



Reconstruction of SARS-CoV-2 outbreaks in a primary school using epidemiological and genomic data

Cécile Kremer^{a,*}, Andrea Torneri^a, Pieter J.K. Libin^{a,b,c}, Cécile Meex^d, Marie-Pierre Hayette^d, Sébastien Bontems^d, Keith Durkin^e, Maria Artesi^e, Vincent Bours^e, Philippe Lemey^c, Gilles Darcis^f, Niel Hens^{a,g,1}, Christelle Meuris^{f,1}

^a Interuniversity Institute for Biostatistics and statistical Bioinformatics, Data Science Institute, Hasselt University, Hasselt, Belgium

^b Artificial Intelligence Lab, Department of Computer Science, Vrije Universiteit Brussel, Brussels, Belgium

^c Department of Microbiology, Immunology and Transplantation, Rega Institute for Medical Research, University of Leuven, Leuven, Belgium

^d Department of Clinical Microbiology, University of Liège, Liège, Belgium

^e Laboratory of Human Genetics, GIGA-Institute, University of Liège, Liège, Belgium

^f Department of Infectious Diseases, Liège University Hospital, Liège, Belgium

^g Centre for Health Economics Research and Modelling Infectious Diseases, Vaccine and Infectious Disease Institute, University of Antwerp, Antwerp, Belgium

ARTICLE INFO

Dataset link: <https://github.com/cecilekremer/school-recon>

Keywords:

SARS-coV-2 transmission
Primary school
Outbreak reconstruction
Whole genome sequences
Screening protocol

ABSTRACT

Mathematical modelling studies have shown that repetitive screening can be used to mitigate SARS-CoV-2 transmission in primary schools while keeping schools open. However, not much is known about how transmission progresses within schools and whether there is a risk of importation to households. During the academic year 2020–2021, a prospective surveillance study using repetitive screening was conducted in a primary school and associated households in Liège (Belgium). SARS-CoV-2 screening was performed via throat washing either once or twice a week. We used genomic and epidemiological data to reconstruct the observed school outbreaks using two different models. The outbreaker2 model combines information on the generation time and contact patterns with a model of sequence evolution. For comparison we also used SCOTTI, a phylogenetic model based on the structured coalescent. In addition, we performed a simulation study to investigate how the accuracy of estimated positivity rates in a school depends on the proportion of a school that is sampled in a repetitive screening strategy. We found no difference in SARS-CoV-2 positivity between children and adults and children were not more often asymptomatic compared to adults. Both models for outbreak reconstruction revealed that transmission occurred mainly within the school environment. Uncertainty in outbreak reconstruction was lowest when including genomic as well as epidemiological data. We found that observed weekly positivity rates are a good approximation to the true weekly positivity rate, especially in children, even when only 25% of the school population is sampled. These results indicate that, in addition to reducing infections as shown in modelling studies, repetitive screening in school settings can lead to a better understanding of the extent of transmission in schools during a pandemic and importation risk at the community level.

1. Introduction

During the early stages of the COVID-19 pandemic in the first half of 2020, several studies have investigated the contribution of children to SARS-CoV-2 transmission with results suggesting that schools did not play a substantial role in driving community transmission (Zimmerman et al., 2021; Mensah et al., 2021; Southall et al., 2021). Mensah et al. (2021) found that SARS-CoV-2 infections in school-aged children followed the same trend as adult cases and only declined after a national

lockdown was implemented while keeping schools open, suggesting community transmission impacts transmission in schools. However, this study was based on data from routine symptom-based surveillance not fully capturing asymptomatic infections in children. Another study reported a positive correlation between community cases and cases in schools in England, with weak evidence suggesting that cases in schools lag behind community cases (Southall et al., 2021). In contrast, a study

* Correspondence to: Agoralaan Gebouw D, 3590 Diepenbeek, Belgium.

E-mail address: cecile.kremer@uhasselt.be (C. Kremer).

¹ These authors contributed equally.

in the US found that in-person schooling poses an increased risk of infection to household members, which can be reduced by implementing mitigation measures in schools (Lessler et al., 2021). Similar findings have been reported in Sweden (Vlachos et al., 2021). In addition, the role of children has likely changed throughout the pandemic in light of the different variants of concern (VoC) circulating. In Belgium, the 2021–2022 academic year in primary schools started in September 2021 without any non-pharmaceutical intervention measures in place. When the school year progressed the number of SARS-CoV-2 cases rose and these interventions (i.e., mask wearing for teachers, isolation of symptomatic children, class closure after two cases) were reinstated (Sciensano, 2021). The rise of the Omicron variant in early 2022 caused a surge in case numbers, resulting in numerous school closures throughout the country.

In addition to structural interventions such as improving air quality in schools through better ventilation systems (Pierce et al., 2022; Callies et al., 2022), screening protocols such as repetitive and reactive screening can also aid in preventing the occurrence of large school outbreaks and consequent closures. In reactive screening, a classroom is tested and quarantined after a certain number of cases is observed, whereas in repetitive screening the entire classroom is tested regularly, regardless of whether a case is observed or not. A recent study using an agent-based model found that in a primary school population, weekly screening of 75% of unvaccinated children could reduce the number of cases by 34% on average, compared to symptom-based testing alone. In addition, the number of school days lost would be reduced up to 80% compared to reactive screening policies (Colosi et al., 2022). In line with this, another modelling study found that repetitive screening greatly reduced the attack rate in school, while reactive screening performed only slightly better than symptomatic isolation (Torneri et al., 2022). Furthermore, using a repetitive screening strategy allowed for increasing the class closure threshold without much impact on the attack rate. Besides limiting transmission, data obtained after implementing a repetitive screening strategy can be used to gain a better understanding of the extent of transmission within schools and the risk of children, as well as school employees, bringing infection into their households (Meuris et al., 2021). Such insights are crucial to ensure effective mitigation measures (Pierce et al., 2022).

During the academic year 2020–2021, a prospective surveillance study using repetitive screening was set up in a primary school located in Liège (Belgium) with the objective to better understand the role of children in SARS-CoV-2 transmission. Results from the first part of this study (September to December 2020) indicated transmission occurred mainly between children and among teachers within the school, with occasional spillover to their households (Meuris et al., 2021). Sequence data of rapidly evolving RNA viruses can provide valuable information on transmission events (Campbell et al., 2018b), as also illustrated specifically for SARS-CoV-2 (Abbas et al., 2022). In this study, we make use of the whole genome sequences that were collected from September 2020 to June 2021 to aid in reconstructing the school outbreaks. Furthermore, using a simulation model, we investigate how the accuracy of estimated weekly positivity rates in a school depends on the proportion of a school that is sampled in a repetitive screening strategy.

