CAMBRIDGE
UNIVERSITY PRESS

Surveillance of SARS-CoV-2 prevalence from repeated pooled testing: application to Swiss routine data

Julien Riou[1,2] , Erik Studer[3], Anna Fesser[3], Tobias Magnus Schuster[3], Nicola Low[1], Matthias Egger[1,4,5] and Anthony Hauser[1,6]

[1]Institute of Social and Preventive Medicine, University of Bern, Bern, Switzerland; [2]Department of Epidemiology and Health Systems, Unisanté, Center for Primary Care and Public Health & University of Lausanne, Lausanne, Switzerland; [3]Federal Office of Public Health, Liebefeld, Switzerland; [4]Centre for Infectious Disease Epidemiology and Research, Faculty of Health Sciences, University of Cape Town, Cape Town, South Africa; [5]Population Health Sciences, Bristol Medical School, University of Bristol, Bristol, UK and [6]INSERM, Sorbonne Université, Pierre Louis Institute of Epidemiology and Public Health, Paris, France

Abstract

Surveillance of SARS-CoV-2 through reported positive RT-PCR tests is biased due to non-random testing. Prevalence estimation in population-based samples corrects for this bias. Within this context, the pooled testing design offers many advantages, but several challenges remain with regards to the analysis of such data. We developed a Bayesian model aimed at estimating the prevalence of infection from repeated pooled testing data while (i) correcting for test sensitivity; (ii) propagating the uncertainty in test sensitivity; and (iii) including correlation over time and space. We validated the model in simulated scenarios, showing that the model is reliable when the sample size is at least 500, the pool size below 20, and the true prevalence below 5%. We applied the model to 1.49 million pooled tests collected in Switzerland in 2021–2022 in schools, care centres, and workplaces. We identified similar dynamics in all three settings, with prevalence peaking at 4–5% during winter 2022. We also identified differences across regions. Prevalence estimates in schools were correlated with reported cases, hospitalizations, and deaths (coefficient 0.84 to 0.90). We conclude that in many practical situations, the pooled test design is a reliable and affordable alternative for the surveillance of SARS-CoV-2 and other viruses.

Introduction

Surveillance of SARS-CoV-2 infections typically relies on event-based surveillance, with positive reverse transcription PCR (RT-PCR) test results being notified to health authorities. These data have proven useful in estimating trends in transmission dynamics over short periods of time. However, they do not reflect the true incidence of SARS-CoV-2 infection in the population for a variety of reasons, including the motivation of individuals to get tested, targeted testing or required testing from particular groups, accessibility of test centres, out-of-pocket expense, and the increasing use of at-home testing. Younger people with COVID-19, those with mild or no symptoms or with more limited access to healthcare may all be underrepresented [1, 2].

An alternative approach to monitoring is the estimation of the population prevalence of SARS-CoV-2 infection using population-based RT-PCR testing of randomly selected samples. Because of the huge financial and logistical costs involved with individually testing the large number of participants required to reach acceptable precision (especially when the true prevalence is low), the pooled testing design has gained popularity. This approach, proposed by Robert Dorfman in 1943 'to weed out all syphilitic men called up for induction' in the United States military [3], pools samples from randomly selected individuals in groups. One test is performed per pool to save resources. If the test result is negative, the entire pool is considered free of infection. If positive, the samples from the pool can be retested individually to identify cases. This approach is commonly used in blood banks [4], for disease screening [5], and other applications. It has also been used during the COVID-19 pandemic as part of the health protection response, to identify and isolate infectious cases [6–11]. In addition to these objectives, it is also possible to use these data to estimate the prevalence of infection in the population from the number of positive pools, even without retesting individuals [12].

Despite the long history of pooled testing, several challenges remain. These relate to the level of prevalence to which the method can sensibly be applied, the optimal pool size, the independence of infection risk within a pool, and the influence of the accuracy of the test used. For prevalence values below 38%, an optimal pool size can be calculated that minimizes the total number of tests [9]. Above this level, the optimal pool size is 1 and the method, therefore, is irrelevant [13]. The assumption that the individuals in pools are independent regarding their risk

of infection is often violated as pools typically group people that are linked (e.g. family members, classmates, or work colleagues). Counter-intuitively, the clustering of cases may improve the efficiency of pooled testing by increasing the separation between infected and non-infected pools [10]. Finally, several approaches have been developed to deal with test accuracy, and particularly imperfect sensitivity [11, 14]. Daon et al. proposed a method accounting for imperfect sensitivity that takes the number of positive tests within a pool into account [15]. Sensitivity might also differ between nasopharyngeal swabs – the reference standard – and saliva samples, which are preferred for children. The uncertainty in these estimates of sensitivity should also be reflected in the results [16]. A final challenge is dealing with the correlation of test results when pooled testing is repeated over time in the same population.

We developed a Bayesian hierarchical model to estimate the prevalence of infection from pooled test data; (1) correcting for imperfect test sensitivity; (2) propagating the uncertainty in test sensitivity into the results; and (3) including correlation over time and space using Gaussian processes (GPs) and a hierarchical structure. We examined the performance of this framework in a simulation study and applied it to real-world data from Switzerland.

Methods

Modelling pooled testing

We aimed to estimate the population prevalence over time, represented by the latent variable $\pi(t)$, from pooled testing data concerning N_t individuals divided into P_t pools of size M_t, of which K_t have a positive test. From the prevalence $\pi(t)$, we computed the probability $\theta(t)$ that a single pooled test returns positive. In case of perfect specificity and sensitivity of RT-PCR tests, the probability $\theta(t)$ is

$$\theta(t) = 1 - (1 - \pi(t))^{M_t}.$$

We then accounted for the imperfect sensitivity (Se) and specificity (Sp) of RT-PCR tests. We followed the approach of Daon et al. who observed that test sensitivity increases with pool size and assumed that imperfect specificity was caused by sample contamination and thus does not depend on pool size [15]. The probability $\theta(t)$ that a single pooled test returns positive becomes:

$$\theta(t) = 1 - Sp(1 - Se \cdot \pi(t))^{M_t}.$$

See Supplementary Material S1, Section 1.1 for more details.

