down overall. This may be due to effects of changes in measures on 18.01.2021 and improvement in compliance and adherence. Our core plots are available on [11] and [12] for real-time monitoring. Further, we will add new Variants of Concern to our webpage once they spread in Switzerland.

Contribution

CC: Conceptualization, Data curation, Formal Analysis, Investigation, Methodology, Resources, Software, Validation, Visualization, Writing - original draft; SN: Data curation, Writing - review & editing; IT: Data Curation, Resources, Software, Validation; MM: Formal Analysis, Writing - review & editing; JSH: Formal Analysis, Writing - review & editing; KPJ: Data curation, Formal analysis, Investigation, Methodology, Software, Validation, Visualization, Writing - review & editing; LF:

medRxiv preprint doi: https://doi.org/10.1101/2021.03.05.21252520; this version posted March 9, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

Software; DD: Data curation, Software, Validation; KJ: Formal analysis, Validation; CA: Methodology, Writing - review & editing; NB: Conceptualization, Funding acquisition, Project administration, Supervision, Writing - review & editing; MA: Conceptualization, Investigation, Writing - review & editing; TS: Conceptualization, Funding acquisition, Investigation, Methodology, Project administration, Supervision, Validation, Writing - original draft; Everyone else: Resources, Writing - review & editing.

Acknowledgement

We thank Richard Neher for useful discussions on the maximum likelihood inference. TS acknowledges funding from the Swiss National Science foundation (Special Call on Coronaviruses; 31CA30_196267 and 31CA30_196348). CA received funding from the European Union Horizon 2020 research and innovation programme - project EpiPose (No 101003688).

References

[1] Erik Volz, Swapnil Mishra, Meera Chand, Jeffrey C. Barrett, Robert Johnson, Lily Geidelberg, Wes R. Hinsley, Daniel J. Laydon, Gavin Dabrera, Áine O'Toole, Roberto Amato, Manon Ragonnet-Cronin, Ian Harrison, Ben Jackson, Cristina V. Ariani, Olivia Boyd, Nicholas J. Loman, John T. McCrone, Sónia Gonçalves, David Jorgensen, Richard Myers, Verity Hill, David K. Jackson, Katy Gaythorpe, Natalie Groves, John Sillitoe, Dominic P. Kwiatkowski, The COVID-19 Genomics UK (COG-UK) Consortium, Seth Flaxman, Oliver Ratmann, Samir Bhatt, Susan Hopkins, Axel Gandy, Andrew Rambaut, and Neil M. Ferguson. Transmission of SARS-CoV-2 Lineage B.1.1.7 in England: Insights from linking epidemiological and genetic data. *medRxiv*, page 2020.12.30.20249034, January 2021.

[2] Kathy Leung, Marcus HH Shum, Gabriel M Leung, Tommy TY Lam, and Joseph T Wu. Early transmissibility assessment of the n501y mutant strains of sars-cov-2 in the united kingdom, october to november 2020. *Eurosurveillance*, 26(1):2002106, 2021.

[3] Nicholas G Davies, Sam Abbott, Rosanna C Barnard, Christopher I Jarvis, Adam J Kucharski, James Munday, Carl AB Pearson, Timothy W Russell, Damien C Tully, Alex D Washburne,

medRxiv preprint doi: https://doi.org/10.1101/2021.03.05.21252520; this version posted March 9, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

et al. Estimated transmissibility and severity of novel sars-cov-2 variant of concern 202012/01 in england. *MedRxiv*, pages 2020–12, 2021.

[4] Andrew Rambaut, Nick Loman, Oliver Pybus, Wendy Barclay, Jeff Barrett, Alesandro Carabelli, Tom Connor, Tom Peacock, David L Robertson, Erik Volz, and on behalf of COVID-19 Genomics Consortium UK (CoG-UK). Preliminary genomic characterisation of an emergent SARS-CoV-2 lineage in the UK defined by a novel set of spike mutations. Technical report, December 2020.

[5] Nervtag meeting on sars-cov-2 variant under investigation vui-202012/01. Available at `https://app.box.com/s/3lkcbxepqixkg4mv640dpvvg978ixjtf/file/756963730457`.

[6] Public health england - investigation of novel sars-cov-2 variant: Variant of concern 202012/01. Available at `https://www.gov.uk/government/publications/investigation-of-novel-sars-cov-2-variant-variant-of-concern-20201201`.

[7] Tyler N. Starr, Allison J. Greaney, Sarah K. Hilton, Daniel Ellis, Katharine H. D. Crawford, Adam S. Dingens, Mary Jane Navarro, John E. Bowen, M. Alejandra Tortorici, Alexandra C. Walls, Neil P. King, David Veesler, and Jesse D. Bloom. Deep Mutational Scanning of SARS-CoV-2 Receptor Binding Domain Reveals Constraints on Folding and ACE2 Binding. *Cell*, 182(5):1295–1310.e20, September 2020.