2. Materials and methods

2.1. Sample collection

Prospective surveillance of SARS-CoV-2 throat carriage was performed among children, their parents, and employees in a single primary school in Liège (Belgium) from September 2020 to June 2021. Weekly screening for SARS-CoV-2 infection through throat washing was performed from September to December 2020. From January to June 2021, screening was done twice a week using the same protocol. An outbreak in March 2021 prompted a change in strategy to three

samples a week for two consecutive weeks. No screening was performed during school holidays. All adults provided informed consent, also on behalf of their children. Mitigation measures evolved during the course of the study, following national guidelines for testing and quarantine in primary schools (Table S3, Supplementary Material). When tested positive for SARS-CoV-2, participants were called by the investigators to fill in a questionnaire about the timing of symptom onset and symptom duration. In addition, they were instructed to isolate in order to limit further spread of the virus. Family members were quarantined according to national guidelines and advice was given on how to try preventing transmission within the household. For full details on the study design, we refer to Meuris et al. (2021).

2.2. SARS-CoV-2 detection and sequencing

Quantitative reverse transcription-PCR was used to detect SARS-CoV-2 and sequencing was performed as described by Freed et al. (2020). RNA was extracted from throat washing (300 μ l) via a Maxwell 48 device using the Maxwell RSC Viral TNA kit (Promega) with a viral inactivation step using Proteinase K, following the manufacturer's instructions. RNA was eluted in 50 μ l of RNase free water. 3.3 μ l of the eluted RNA was combined with 1.2 μ l of SuperScript IV VILOTM Master Mix and 1.5 μ l of H₂O to carry out Reverse Transcription. This was incubated at 25 °C for 10 min, 50 °C for 10 min, and 85 °C for 5 min. PCR was carried out using Q5[®] High-Fidelity DNA Polymerase (NEB) and primers to obtain 1200bp amplicons as described by Freed et al. (2020). PCR conditions followed the recommendations in the sequencing protocol of the ARTIC Network. The samples were multiplexed following the manufacturer's recommendations using the Oxford Nanopore Native Barcoding Expansion kits 1–12, 13–24, and 96, in conjunction with Ligation Sequencing Kit 109 (Oxford Nanopore). Sequencing was carried out on a Minion using R9.4.1 flow cells.

2.3. Statistical analysis

SARS-CoV-2 positivity in adults and children was compared using a mixed effects logistic regression model with random intercepts for classroom and household, accounting for the clustering of individuals. For all other analyses, differences between two groups were compared using Pearson's chi squared, Fisher's exact, Student's *t*, or Mann–Whitney *U*-tests, as appropriate. Results were considered statistically significant at $p < .05$. Analyses were performed using R version 4.1.0.

2.4. Sequence alignment and phylogenetic analysis

Quality assessment of the whole genome sequences was performed using Nextclade (Aksamentov et al., 2021). The sequences were then aligned to the Wuhan-Hu-1 (NC_045512.2) reference sequence using VIRULIGN, which computes a codon correct multiple-sequence alignment relative to this reference sequence (Libin et al., 2018; Hufsky et al., 2020). The maximum likelihood phylogenetic tree was inferred using IQ-TREE v1.6.12 (Nguyen et al., 2014) with ultrafast bootstrap approximation (UFBoot) using 1000 bootstrap replicates to assess branch support (Hoang et al., 2017). ModelFinder (Kalyaanamoorthy et al., 2017) was used to find the best fitting substitution and rate heterogeneity across sites (RHAS) model based on a full tree search for each combination of models. The Bayesian Information Criterion (BIC) was used to identify the optimal combination. TreeTime v0.8.5 (Sagulenko et al., 2018) was then used to create a time-scaled phylogeny based on the maximum likelihood tree, rooted on the reference sequence which represents an isolate obtained early in the pandemic.

2.5. Outbreak reconstruction

To gain insight on the extent of transmission within the school, reconstruction of the outbreaks was performed using two different models, Outbreaker2 and SCOTTI (Campbell et al., 2019; De Maio et al., 2016; Bouckaert et al., 2014). The Outbreaker2 model combines information on the generation interval and contact patterns with a model of sequence evolution in a Bayesian framework (Campbell et al., 2019, 2018a). The model considers, for each case $i = 1, \dots, N$, the probability of a proposed transmission history given the sample collection time t_i and a genome sequence s_i . A contact model accounting for partial sampling (by estimating the proportion of cases sampled, π) and the presence of non-infectious contacts between cases is also included. Transmission events are modelled using augmented data, i.e., a case's infection time T_i^{inf} , most recently sampled ancestor (MRSA) α_i , and the number of generations ($\kappa_i \geq 1$) separating i and α_i . We assumed the generation interval to be Gamma-distributed with a mean of 4.7 days and standard deviation of 2.5 days (Meuris et al., 2021). The distribution of the time-to-collection was assumed to follow a Gamma distribution with a mean of 7 days and standard deviation of 2 days. In a baseline scenario, prior distributions supporting higher values for the proportion of contacts reported (ϵ) and lower values for the probability of non-infectious contacts (λ) were chosen in order to emphasize reported contacts when assigning ancestries (Table S1, Supplementary Material) (Campbell et al., 2019). For the mutation rate an exponential prior with a mean of 10^{-5} substitutions site $^{-1}$ generation $^{-1}$ was used (Wang et al., 2022). The MCMC chain was run for 10^6 iterations with a thinning frequency of 1/100 and a burn-in of 10%. A sensitivity analysis assessing the impact of the choice of prior distributions was performed using a flat Beta(1, 1) prior for each of the parameters π , ϵ , and λ .

For comparison to the Outbreaker2 model we used SCOTTI which, as part of the BEAST2 platform for Bayesian phylogenetic analysis, is based on a structured coalescent model that enables outbreak reconstruction while accounting for within-host evolution and non-sampled cases (De Maio et al., 2016; Bouckaert et al., 2014; Suchard et al., 2018). Each case is modelled as a distinct population, and transmissions between cases are modelled as migration events. Non-sampled cases are modelled by dynamically increasing or decreasing the number of populations. The model assumes that transmission is a priori equally likely between every pair of cases, i.e., the migration rate is the same between every pair of cases for the time they are both exposed. It is further assumed that all cases have the same, constant, within-host pathogen evolution dynamics. In addition to genome sequences, SCOTTI incorporates epidemiological information by only allowing cases to transmit the virus during a specified exposure window. We defined the exposure window as being from 7 days before symptom onset (or sampling for asymptomatic cases) until two days after the sequenced sample was obtained (i.e., assuming cases self-isolate within two days after the sample was obtained). For cases that had symptom onset more than 5 days after the date of their positive sample, the lower bound of this interval was set to one day before the sequenced sample was taken. The best fitting substitution and RHAS model identified by ModelFinder was used. A strict molecular clock was assumed, using an exponential prior with a mean of 2×10^{-6} substitutions site $^{-1}$ day $^{-1}$ (Wang et al., 2022). Because SCOTTI estimates several phylogenetic parameters in addition to inferring the transmission network, the analysis was done separately for three time periods to account for different VoCs (i.e., October–November 2020, December 2020, March 2021). To ensure a high enough but realistic upper bound, although arbitrary, the maximum number of cases (sampled and non-sampled) was set to 10 times the observed number of sequences, with as minimum the number of sequenced samples during each period. The MCMC chain was run for 2×10^7 iterations with a thinning frequency of 1/2000 and a burn-in of 10%.