Inference

The sensitivity and the specificity were treated differently. We assumed a fixed specificity of 100% (this assumption was relaxed in a sensitivity analysis, Supplementary Material S1, Section 2.3). On the other hand, due to varying estimates reported for the sensitivity, we treated it as a free parameter, propagating the uncertainty about Se into the population prevalence estimate. The sensitivity Se was jointly estimated alongside the prevalence using the number of positive RT-PCR tests L among M individuals with confirmed SARS-CoV-2 infection:

$$\Pr(k = L) = \text{Binomial}(M, Se).$$

For this, we used data from 20 studies, with a total sample size of 8,026 [17], resulting in a mean Se value of 84.6%. Last, we linked this probability to the pooled test results using a binomial likelihood:

$$\Pr(k = K_t) = \text{Binomial}(P_t, \theta(t)).$$

Information about prior distributions and parameter choice can be found in Supplementary Material S1, Section 1.2.

Spatio-temporal structure

We used inverse-logit transformed GPs to express the temporal correlation of prevalence $\pi(t)$ [18]:

$$\pi(t) \sim \text{logit}^{-1}(\text{GP}(\mu, \Sigma))$$

where the GP is defined by a mean μ and a kernel Σ. Σ is characterized by an exponentiated quadratic kernel, where the covariance between prevalence at two time points decreases exponentially with the interval between these points. The two parameters of the exponentiated quadratic kernel, that is, the length scale and the variance parameters, were estimated during the fitting procedure. We used a hierarchical structure with two levels expressing the spatial correlation of prevalence. To provide an aggregated prevalence estimated for each higher-level area while allowing for some variation between the subareas belonging to the same higher-level area, we used a beta-binomial distribution [19]:

$$\Pr(k = K_{i,j,t}) = \text{Beta-binomial}(P_{i,j,t}, \kappa\theta_j(t), \kappa(1 - \theta_j(t))),$$

where i refers to the subarea, j refers to the higher-level area, and κ is a dispersion parameter. This approach can be adapted to any two nested geographical areas to estimate the prevalence at the highest level. For Switzerland, the subareas correspond to the cantons (NUTS-3 in the Eurostat nomenclature of territorial units), which are nested within regions (NUTS-2) or the entire country [20].

Simulation study

We performed a simulation study to validate the ability of the model including a spatiotemporal structure (thereafter named *GP model*) to estimate population prevalence. We simulated four scenarios mimicking different dynamics of SARS-CoV-2 prevalence in five hypothetical subareas over 30 weeks, using a deterministic susceptible-exposed-infected-recovered (SEIR) model (Supplementary Material S1, Section 1.3). In scenario 1, the prevalence was kept relatively stable at around 2.5%, while it increased from 0% to 5% in scenario 2, increased up to 7.5% in one wave in scenario 3, and increased up to 10% in two successive waves in scenario 4 (Figure 1a). We introduced heterogeneity in prevalence across subareas with a beta distribution. For each scenario and region, we simulated pooled test data for each week of different pool sizes (5, 10, or 20) and total sample sizes (100, 500, 1,000, or 5,000, equally spread over the five subareas). We then applied the GP model to the simulated data to estimate the weekly overall prevalence and compared estimates to the true values used in the simulation. For comparison, we also assessed the performance of a *naive* model that ignored the temporal correlation and estimated the prevalence for each week independently. We repeated the procedure 100 times for each scenario and combination of pool size and total sample size, and calculated two metrics: (1) the root mean squared error (RMSE) of the point estimate of prevalence, measuring *accuracy*, and (2) the half-width of the 95% credible
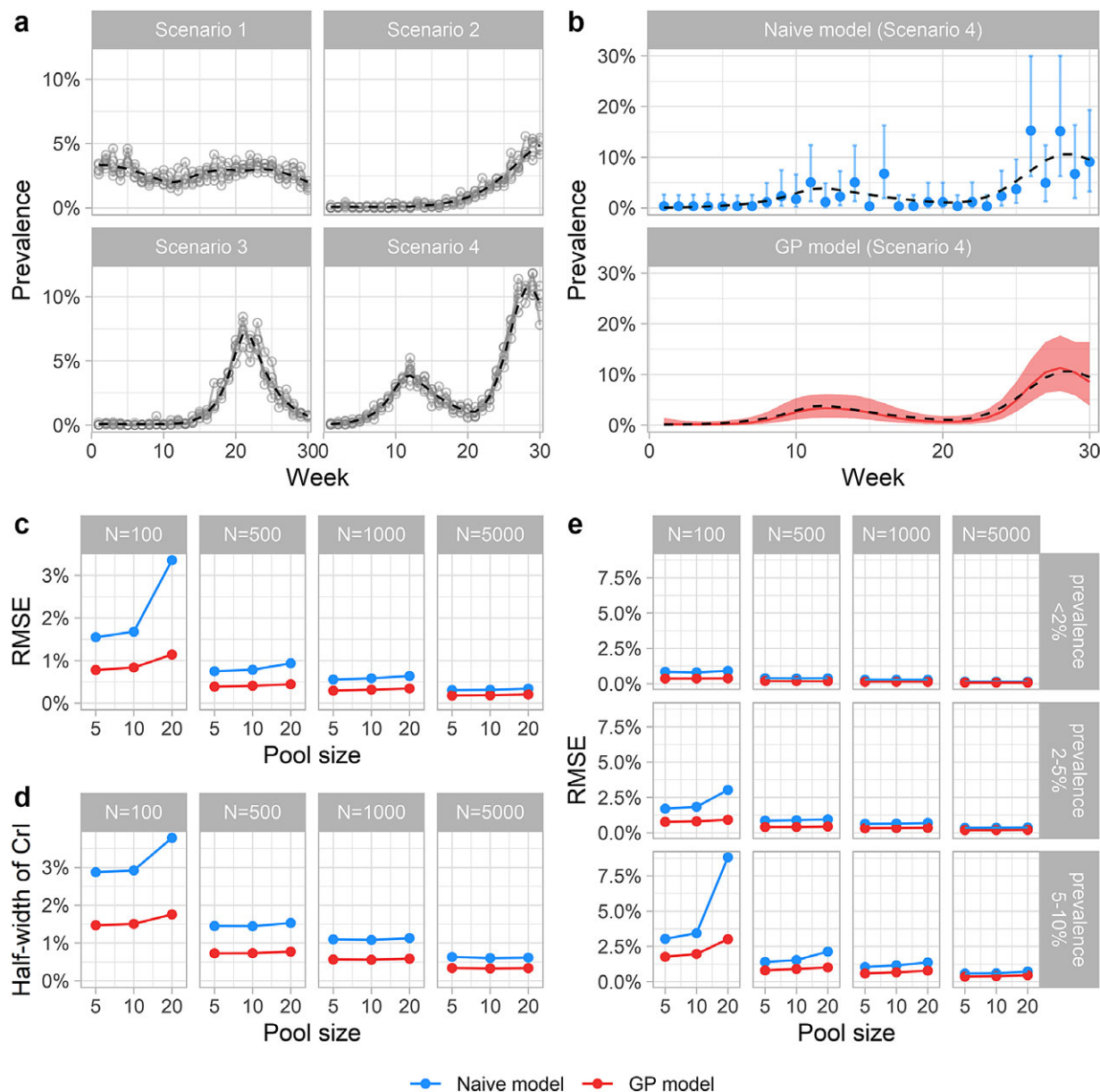
Figure 1. Simulation study (a) Three scenarios of prevalence over time were used to simulate pooled test data. The dashed line shows the overall prevalence, and the circles show the prevalence values across five subareas. (b) Example of model fit to scenario 3 for the naive model and the Gaussian process (GP) model. (c) Root mean squared error (RMSE) measuring the accuracy of prevalence estimates obtained from the naive model or the GP model (lower values are better), for each combination of pool size P (5, 10, or 20) and total sample size N (100, 500, 1,000, or 5,000). (d) Half-width of the 95% credible interval measuring the sharpness of prevalence estimates obtained from the naive model or the GP model (lower values are better). (e) RMSE of prevalence estimates obtained from the naive model or the GP model according to the population prevalence (<2%, 2–5%, and 5–10%).