[8] Ana Rita Goncalves Cabecinhas, Tim Roloff, Madlen Stange, Claire Bertelli, Michael Huber, Alban Ramette, Chaoran Chen, Sarah Ann Nadeau, Yannick Gerth, Sabine Yerly, Onya Opota, Trestan Pillonel, Tobias Schuster, Cesar M. J. A. Metzger, Jonas Sieber, Michael Bel, Nadia Wohlwend, Christian Baumann, Michel C. Koch, Pascal Bittel, Karoline Leuzinger, Myrta Brunner, Franziska Suter-Riniker, Livia Berlinger, Kirstine K. Soegaard, Christiane Beckmann, Christoph Noppen, Maurice Redondo, Ingrid Steffen, Helena M. B. Seth-Smith, Alfredo Mari, Reto Lienhard, Martin Risch, Oliver Nolte, Isabella Eckerle, Gladys Martinetti Lucchini, Emma B. Hodcroft, Richard A. Neher, Tanja Stadler, Hans H. Hirsch, Stephen L. Leib, Lorenz Risch, Laurent Kaiser, Alexandra Trkola, Gilbert Greub, and Adrian Egli. SARS-CoV-2 N501Y introductions and transmissions in Switzerland from beginning of October 2020 to February 2021 - implementation of Swiss-wide diagnostic screening and whole genome sequencing. *medRxiv*, page 2021.02.11.21251589, February 2021.

medRxiv preprint doi: https://doi.org/10.1101/2021.03.05.21252520; this version posted March 9, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

[9] Emma B. Hodcroft, Moira Zuber, Sarah Nadeau, Iñaki Comas, Fernando González Candelas, SeqCOVID-SPAIN Consortium, Tanja Stadler, and Richard A. Neher. Emergence and spread of a SARS-CoV-2 variant through Europe in the summer of 2020. *medRxiv*, page 2020.10.25.20219063, October 2020.

[10] Transmission of sars-cov-2 variants in switzerland. Available at `https://ispmbern.github.io/covid-19/variants/`.

[11] Sars-cov-2 variants of concern in switzerland. Available at `https://cevo-public.github.io/Quantification-of-the-spread-of-a-SARS-CoV-2-variant/`.

[12] Sars-cov-2 variants of concern in switzerland. Available at `https://sciencetaskforce.ch/nextstrain-phylogentische-analysen/`.

[13] github repository for "quantification of the spread of sars-cov-2 variant b.1.1.7 in switzerland"; chen et al. 2021. https://github.com/cevo-public/Quantification-of-the-spread-of-a-SARS-CoV-2-variant.

[14] Sarah A Nadeau, Timothy G Vaughan, Jérémie Sciré, Jana S Huisman, and Tanja Stadler. The origin and early spread of sars-cov-2 in europe. *Proceedings of the National Academy of Sciences*, 118(9), 2021.

[15] Illumina covidseq test. Available at `https://emea.illumina.com/products/by-type/ivd-products/covidseq.html`.

[16] Artic v3 multiplex pcr amplicon protocol. Available at `https://artic.network/`.

[17] Health 2030 genome center github. Available at `https://github.com/health2030genomecenter`.

[18] Susana Posada-Céspedes, David Seifert, Ivan Topolsky, Kim Philipp Jablonski, Karin J Metzner, and Niko Beerenwinkel. V-pipe: a computational pipeline for assessing viral genetic diversity from high-throughput data. *Bioinformatics*, 01 2021. btab015.

[19] Madlen Stange, Alfredo Mari, Tim Roloff, Helena MB Seth-Smith, Michael Schweitzer, Myrta Brunner, Karoline Leuzinger, Kirstine K Søgaard, Alexander Gensch, Sarah Tschudin-Sutter,

medRxiv preprint doi: https://doi.org/10.1101/2021.03.05.21252520; this version posted March 9, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

et al. Sars-cov-2 outbreak in a tri-national urban area is dominated by a b. 1 lineage variant linked to mass gathering events. *medRxiv*, 2020.

[20] Luis-Miguel Chevin. On measuring selection in experimental evolution. *Biology letters*, 7(2):210–213, 2011.

[21] Skipper Seabold and Josef Perktold. statsmodels: Econometric and statistical modeling with python. In *9th Python in Science Conference*, 2010.

[22] Jana S. Huisman, Jérémie Scire, Daniel C. Angst, Richard A. Neher, Sebastian Bonhoeffer, and Tanja Stadler. Estimation and worldwide monitoring of the effective reproductive number of SARS-CoV-2. *medRxiv*, page 2020.11.26.20239368, November 2020.

[23] Anne Cori, Neil M. Ferguson, Christophe Fraser, and Simon Cauchemez. A New Framework and Software to Estimate Time-Varying Reproduction Numbers During Epidemics. *American Journal of Epidemiology*, 178(9):1505–1512, November 2013.

[24] Coronavirus: Federal council extends and tightens measures. Available at `https://www.admin.ch/gov/en/start/documentation/media-releases/media-releases-federal-council.msg-id-81967.html`.

[25] Real-time estimates of the reproductive number for sars-cov-2. Available at `https://ibz-shiny.ethz.ch/covid-19-re-international/`.