Convergence of both methods was assessed using trace plots. One case known to have not contributed to transmission and three sporadic cases in May–June not linked to onward transmission within the sampled population were excluded from the reconstruction. Reinfections were considered to be separate cases. For both models, we quantified the uncertainty related to inferred ancestry using the Shannon entropy of the posterior distribution of potential infectors for each case (Shannon, 1997). Given that a case is assigned K possible infectors with frequency f_k , ($k = 1, \dots, K$), the entropy is defined as

$$-\sum_{k=1}^K f_k \log(f_k)$$

such that an entropy of 0 indicates complete posterior support for a given infector. Higher values indicate a larger number of plausible infectors, hence higher uncertainty in ancestry assignment (Campbell et al., 2019). In addition we reconstructed the outbreak without using the genome sequence data to investigate whether inclusion of these data reduced the uncertainty in inferred ancestry.

2.6. Simulation study

A previously developed individual-based model representing a population of primary school children aged 6–12 years old and their teachers was used to investigate how observed positivity rates reflect the true positivity rate depending on the proportion of a school that is sampled in a repetitive screening strategy. Full details of the simulation model are available elsewhere (Torneri et al., 2022, 2020). Briefly, children are assigned to one of 23 classes and interaction among children can occur within as well as between classes. The within- and between-class contact rates for children were set according to the results of a contact data survey that took place in Belgium (Hoang et al., 2019), but we assumed that in a pandemic setting the number of between-class contacts would be reduced to a level of 20% compared to pre-pandemic behaviour. Each class is assigned a number of teachers at a ratio of 1:9 (Departement Onderwijs en Vorming, 2023). Since the type of close contact necessary for transmission would be rare between teachers from different classes during a time in which mitigation measures such as mask wearing are in place, we assumed that teachers only interact with the children and other teachers in their class.

In the model, a contact between an infected and susceptible individual leads to transmission according to a time-varying probability, based on the time since infection. This time-dependent infectiousness is set to represent the shape of the viral load curve for a SARS-CoV-2 infection, under the assumption that a higher viral load corresponds to a higher transmission probability (Buonanno et al., 2020). Infection can be asymptomatic or symptomatic, with symptomatic individuals developing symptoms at the peak of their infectiousness (He et al., 2020; Sun et al., 2021), at which time they can be detected and are placed in isolation for 10 days (Chang et al., 2020). Asymptomatic and symptomatic individuals are assumed to have the same viral progression (Zhou et al., 2020; Zou et al., 2020), but the relative infectiousness of asymptomatic individuals compared to symptomatic individuals was set to 0.5 (Davies et al., 2020). The probability of symptomatic infection was set to 30% for children and 50% for adults (Sinha et al., 2020; Poletti et al., 2021). In a sensitivity analysis, these were set to 65% for children and 85% for adults as observed in the surveillance study. It is assumed that asymptomatic individuals are as infectious as symptomatic individuals, and that children are half as susceptible as adults (Davies et al., 2020). In addition, we assumed that 10% of children acquired immunity from previous infection, and that 30% of adults were immune due to previous infection or vaccination (Sciensano, 2022). In a sensitivity analysis, this was set to 50% for adults.

Classes were closed when the number of detected cases exceeds a threshold of 4, while schools were assumed to remain open regardless of the total number of cases. The sensitivity of PCR tests on saliva samples was assumed to be 86% (Libin et al., 2021). Every week,

two susceptible children are assumed to acquire infection outside the school environment, accounting for disease importation or seeding. We simulated 100 outbreaks for 100 days under different sampling scenarios, considering weekly testing applied to 25%, 50%, 75%, and 100% of the school population at random. As a sensitivity analysis, scenarios assuming twice weekly testing or symptomatic isolation were considered. We investigated how the accuracy of estimated weekly positivity rates, quantified as the absolute deviation between observed and true positivity rates, depends on the proportion of a school that is sampled. In addition, we investigated the accuracy of estimated positivity rates among children and adults separately.

3. Results

3.1. SARS-CoV-2 positivity in the study population

The total study population included 240 individuals, of which 88 children (36.7%) aged 5 to 13 years, 110 parents of these children (45.8%), 14 school employees (5.8%), 22 teachers (9.2%), and 5 participants (2.1%) that were both teacher and parent of a child included in the study. The children and parents made up 66 households, including 22 sibling pairs. The children and teachers were part of 20 class groups at the primary school and kindergarten level, and made up approximately 25% of the school population. Of the adults, 55.9% had received at least one vaccination dose by the end of the study, with 28.3% being fully vaccinated (Fig. S1, Supplementary Material). Of the vaccinated adults, 14.1% were teachers or school employees. Children were not yet eligible for vaccination at the time of the study. Quantitative reverse transcription-PCR was used to detect SARS-CoV-2 and genome sequencing of positive samples was performed. Over the entire study period, 61 individuals (25.4%) tested positive at least once for SARS-CoV-2, of which 22 children (25.0% of all participating children) and 39 adults (25.7% of all participating adults). There were two reinfections during the study period, concerning two healthy (i.e. not immunocompromised) adults. When accounting for clustering of individuals in classrooms and households using a mixed effects logistic regression model, there was no significant difference in positivity for children compared to adults (conditional odds ratio 0.99 (bootstrap 95% CI 0.44–2.10), $p = .979$). The intraclass correlation (ICC) was 0.39 for household and 0.05 for classroom. Before January 2021, the overall positivity rate was 4.2% in adults (52/1249 samples) compared to 2.7% in children (18/667 samples). Between January and June 2021, the overall positivity rate was 0.6% in adults (25/4118 samples) compared to 1.7% in children (41/2465 samples). Overall, the positivity rate was 2.5% (218/8863 samples; Fig. S2, Supplementary Material). Data on symptoms experienced by infected individuals were available for 59 cases. Over the full study period, 45 cases were symptomatic, while 14 remained asymptomatic, and children were not more often asymptomatic compared to adults (35.0% vs 16.2%, $p = .184$). Cycle threshold (Ct) values at diagnosis did not differ between children and adults in either time period, nor between time periods ($p = 1.000$ for week 1–15, $p = .299$ for week 16–40, $p = .655$ between periods; Fig. S3, Supplementary Material). Ct values at diagnosis also did not differ between symptomatic and asymptomatic individuals ($p = .592$).