interval (95%CrI) around the prevalence estimate, measuring *sharpness* [21].

Analysis of Swiss data

We applied the model to the Swiss weekly pooled test data from 19 April 2021 to 29 August 2022. During this period, repeated pooled testing was conducted as part of the wider health protection response, with the primary objective of identifying and isolating infectious cases of SARS-CoV-2 infection. It was not designed to produce prevalence estimates, so the data collection was secondary to the immediate use of information by local health authorities. The program was implemented in each canton independently but

followed similar modalities. The implementation was supported by the federal government with seed financing of the software and logistical infrastructure, and cost coverage of each test. All samples were tested using RT-PCR, and most cantons used the same infrastructures and laboratories. All laboratories were accredited. The data were collected by the 26 Swiss cantons and it was legally required to be sent to the Federal Office of Public Health (FOPH) weekly. Each canton had to report the weekly numbers of total pools, positive pools and number of participants. We summarized the data by the seven Swiss NUTS-2 regions (Central, Eastern, Lake Geneva, Middle ('Mittelland'), Northwest, Ticino, and Zurich [20]) to estimate weekly regional prevalence estimates from multiple observations at the cantonal level (except for canton

Ticino which is also a NUTS-2 region by itself). We also summarized cantonal-level observations to obtain country-level estimations of prevalence for each week with the same approach. We used multiple imputation to impute missing pool sizes (0.7% of cases, Supplementary Material S1, Section 1.4) [22].

Pooled tests with a pool size of 4 or larger were financed by the state and performed in a selection of (1) schools, (2) care centres, and (3) workplaces in order to reduce the need for global control measures indiscriminate to local conditions. Samples from schools included children and their teachers. The care setting included long-term care facilities, elderly people homes and hospitals and comprised both staff and patients or residents. Workplaces included several companies and public administrations. In most cases, saliva samples were pooled on-site, before being transported to the laboratory, but the pooling could also be done at the laboratory. There were standardized recommendations for all processes including sample collection and pooling, transportation, waste management, and molecular analysis (Supplementary Material S2). In all settings, individuals could contribute repeatedly over successive weeks.

We considered the three settings separately and obtained prevalence estimates over time for each setting and NUTS-2 region, and at the national level. We used Spearman's rank correlation coefficient on posterior samples to compare the prevalence estimates across areas and settings and with other data on the SARS-CoV-2 epidemic in Switzerland (counts of reported cases of SARS-CoV-2 infection, COVID-19 hospitalizations and COVID-19 deaths). We performed all the analyses in R version 4.2.1 [23] and Stan version 2.29.1 [24]. The code is available from https://github.com/erikstuder/poolprevBAG. We also developed the R package poolprev that can be used to apply these methods in other settings (https://github.com/anthonyhauser/poolprev).

Results

Simulation study

The simulation study demonstrated the ability of the GP model to accurately estimate prevalence from pooled test data (Figure 1a).

Figure 1b illustrates the superior model fit of the GP model compared to the naïve model for scenario 3. The RMSE from the GP model was consistently below or equal to 1.5 percentage points across combinations of pool size (5, 10, and 20) and total sample size (100, 500, 1,000, and 5,000) (Figure 1c). Both the accuracy (measured by the RMSE) and the sharpness (measured by the half-width of the 95%CrI) of the prevalence estimates increased with larger sample sizes and, to a small extent, with smaller pool sizes (Figure 1c,d). The quality of the estimates dropped when the sample size was small ($N = 100$) and the pool size was large ($N = 20$). This deterioration of quality was, however, observed only for higher prevalence values (Figure 1e), suggesting that sample sizes of 500 or above with pool sizes between 5 and 10 are sufficient to produce reliable estimate for prevalence values below 5%. Compared to the naive model ignoring the temporal correlation, the GP model had better accuracy and sharpness in all situations.