3.2. Compliance to the study protocol

Among all participants, the median duration of participation was 38 weeks (IQR 25–40). Between September and December 2020, the median number of samples per participant was 12 (IQR 10–13), reflecting high adherence to the study protocol of one sample every week for 15 weeks. Adherence was in general a bit lower between January and June 2021 (week 16 to 40), with a median of 34 (IQR 21–39) samples per participant. The school was closed from October 30 until November 15, 2020, resulting in a drop in the number of samples collected during weeks 7 and 8. In addition, week 34 included Ascension Day (13

May 2021), resulting in a long weekend which may explain the drop in samples collected. Overall, adherence was higher among children and parents compared to school staff (Fig. 1). Of the 235 participants included from January 2021 onward, 88 (37.5%) dropped out of the study before the last week of sampling. Among those who did not complete the study, 18.3% had previously tested positive for SARS-CoV-2 during the study period and 17.2% had received a first vaccination dose before their last sample.

3.3. Sequence alignment and phylogenetic analysis

Forty WGS were available for 36 distinct cases (59% of all cases). Quality assessment of the raw sequence data identified no issues concerning frameshifts, premature stop codons, mixed sites, clusters of mutations, or excess private mutations. There were four individuals from whom two samples were sequenced. In the following analyses, the earliest sampled WGS for each confirmed case was used. The reasons for excluding the later sequenced samples were (i) identical sequence, (ii) only a partial genome due to high Ct sample, and (iii) a high proportion of ambiguous nucleotides. For all sequences, the region encoding the envelope protein was excluded due to alignment issues caused by missing nucleotides in 12 sequences. Mutations in the envelope protein have only been reported in the Beta (B.1.351) variant (Mohammad et al., 2021), which has not been observed in this study. The most frequently observed amino acid changes compared to the reference sequence were the P4715L mutation in the ORF1ab gene occurring in 100% of the sequences, and the D614G mutation in the spike glycoprotein occurring in 91.7% of the sequences (Fig. S4, Supplementary Material). A previous study has found these to be the most prevalent mutations in WGS sampled between December 2019 and September 2020 (i.e., wild type SARS-CoV-2) (Omotoso et al., 2021). The ORF8 Q27* knockout mutation was detected in the 7 sequences from the March cluster and is a known defining mutation for the Alpha VoC (Kistler et al., 2021). The best fitting substitution and RHAS model was a TN93 model with empirical base frequencies and no rate heterogeneity across sites (Tamura and Nei, 1993). The sequenced cases belong to four different pangolin lineages, i.e., B.1.221 (47.2%), B.1.160.28 (30.6%), B.1.389 (2.5%), and B.1.1.7 (19.4%), as shown in the time-scaled phylogenetic tree (Fig. 2).

3.4. Outbreak reconstruction

Trace plots indicated adequate convergence of the MCMC chain for both models (Figs. S5 and S6, Supplementary Material). Table S1 (Supplementary Material) shows the assumed prior distributions and posterior estimates with 95% credible intervals (CrI) for the parameters of the Outbreaker2 model under a baseline and sensitivity scenario. In the consensus transmission tree, defined as the tree with the modal posterior infector for each case, 27 direct transmission events (i.e. between two observed cases, without unobserved intermediate cases) occurred in school, 14 in households, and 15 were indirect transmission events (i.e. with unobserved intermediate cases). Of the direct transmission events, 12 (29.3%) occurred among children, 16 (39.0%) among adults, 10 (24.4%) from child to adult, and 3 (7.3%) from adult to child (Fig. 3). Based on the posterior distribution of ancestors, direct transmission was most likely to have occurred between children or from children to adults during the March outbreak when the Alpha VoC was circulating (Fig. S7, Supplementary Material). In October (case 1 to 38) and December (case 39 to 47) it was more likely to have unobserved intermediate cases (i.e., $\kappa > 1$) compared to March (case 48 to 59) when sampling was done three times a week (Fig. S8, Supplementary Material). The use of uninformative priors did not impact estimates of the model parameters (Table S1, Supplementary Material). Another sensitivity analysis assuming the generation and time-to-collection interval to be uniformly distributed over 10 days led

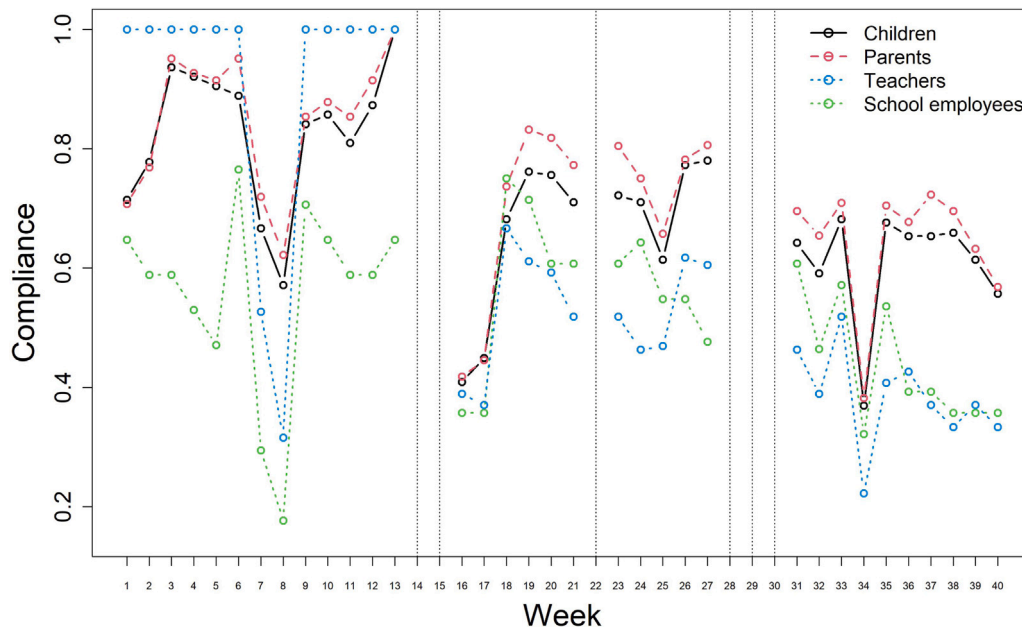


Fig. 1. Compliance to the study protocol over time by participant group. Compliance is defined as the number of samples obtained divided by the expected number of samples for each week. Dotted vertical lines indicate weeks of school holidays when no samples were collected.

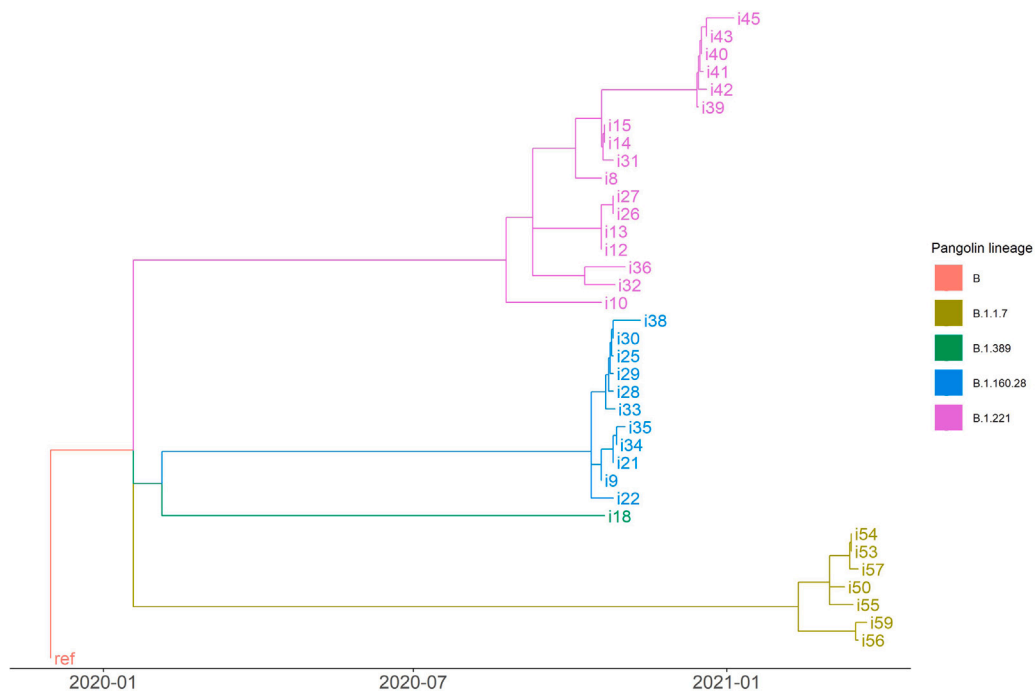


Fig. 2. Time-scaled maximum likelihood tree. The tree is rooted on the reference sequence (ref) which represents an isolate obtained early in the pandemic. Branches are coloured according to pangolin lineage.

to similar results, with a slightly higher estimate for the proportion of cases sampled.

Table S2 (Supplementary Material) shows the posterior estimates and 95% highest posterior density (HPD) intervals for the parameters of the SCOTTI model for each of three separately analysed time periods. When using SCOTTI, for 24 (66.7%) of the 36 cases included in this analysis, the highest probability of direct infection was linked to transmission in school. Especially for the cases linked to the October cluster, SCOTTI assigned only very low probability to observed infectors, with most cases assigned to have been infected by a non-sampled case. Because SCOTTI only uses the cases for which a genome sequence

is available, these non-sampled cases could in fact be the observed cases for which no sequence was available and that were assigned as maximum posterior ancestry by Outbreaker2.

Uncertainty related to inferred ancestry was quantified using the Shannon entropy of the posterior distribution of potential infectors for each case (Shannon, 1997), with a higher entropy indicating more uncertainty in ancestry assignment (Campbell et al., 2019). For the Outbreaker2 model, the average entropy among all cases was 1.43 when including and 1.60 without including sequence data. For SCOTTI, the average entropy was 1.46 among the 36 cases included in this analysis. Using Outbreaker2 without sequence data, most uncertainty

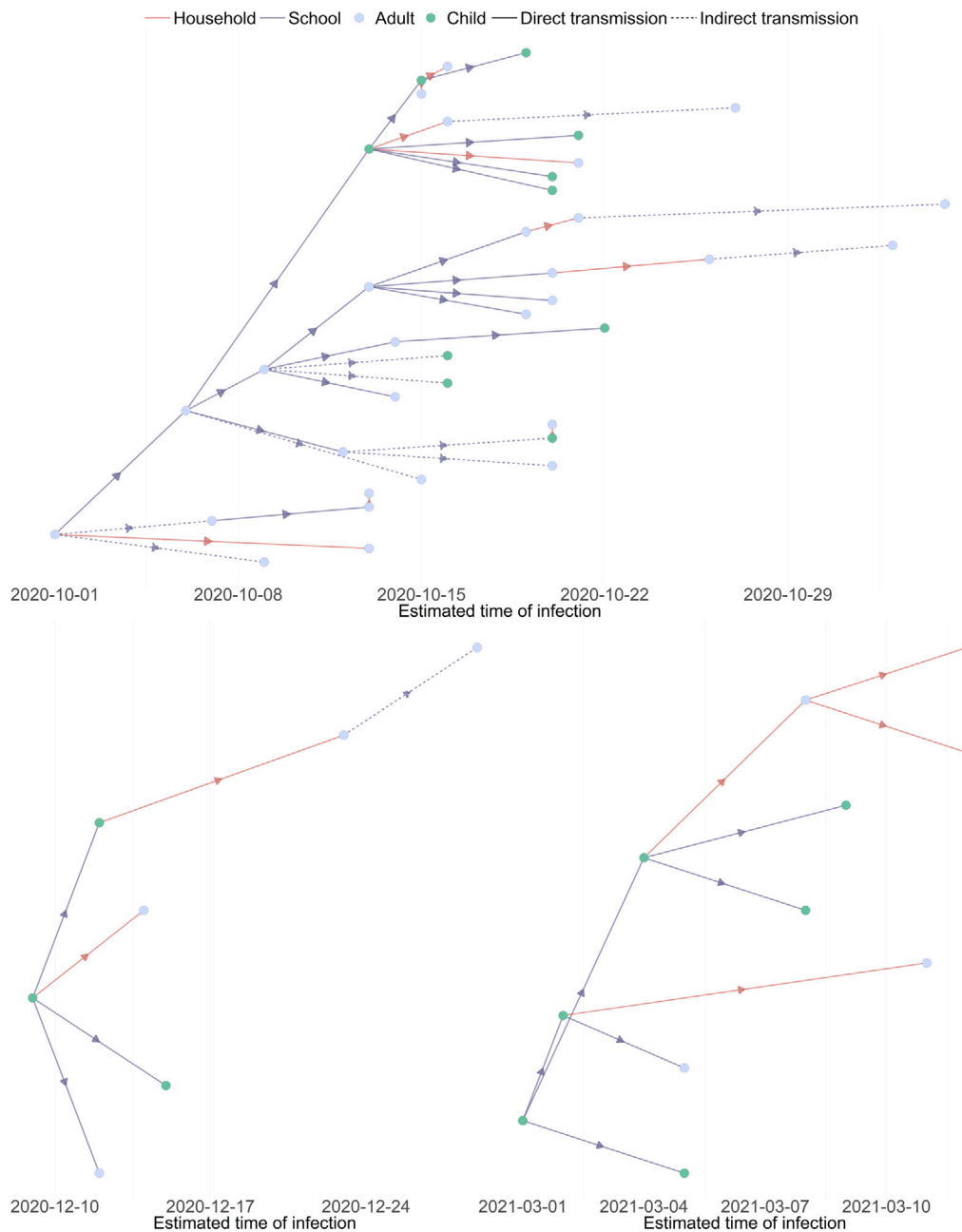


Fig. 3. Consensus transmission tree under the baseline scenario from the Outbreaker2 model. The consensus tree is defined as the tree with the modal posterior ancestor for each case. Indirect transmission (dashed line) means at least one unobserved intermediate case is present.

is observed for the cases in October (case 1 to 38), and this uncertainty is in general slightly reduced when including sequence data in the outbreak reconstruction. For most cases observed during the December and March outbreaks, there is more uncertainty in ancestry assignment when using SCOTTI compared to Outbreaker2 including sequence data (Fig. 4).