Application to Swiss data

A total of 1,439,984 pooled tests were done in Switzerland over the study period (Table 1). Of these, 837,278 (58%) were from samples collected in schools, 169,634 (12%) in care settings, and 433,072 (30%) in workplaces (Figure 2a). The number of pooled tests varied over time (for example, in schools depending on holidays). Data collection stopped in schools in April 2022 and continued at lower levels in the other settings until August 2022. The distribution of pooled tests across the seven Swiss regions was unequal, with few data for the canton Ticino, which was excluded from region-level analyses. The average number of individuals per pool ranged between 4 and 48, with a median of 6 (interquartile range 5–8) (Table 1 and Figure 2b). The proportion of positive pools varied over time, with the average proportion of positive pools at the country level following similar patterns in the three settings (Figure 2c).

Based on these data, we estimated SARS-CoV-2 prevalence in the three settings over time at the national level (Figure 3). In schools, there were three waves of increasing magnitude: a first small wave that peaked in September 2021 at an estimated

Table 1. Pooled test data collected for monitoring SARS-CoV-2 in Switzerland between 19 April 2021 and 29 August 2022

Setting	Total number of pools	Number of positive pools	Median pool size
Overall	1,439,984	73,525 (5%)	6 (5–8)
By setting	–	–	–
Schools	837,278 (58%)	42,471 (5%)	7 (6–10)
Care centres	169,634 (12%)	12,624 (7%)	6 (5–7)
Selected workplaces	433,072 (30%)	18,430 (4%)	6 (5–7)
By setting	–	–	–
Central	203,013 (14%)	11,365 (6%)	6 (5–7)
Eastern	306,709 (21%)	15,849 (5%)	7 (5–9)
Lake Geneva	37,208 (3%)	2,464 (7%)	6 (5–7)
Mittelland	247,117 (17%)	10,877 (4%)	6 (5–7)
Northwest	185,955 (13%)	9,515 (5%)	8 (5–11)
Ticino	453 (0%)	53 (12%)	5 (5–9)
Zurich	459,529 (32%)	23,402 (5%)	6 (6–7)

Note: In the column 'Total number of pools', the percentages in bracket indicates the distributions of the pools over the three settings and the seven regions. In the column 'Number of positive pools', the percentages in bracket are the pool positivities. In the column "Median pool size", the numbers in bracket indicate the interquartile ranges.
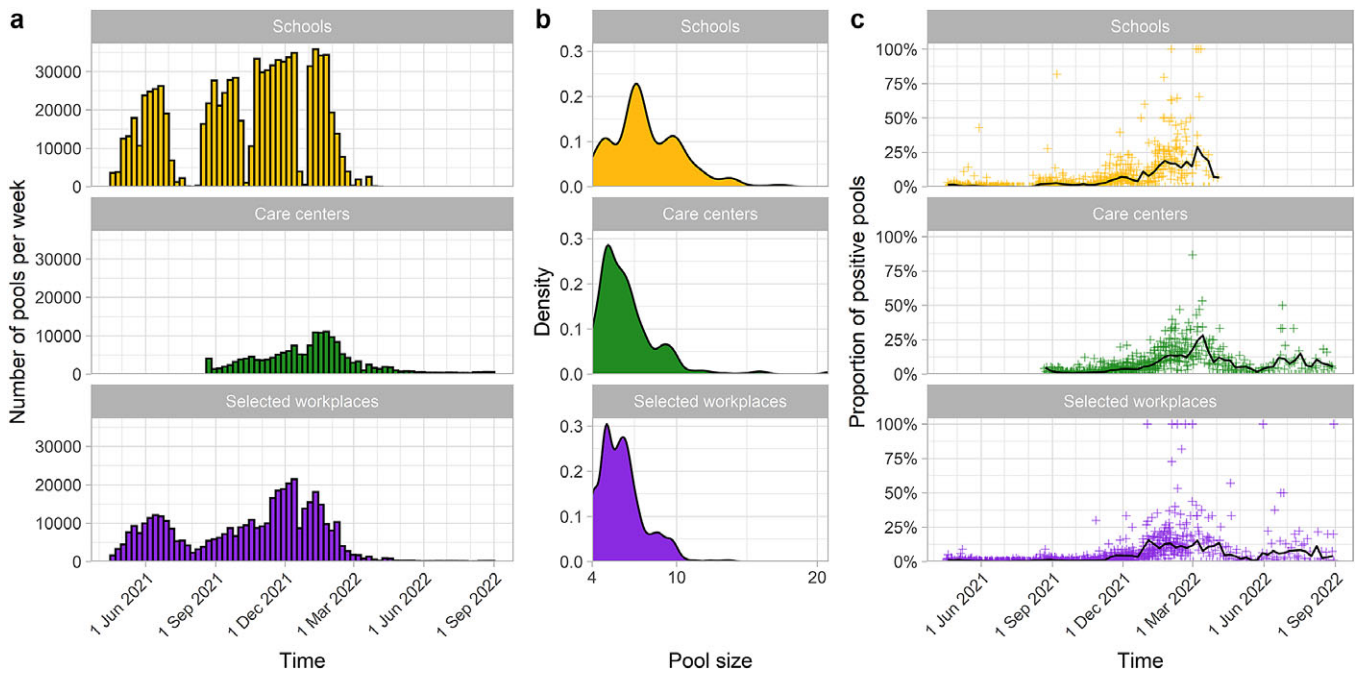
Figure 2. Pooled test data collected for monitoring SARS-CoV-2 in Switzerland between 19 April 2021 and 29 August 2022 in three types of setting (schools, care centres, and selected workplaces): (a) the total number of pools per week (truncated at 20); (b) distribution of the pool size (cut at 30); and (c) proportion of positive pools over time (the line corresponds to weekly average).
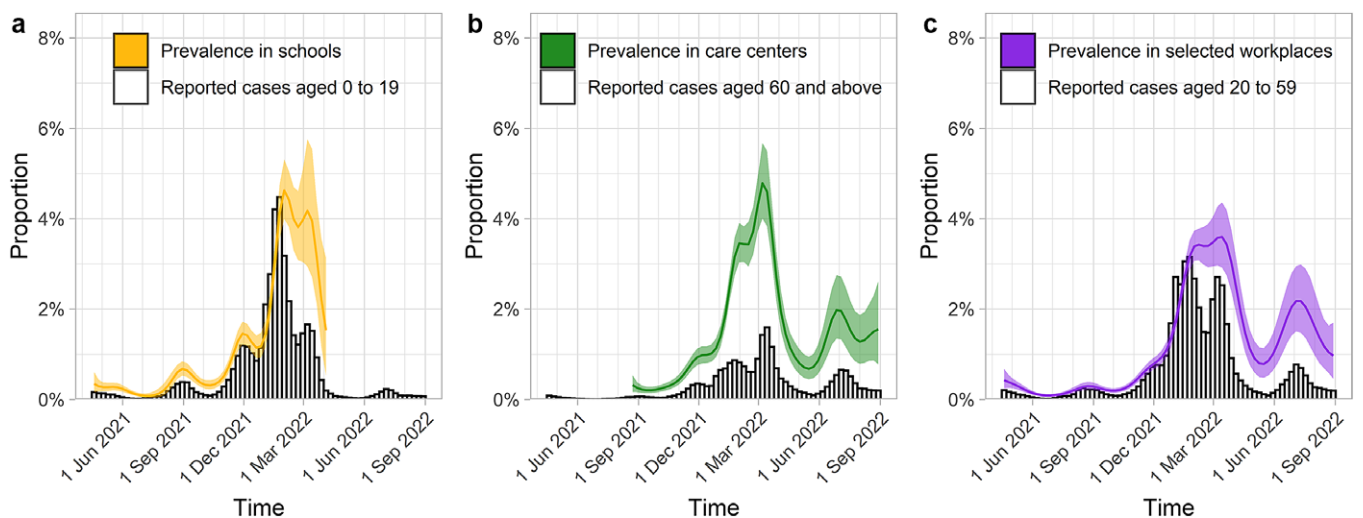
Figure 3. SARS-CoV-2 prevalence estimated from pooled test data in Switzerland between 19 April 2021 and 29 August 2022 in three settings: schools (panel a), care centres (panel b), and selected workplaces (panel c). The lines represent the posterior means, and the coloured areas the 95% credible intervals. The white bars show the weekly number of reported cases of SARS-CoV-2 infections in corresponding age groups scaled by population (0–20 for schools, above 60 for care centres, and 21–60 for selected workplaces).