3.5. Simulation study

We found that the true positivity rate is underestimated during the first weeks of an outbreak. After some introductions to get an outbreak going, the deviation between true and observed positivity rates as well as the variation in this deviation decreases for an increasing sampling proportion. Repetitive testing at a frequency of twice instead of once weekly generally does not result in a big improvement in estimation of the positivity rate. In contrast, a strategy of only testing and isolating

symptomatic individuals results in a consistent underestimation of the true positivity rate (Fig. 5). Increasing the probability of symptomatic infection (to 65% for children and 85% for adults) or increasing the proportion of immune adults to 50% did not have an impact on these results (Fig. S9, Supplementary Material). Estimation of the weekly positivity rate in children was in general more accurate than the rate in adults (Fig. S10, Supplementary Material).

4. Discussion

In the present study, SARS-CoV-2 infection was equally present among adults and children, consistent with findings from previous studies (Meuris et al., 2021; Ulyte et al., 2021; Callies et al., 2022). We found no significant difference in the proportion of asymptomatic cases between children and adults during the second part of the study, while it was observed that children were more often asymptomatic

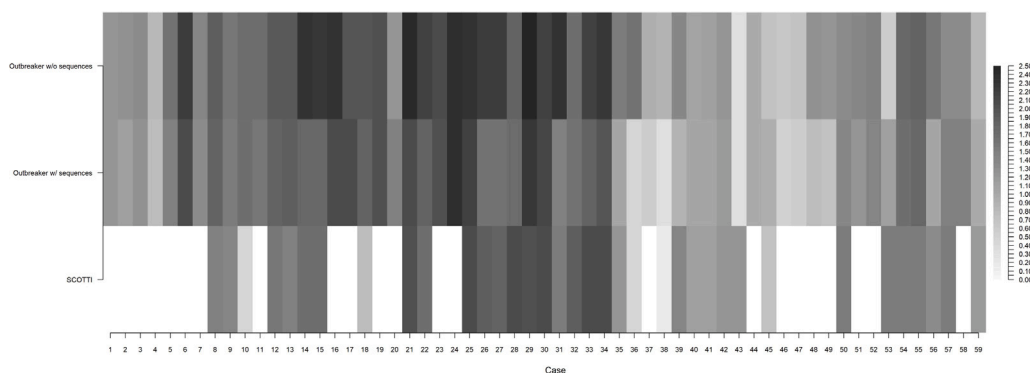


Fig. 4. Uncertainty in ancestry assignment for the different models. For each case, uncertainty is quantified as the Shannon entropy of the posterior distribution of potential infectors. Only 36 cases are included in the SCOTTI model. Higher values indicate a larger number of plausible infectors, hence higher uncertainty in ancestry assignment.

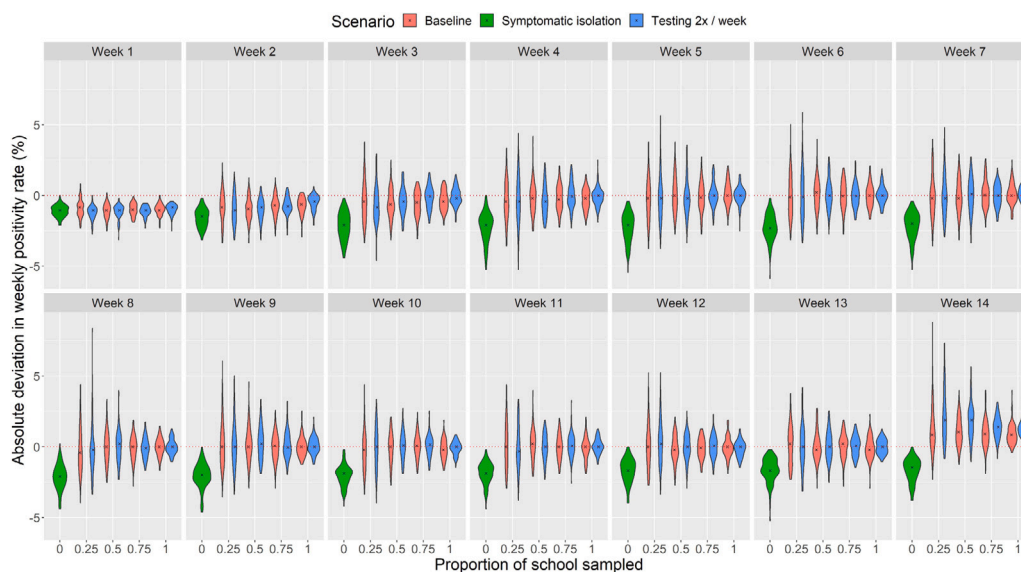


Fig. 5. Accuracy of estimated weekly positivity rates. Accuracy is defined as the absolute deviation (observed minus true positivity rate) in estimates of the weekly positivity rate (in %) for different proportions of the school sampled. Black crosses represent the median deviation among 100 simulated outbreaks. In the baseline scenario, testing was performed once a week.

during September–December 2020 (Meuris et al., 2021). This may be explained by circulation of the Alpha VoC in the cluster observed in March 2021, for which it has been shown that children have a higher probability of developing symptoms compared to wild-type SARS-CoV-2 (Gorgels et al., 2022). In line with a previous study that found no effect of age on Ct values after adjusting for the effects of self-reported symptom status and number of positive genes (Walker et al., 2021), we observed no difference in Ct values between children and adults. As a proxy for viral load, Ct values may inform an individual’s infectiousness (Trunfio et al., 2022), suggesting that children and adults may have been equally infectious in this study. From April 2021 onward, only three sporadic cases were observed. Vaccination of adults could have had a protective effect on the school community. As a lot of parents included in this study are healthcare workers, about half of them had been vaccinated relatively early on. However, only 14.1% of the vaccinated adults were teachers, and contact between parents and children that are not their own was assumed to have been limited during the study period. Even if all teachers would have been vaccinated, they only make up a small part of the school population and have lower contact rates with students compared to students amongst each other, hence vaccination of children may be needed to limit infections in schools and decrease the risk of importation at the community level (Colosi et al., 2022).

Because the two models used for outbreak reconstruction differ in their estimation procedure they can only be compared in a broad sense in terms of inference of the transmission network (Firestone et al., 2019). The main purpose of outbreak reconstruction in this study was to infer the extent of transmission within the school environment. Using SCOTTI, we found that for 66.7% of cases transmission most likely occurred at school. Similarly, using Outbreaker2 we estimated that 65.9% of direct infection occurred within the school and furthermore that more than half of direct transmission events most likely originated from a child. These results imply that children, as well as teachers, pose a risk of bringing infection into their households, from which it can spread further into the community. In line with this, a recent retrospective cohort study in students, school personnel, and their household members found that younger age groups had a high contribution to the spread of infection (Manica et al., 2022). Similarly, two recent studies in The Netherlands and Switzerland found evidence of large-scale transmission among children, school personnel, and introduction into their households during dominance of the Alpha VoC (Gorgels et al., 2022; Lorth et al., 2022). Uncertainty in outbreak reconstruction was lowest for the Outbreaker2 model that uses genomic sequence data in addition to epidemiological data. However, the average entropy for SCOTTI, which does not use information on contacts, and Outbreaker2 was similar, suggesting that the addition of contact data may not be very informative in a school setting.

From January to June 2021, we observed lower positivity rates compared to before January 2021, reflecting the lower background incidence in the Belgian population at that time. In addition, positivity rates were higher for children compared to adults from January to June 2021. In our simulation study, observed weekly positivity rates were found to be a good approximation to the true weekly positivity rate, especially in children, even when only a small proportion of the school is tested regularly. In practice, this can be used to evaluate and if necessary adapt mitigation measures in school. Symptomatic isolation has already been shown not to be very effective in reducing SARS-CoV-2 infections in schools (Colosi et al., 2022; Torneri et al., 2022), and we found that data obtained from such a strategy results in consistent underestimation of positivity rates. Further research is needed to investigate how the accuracy of outbreak reconstruction depends on which proportion of a school population is sampled and how many of these samples are sequenced.

Adherence to the screening protocol has been assumed crucial for a repetitive testing strategy to be effective (Colosi et al., 2022). We observed moderate-to-high adherence to either once or twice a week testing, with adherence in general being higher among children and parents compared to school staff. In contrast, there was also substantial loss to follow-up. Although this cannot be directly inferred from the available data, possible reasons for dropout and reduced adherence include fatigue, vaccination, stress generated by positive results, and previous infection of the individual or in their household leading to a sense of protection. A recent modelling study found that regardless of the level of compliance, repetitive testing always leads to a greater reduction in final outbreak size compared to reactive screening, although the reduction in final size increased with higher levels of compliance (Torneri et al., 2022). When social mixing between classes is high, as in primary schools, other classes may have already been affected before the first case is detected, making a reactive screening strategy less efficient (Liu et al., 2022). In contrast, with repetitive testing, more cases that would otherwise go unnoticed will be detected and isolation can be applied only to those cases during their infectious period, instead of quarantining the entire class.

This study has several limitations. Assessment of symptoms was retrospective and hence could be biased by the knowledge of having been infected. Especially in children, it may have been the case that very mild symptoms with a duration of less than one day have been reported. The Outbreaker2 model makes a number of simplifying assumptions (Campbell et al., 2019). Contacts are undated, hence the model does not consider that these contacts only could have led to transmission if they occurred during the infectious period of the infector. However, in a school setting, the same individuals will be in contact with each other almost daily. Possible within-host evolution is not accounted for, hence the genetic likelihood only depends on the number of transmission events separating two cases and not on time. In addition, the genetic likelihoods are assumed to be independent while in fact the genetic relatedness between two cases, A infecting B, is dependent on the infector of case A. SCOTTI overcomes some of these limitations by explicitly using the sequence data rather than only the pairwise genetic distances and allowing transmission to occur only within a specified exposure window. However, the assumption that transmission is equally likely between every pair of cases during their exposure window may be reasonable for transmission among children in the school environment but not for transmission between children and adults who are not their teacher or parent. Although it is possible to relax this assumption to account for known contact structure by changing the default migration model, such a configuration of SCOTTI would be very computationally demanding (De Maio et al., 2015). Furthermore, SCOTTI does not include the non-sequenced cases, thereby excluding a large part of the available information which is reflected in the high probabilities assigned to non-sampled cases being the most likely ancestor.

5. Conclusions

Despite these limitations, the present study shows that it is worthwhile to implement repetitive screening in a school setting. In addition to reducing infections, important insights can be obtained on the role of schools in transmission of infectious diseases such as the extent of transmission happening within the school environment. In particular, we found that the majority of SARS-CoV-2 transmission occurred in school, with importation to students' and staff members' households. Challenges to implementing such a strategy include the cost of PCR testing as well as quickly obtaining test results to ensure fast isolation of positive individuals in order to prevent onward transmission within as well as outside the school. Testing capacity can be quickly overwhelmed when case numbers are high. In periods of low incidence, pooled testing could be a more cost-effective approach (Libin et al., 2021). Alternatively, the use of antigen tests could be considered (Liu et al., 2022). Besides educational disadvantages, it is important to minimize SARS-CoV-2 infections in children in order to avoid possible long-term health disadvantages caused by conditions such as long-COVID or multisystem inflammatory syndrome, which can occur even after mild or asymptomatic disease (Lopez-Leon et al., 2022; Jiang et al., 2020; Pierce et al., 2022).

Ethics approval and consent to participate

This study was approved by the ethics committee of Liège University Hospital (number: 2020–241), the school directors, and the local authorities. Participants provided written informed consent. This study was conducted in accordance with international ethical standards (Declaration of Helsinki 1964).

Funding

The study was funded by Fondation Léon Fredericq and by the Liège University Hospital Research funds. P.J.K.L. gratefully acknowledges support from the Research Foundation Flanders (FWO), Belgium via postdoctoral fellowship 1242021N (P.J.K.L.). P.J.K.L. also acknowledges support from the Research council of the Vrije Universiteit Brussel (OZR-VUB), Belgium via grant number OZR3863BOF, and from the Flemish Government through the AI Research Program. G.D. is an FNRS postdoctoral clinical master specialist. P.L. acknowledges support by the European Research Council under the European Union's Horizon 2020 research and innovation programme (grant agreement number 725422 - ReservoirDOCS) and the Research Foundation – Flanders ('Fonds voor Wetenschappelijk Onderzoek – Vlaanderen', G066215N, G0D5117N and G0B9317N). N.H. and A.T. acknowledge funding from the European Union's Horizon 2020 research and innovation programme - project EpiPose (Grant agreement number 101003688). This project was supported by the VERDI project (101045989), funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the Health and Digital Executive Agency. Neither the European Union nor the granting authority can be held responsible for them.