prevalence of 0.7%, a second short wave peaking in early December 2021 at about 1.4%, immediately followed by a third, larger wave from January to March 2022 that peaked at 4.6% prevalence. The prevalence trajectory was similar in care centres and selected workplaces, with a large wave from December 2021 to the end of March 2022 peaking at 4.8% and 3.6%, respectively. While pooled testing in schools was discontinued at this time, it continued in the care centres and workplaces, with another wave in the summer of 2022 peaking at around 2%.

Prevalence in schools was generally slightly higher than in the other two settings (Figure 4a). This was especially apparent during the early phase of the largest wave, around February 2022, suggesting an earlier peak in schools. The dynamics of SARS-CoV-2 prevalence across regions of Switzerland were generally synchronous, with important differences in magnitude (Figure 4b,c). The largest wave of December 2021 to March 2022 started slightly earlier in the Lake Geneva region in all three settings, but reached higher values in the Central region, especially in schools. We also
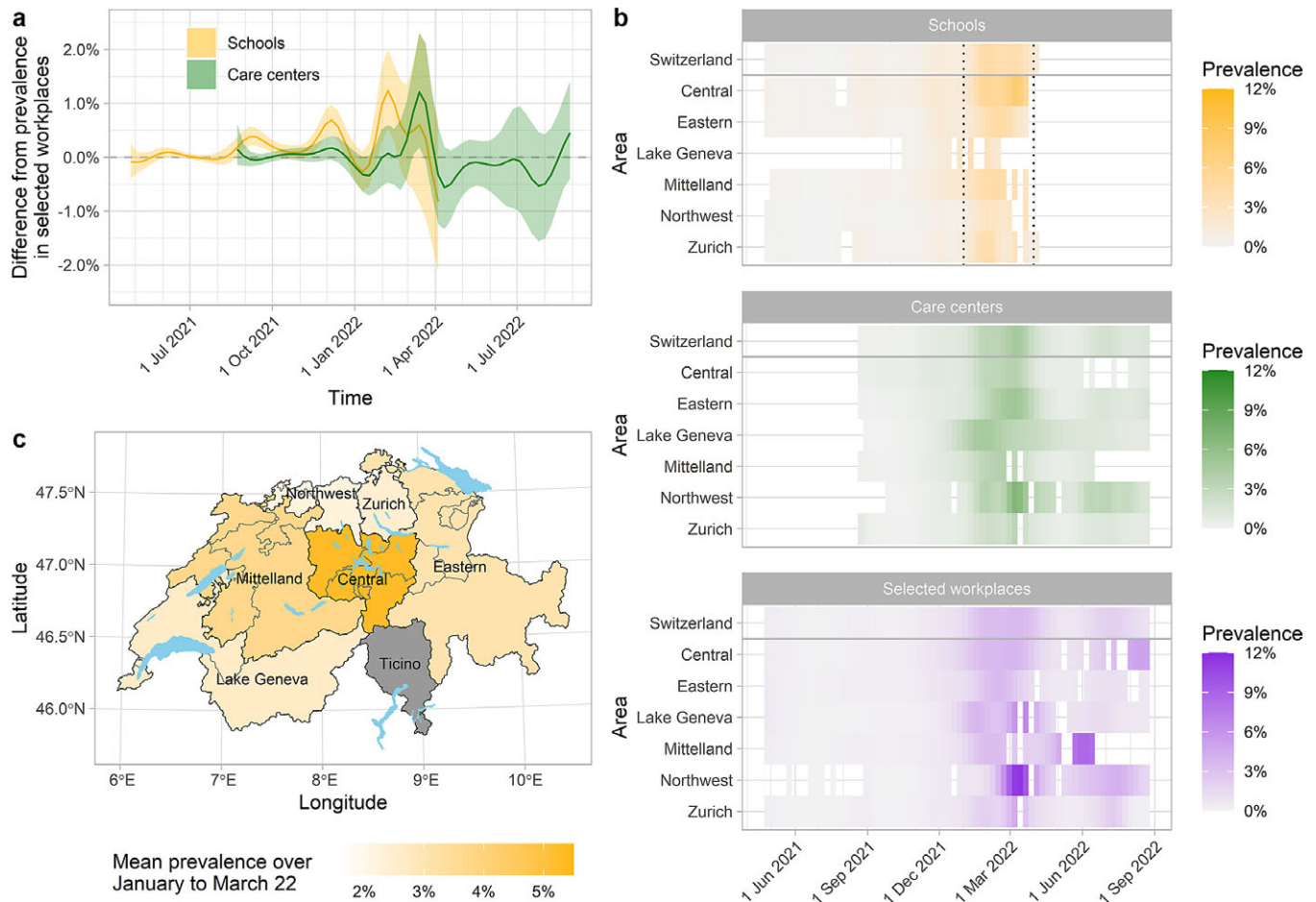
Figure 4. (a) Absolute difference (with 95% credible interval) in SARS-CoV-2 prevalence estimated from pooled test data at the national level across settings (selected workplaces are taken as the reference). (b) SARS-CoV-2 prevalence estimated from pooled test data in six regions of Switzerland between 19 April 2021 and 29 August 2022 in three settings (vertical dotted lines indicate the period used in panel C). (c) Mean prevalence estimated from pooled test data in schools at the regional level between 1 January and 31 March 2022 (Ticino is excluded).

found very high estimates in the Northwest region during this period but concentrated in care centres and workplaces. The steep rise in prevalence estimates observed only in June 2022 in the selected workplaces of region Mittelland appears to be an artefact based on few data points, with very large uncertainty intervals (Supplementary Material S1, Section 2.1, Supplementary Figure S1).

Prevalence estimates from pooled testing were highly correlated across the three settings, with Spearman's correlation coefficient ranging from 0.91 to 0.95 (Table 2). Correlations of prevalence estimates across areas and settings were also generally high, with some exceptions (Supplementary Material S1, Section 2.2, Supplementary Figure S2). The correlation of prevalence estimates from pooled test data with event-based surveillance data varied (Table 2). The estimated prevalence in schools was highly correlated with the weekly counts of reported cases, hospitalizations and deaths (Spearman's correlation coefficients 0.84 to 0.90). For care centres and workplaces, the correlation with event-based surveillance data was weaker (range of coefficients 0.21–0.84). The correlation coefficients for reported cases across the three age groups were consistently high for the estimates from schools (range 0.89–0.93) but lower for the estimates from care centres and workplaces (range 0.38–0.93). The exception was the age group 60 years and above for which the correlation was high for estimates from all three settings (Table 2). Across regions, correlations between reported

cases and prevalence were lower overall but highest for the prevalence estimates from the school setting (Supplementary Material S1, Section 2.1, Supplementary Figure S1).

Discussion

In this study, we demonstrated that the pooled test design can be useful to monitor the prevalence of SARS-CoV-2 infections when analyzed with the appropriate tools. We developed a reliable approach to analyze pooled test data while accounting for the spatial and temporal structure of the data and imperfect test performance. This approach was thoroughly validated in a simulation study and applied to real data from Switzerland. We estimated prevalence levels up to 4–5% during the omicron wave of winter 2021–2022 in three different settings: schools, care centres, and workplaces. Despite the apparent noisiness of the pooled test data, the estimated trajectories of prevalence were consistent across settings at the national level, showing high correlation across settings. These trajectories were also aligned with external data about the dynamics of reported cases of SARS-CoV-2 infection.

Reported incidence has been largely used to track SARS-CoV-2 transmission. Estimates of SARS-CoV-2 prevalence were rarely accessible due to the large sample size required to measure low prevalence levels. The Office of National Statistics (ONS)

Table 2. Correlation between SARS-CoV-2 prevalence estimated from pooled test data and other indicators of the dynamics of SARS-CoV-2 in Switzerland (Spearman's rank correlation coefficient with 95% credible interval)

	Prevalence (Switzerland)		
	Schools	Care centres	Selected workplaces
Prevalence (Switzerland)			
Schools		0.91 (0.87 to 0.96)	0.93 (0.88 to 0.97)
Care centres	0.91 (0.87 to 0.96)		0.95 (0.90 to 0.98)
Selected workplaces	0.93 (0.88 to 0.97)	0.95 (0.90 to 0.98)	
Event–based indicators			
Reported cases (all)	0.90 (0.87 to 0.94)	0.73 (0.66 to 0.80)	0.84 (0.80 to 0.87)
COVID–19 hospitalizations (all)	0.87 (0.83 to 0.90)	0.54 (0.45 to 0.62)	0.67 (0.63 to 0.71)
COVID–19 deaths (all)	0.84 (0.81 to 0.86)	0.21 (0.11 to 0.31)	0.46 (0.41 to 0.50)
Reported cases by age group:			
Aged 0–19	0.89 (0.86 to 0.92)	0.38 (0.30 to 0.47)	0.53 (0.49 to 0.58)
Aged 20–59	0.89 (0.85 to 0.92)	0.75 (0.69 to 0.82)	0.84 (0.81 to 0.88)
Aged 60 and above	0.93 (0.90 to 0.96)	0.89 (0.84 to 0.93)	0.93 (0.91 to 0.95)
Reported cases by region:			
Central	0.91 (0.86 to 0.95)	0.63 (0.45 to 0.80)	0.69 (0.52 to 0.83)
Eastern	0.89 (0.82 to 0.94)	0.58 (0.42 to 0.73)	0.80 (0.70 to 0.89)
Lake Geneva	0.55 (0.14 to 0.86)	0.74 (0.43 to 0.91)	0.78 (0.62 to 0.89)
Mittelland	0.92 (0.86 to 0.97)	0.84 (0.69 to 0.93)	0.66 (0.59 to 0.77)
Northwest	0.93 (0.88 to 0.96)	0.56 (0.41 to 0.64)	0.70 (0.61 to 0.76)
Zurich	0.92 (0.89 to 0.94)	0.74 (0.64 to 0.82)	0.89 (0.84 to 0.92)

coronavirus infection survey in the United Kingdom and the REACT-1 study in England collected samples from hundreds of thousands of individuals to estimate prevalence over time. Here, we took advantage of pooled test data collected locally with limited oversight to reconstruct the dynamics of SARS-CoV-2 transmission in Switzerland [25, 26]. Prevalence estimates provide additional insights about the trends based on reported incidence, mainly from newly symptomatic people, which can lead to misleading interpretations due to variations in testing levels. For example, the reported incidence in 0–19 year olds dropped in March 2022, while the estimated prevalence in schools remained high. As this period coincides with a sharp decrease in testing in the 0–19 age group [27], this suggests that the observed decreasing trend is likely wrong, and that reported incidence failed to capture the transmission dynamics.

Our estimates of SARS-CoV-2 prevalence in Switzerland provided other important insights. We found high prevalence in the three settings, reaching 4–5% at the national level during the omicron wave of winter 2021–2022. Such high prevalence levels have been observed in other European countries around the same time, for example, in England where the REACT-1 study estimated a peak prevalence of 6.4% in March 2022 [26]. The comparison of prevalence estimates in the three settings shows a temporal lag between schools and working places or care centres. This suggests that the omicron wave in Switzerland occurred earlier in schools, supporting the literature about the role of schools in enhancing of spread of respiratory infections [28]. The lower immunity levels in children, whose access to vaccination was limited, could also

explain this discrepancy, which was also observed in the REACT-1 study [26].

While the dynamics of prevalence were likely reliably captured, the prevalence estimated from these Swiss pooled testing data could be biased downwards. Pooled tests conducted in schools or workplaces are likely to at least partially exclude infected students, teachers, and workers showing symptoms, as these people would be encouraged to stay at home, and only asymptomatic or mild cases would remain in the pool. This likely creates a selection bias, leading to the underestimation of the community prevalence (although still representing the prevalence in the settings themselves). This bias is expected to be smaller in care centres, as patients and residents should remain even if symptomatic, although they could be discouraged from participating. Care settings are also more likely to implement stringent control measures and might have a lower prevalence, so direct comparison across settings does not allow for a better characterization of this potential selection bias. Interestingly, the pooled test data from schools and workplaces can be considered as complementary to reported cases, the former approach focusing on the asymptomatic and paucisymptomatic portion of the infected population, while the latter focuses on newly symptomatic people who receive a test leading to a notification to the surveillance system.

The general lack of information about individuals and local practices constitutes another important limitation of this work. The modalities of selection of individuals participating in pooled testing, as well as test sensitivity and specificity, could vary across time and space in unexpected fashion, with incentives to participate

or not participate depending on local conditions and policies. The quality of the data reported by the local teams may also vary. As the primary objective of the program was practical and aimed at providing information to local authorities, a systematic protocol was not immediately established at the time of the earliest testing, and some information was not recorded. For instance, we could not access the detailed results of each pool, but only obtained the proportion of positive pools, together with the average pool sizes by week and canton. This could lead to an underestimation of the uncertainty in our prevalence estimates. For the same organizational reasons, pooled test data were not available at some times and places, emphasizing the importance of using Gaussian processes to stabilize the estimates over time.

Our approach has several important strengths, despite the limitations. The simulation study showed that our model based on Gaussian processes could provide reliable prevalence estimates even with sample sizes as small as 100 if the true prevalence remains below 5% (which was the case in most settings throughout the SARS-CoV-2 pandemic). We accounted for the temporal correlation of prevalence and thus limit the impact of missing or scarce data on the estimates. We modelled the imperfect sensitivity of RT-PCR in the context of pool testing, and propagated all sources of uncertainty into the results. We provide an R package that makes the application of our method easy with minimal pool test data (number of pools, number of positive pools, and pool size). Since the beginning of the SARS-CoV-2 pandemic, other approaches that share commonalities with our model have been proposed to appropriately analyze pooled test data. McLure et al. have developed a R package that also provides flexible functions to model the prevalence over time, but does not account for test sensitivity [29]. The European Center for Disease Prevention and Control has also developed methods that adjusted for imperfect test sensitivity, but they did not account for the correlation of pooled test results over time [30].

Conclusion

Pooled testing can be used as a reliable approach to monitor the dynamics of SARS-CoV-2, especially as part of a wider health protection response aimed at mitigating the immediate consequences of an outbreak by identifying and isolating infectious cases. It is more affordable than alternatives based on event-based surveillance or large-scale prevalence studies, but can provide highly accurate estimates with relatively small sample sizes as long as (1) prevalence remains lower than 5% and (2) appropriate tools are used to account for imperfect testing and space–time correlation. For these reasons, it could be considered in pandemic preparedness plans as a potential addition to traditional surveillance strategies in situations of low to intermediate circulation of SARS-CoV-2 and other viruses.

References

[1] **Lipsitch M**, et al. (2015) Potential biases in estimating absolute and relative case-fatality risks during outbreaks. *PLoS Neglected Tropical Diseases* **9**(7), e0003846.

[2] **Tancredi S, Anker D, Rosella L and Chiolero A** (2021) Elimination of covid-19: Beware of surveillance bias. *BMJ* **374**, n2126.

[3] **Dorfman R** (1943) The detection of defective members of large populations. *The Annals of Mathematical Statistics* **14**(4), 436–440.

[4] **García Z**, et al. (1996) Evaluation of a pooling method for routine anti-HCV screening of blood donors to lower the cost burden on blood banks in countries under development. *Journal of Medical Virology* **49**(3), 218–222.

[5] **Gaydos CA** (2005) Nucleic acid amplification tests for gonorrhea and chlamydia: Practice and applications. *Infectious Disease Clinics* **19**(2), 367–386.

[6] **Barak N**, et al. (2021) Lessons from applied large-scale pooling of 133,816 SARS-CoV-2 RT-PCR tests. *Science Translational Medicine* **13**(589), eabf2823.

[7] **Praharaj I**, et al. (2020) Pooled testing for COVID-19 diagnosis by real-time RT-PCR: A multi-site comparative evaluation of 5-& 10-sample pooling. *The Indian Journal of Medical Research* **152**(1–2), 88.

[8] **Simas A**, et al. (2021) Pooling for SARS-CoV2 surveillance: Validation and strategy for implementation in k-12 schools. *Frontiers in Public Health* 2088 **9**, 789402.

[9] **Hanel R and Thurner S** (2020) Boosting test-efficiency by pooled testing for SARS-CoV-2—Formula for optimal pool size. *PLoS One* **15**(11), e0240652.

[10] **Comess S, Wang H, Holmes S and Donnat C** (2022) Statistical modeling for practical pooled testing during the COVID-19 pandemic. *Statistical Science* **37**(2), 229–250.

[11] **Pikovski A and Bentele K** (2020) Pooling of coronavirus tests under unknown prevalence. *Epidemiology & Infection* **148**, e183.

[12] **Bilder CR** (2019) Group testing for estimation. In *Wiley StatsRef: Statistics Reference Online*. Hoboken, New Jersey: John Wiley & Sons, Ltd, pp. 1–11. https://doi.org/10.1002/9781118445112.stat08231.

[13] **Ungar P** (1960) The cutoff point for group testing. *Communications on Pure and Applied Mathematics* **13**(1), 49–54.

[14] **Aprahamian H, Bish DR and Bish EK** (2020) Optimal group testing: Structural properties and robust solutions, with application to public health screening. *INFORMS Journal on Computing* **32**(4), 895–911.

[15] **Daon Y, Huppert A and Obolski U** (2021) An accurate model for SARS-CoV-2 pooled RT-PCR test errors. *Royal Society Open Science* **8**(11), 210704. https://doi.org/10.1098/RSOS.210704.

[16] **Gelman A and Carpenter B** (2020) Bayesian analysis of tests with unknown specificity and sensitivity. *Journal of the Royal Statistical Society: Series C (Applied Statistics)* **69**(5), 1269–1283.

[17] **Marando M, Tamburello A, Gianella P, Taylor R, Bernasconi E and Fusi-Schmidhauser T** (2022) Diagnostic sensitivity of RT-PCR assays on nasopharyngeal specimens for detection of SARS-CoV-2 infection: A systematic review and meta-analysis. *Caspian Journal of Internal Medicine* **13**(Suppl 3), 139–147. https://doi.org/10.22088/cjim.13.0.139.

[18] **Riutort-Mayol G, Bürkner P-C, Andersen MR, Solin A and Vehtari A** (2022) Practical Hilbert Space Approximate Bayesian Gaussian Processes

for Probabilistic Programming. Available at http://arxiv.org/abs/2004.11408 (accessed 13 December 2023).

[19] **Kim J and Lee J-H** (2017) The validation of a beta-binomial model for overdispersed binomial data. *Communications in Statistics-Simulation and Computation* **46**(2), 807–814.

[20] **Eurostat** Overview - regions and cities - Eurostat. Available at https://ec.europa.eu/eurostat/web/regions-and-cities/overview (accessed 17 August 2023).

[21] **Gneiting T**, **Balabdaoui F and Raftery AE** (2007) Probabilistic forecasts, calibration and sharpness. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* **69**(2), 243–268.

[22] **Azur MJ**, **Stuart EA**, **Frangakis C and Leaf PJ** (2011) Multiple imputation by chained equations: What is it and how does it work? *International Journal of Methods in Psychiatric Research* **20**(1), 40–49.

[23] **R Core Team** (2022) *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. Available at https://www.R-project.org/ (accessed 28 November 2023).

[24] **Carpenter B**, et al. (2017) Stan: A probabilistic programming language. *Journal of Statistical Software* **76**(1), 1.

[25] **Pouwels KB**, et al. (2021) Community prevalence of SARS-CoV-2 in England from april to november, 2020: Results from the ONS coronavirus infection survey. *The Lancet Public Health* **6**(1), e30–e38.

[26] **Elliott P**, et al. (2022) Twin peaks: The omicron SARS-CoV-2 BA. 1 and BA. 2 epidemics in England. *Science* **376**(6600), eabq4411.

[27] **Federal Office of Public Health** (2023) COVID-19 switzerland | coronavirus | dashboard. Available at https://www.covid19.admin.ch/en/overview.

[28] **Cauchemez S**, et al. (2011) Role of social networks in shaping disease transmission during a community outbreak of 2009 H1N1 pandemic influenza. *Proceedings of the National Academy of Sciences* **108**(7), 2825–2830.

[29] **McLure A**, **O'Neill B**, **Mayfield H**, **Lau C and McPherson B** (2021) PoolTestR: An r package for estimating prevalence and regression modelling for molecular xenomonitoring and other applications with pooled samples. *Environmental Modelling & Software* **145**, 105158.

[30] **E. C. for Disease Prevention** (2020) Methodology for Estimating Point Prevalence of SARS-CoV-2 Infection by Pooled RT-PCR Testing. Available at https://www.ecdc.europa.eu/en/publications-data/methodology-estimating-point-prevalence-sars-cov-2-infection-pooled-rt-pcr (accessed 1 December 2023).