CRedit authorship contribution statement

Cécile Kremer: Conceptualization, Methodology, Software, Formal analysis, Data curation, Writing – original draft, Visualization, Writing – review & editing. **Andrea Torneri:** Conceptualization, Methodology, Software, Writing – review & editing. **Pieter J.K. Libin:** Conceptualization, Methodology, Software, Writing – review & editing. **Cécile Meex:** Investigation. **Marie-Pierre Hayette:** Investigation. **Sébastien Bontems:** Investigation. **Keith Durkin:** Investigation, Data curation. **Maria Artesi:** Investigation, Data curation. **Vincent Bours:** Investigation, Data curation. **Philippe Lemey:** Writing – review & editing. **Gilles**

Darcis: Conceptualization, Investigation, Resources, Writing – review & editing, Supervision, Project administration, Funding acquisition. **Niel Hens:** Conceptualization, Methodology, Writing – review & editing, Supervision, Project administration, Funding acquisition. **Christelle Meuris:** Conceptualization, Investigation, Resources, Writing – review & editing, Supervision, Project administration, Funding acquisition.

Declaration of competing interest

Christelle Meuris received grants from Fondation on Léon Fredericq and FIRS during the conduct of the study. Niel Hens reports that the Universities of Antwerp and Hasselt obtained grants from several vaccine manufacturers for specific studies aimed at modelling the spread of infectious diseases outside the submitted work for which Niel Hens is the principal investigator. Niel Hens obtains no personal remuneration. All other authors declare no competing interests.

Data availability

Data and code are available on Github (<https://github.com/cecilekremer/school-recon>).

Acknowledgements

The computational resources and services used in this work were provided by the VSC (Flemish Supercomputer Center), funded by the Research Foundation Flanders (FWO) and the Flemish Government department EWI.

Appendix A. Supplementary data

Supplementary material related to this article can be found online at <https://doi.org/10.1016/j.epidem.2023.100701>.

References

- Abbas, M., Cori, A., Cordey, S., Laubscher, F., Robalo Nunes, T., Myall, A., Salamun, J., Huber, P., Zekry, D., Prendki, V., Iten, A., Vieux, L., Sauvan, V., Graf, C.E., Harbarth, S., 2022. Reconstruction of transmission chains of SARS-CoV-2 amidst multiple outbreaks in a geriatric acute-care hospital: a combined retrospective epidemiological and genomic study. In: Schiffer, J.T., Garrett, W.S., Opatowski, L., Pham, T. (Eds.), *ELife* (ISSN: 2050-084X) 11, e76854. <http://dx.doi.org/10.7554/eLife.76854>.
- Aksamentov, I., Roemer, C., Hodcroft, E.B., Neher, R.A., 2021. Nextclade: clade assignment, mutation calling and quality control for viral genomes. *J. Open Source Softw.* 6 (67), 3773. <http://dx.doi.org/10.21105/joss.03773>.
- Bouckaert, R., Heled, J., Kühnert, D., Vaughan, T., Wu, C.-H., Xie, D., Suchard, M.A., Rambaut, A., Drummond, A.J., 2014. BEAST 2: A software platform for Bayesian evolutionary analysis. *PLoS Comput. Biol.* 10 (4), e1003537.
- Buonanno, G., Stabile, L., Morawska, L., 2020. Estimation of airborne viral emission: Quanta emission rate of SARS-CoV-2 for infection risk assessment. *Environ. Int.* 141, 105794.
- Callies, M., Kabouche, I., Desombere, I., Merckx, J., Roelants, M., Vermeulen, M., Duysburgh, E., 2022. Measures for infection prevention and control of SARS-CoV-2 in Belgian schools between december 2020 and june 2021: A prospective cohort study. *MedRxiv* <http://dx.doi.org/10.1101/2022.04.12.22273722>.
- Campbell, F., Cori, A., Ferguson, N., Jombart, T., 2019. Bayesian inference of transmission chains using timing of symptoms, pathogen genomes and contact data. *PLoS Comput. Biol.* 15 (3), 1–20. <http://dx.doi.org/10.1371/journal.pcbi.1006930>.
- Campbell, F., Didelot, X., Fitzjohn, R., Ferguson, N., Cori, A., Jombart, T., 2018a. outbreaker2: A modular platform for outbreak reconstruction. *BMC Bioinformatics* 19 (11), 1–8.
- Campbell, F., Strang, C., Ferguson, N., Cori, A., Jombart, T., 2018b. When are pathogen genome sequences informative of transmission events? *PLoS Pathog.* 14 (2), e1006885.
- Chang, D., Mo, G., Yuan, X., Tao, Y., Peng, X., Wang, F.-S., Xie, L., Sharma, L., Dela Cruz, C.S., Qin, E., 2020. Time kinetics of viral clearance and resolution of symptoms in novel coronavirus infection. *Am. J. Respir. Crit. Care Med.* 201 (9), 1150–1152.
- Colosi, E., Bassignana, G., Contreras, D.A., Poirier, C., Boëlle, P.-Y., Cauchemez, S., Yazdanpanah, Y., Lina, B., Fontanet, A., Barrat, A., et al., 2022. Screening and vaccination against COVID-19 to minimise school closure: a modelling study. *Lancet Infect. Dis.*
- Davies, N.G., Klepac, P., Liu, Y., Prem, K., Jit, M., Eggo, R.M., 2020. Age-dependent effects in the transmission and control of COVID-19 epidemics. *Nature Med.* 26 (8), 1205–1211.
- De Maio, N., Wu, C.-H., O'Reilly, K.M., Wilson, D., 2015. New routes to phylogeography: A Bayesian structured coalescent approximation. *PLoS Genet.* 11 (8), e1005421.
- De Maio, N., Wu, C.-H., Wilson, D.J., 2016. SCOTTI: efficient reconstruction of transmission within outbreaks with the structured coalescent. *PLoS Comput. Biol.* 12 (9), e1005130.
- Departement Onderwijs en Vorming, 2023. Vlaams onderwijs in cijfers 2021–2022. <https://www.vlaanderen.be/publicaties/vlaams-onderwijs-in-cijfers>.
- Firestone, S.M., Hayama, Y., Bradhurst, R., Yamamoto, T., Tsutsui, T., Stevenson, M.A., 2019. Reconstructing foot-and-mouth disease outbreaks: A methods comparison of transmission network models. *Sci. Rep.* 9 (1), 1–12.
- Freed, N.E., Vlkova, M., Faisal, M.B., Silander, O.K., 2020. Rapid and inexpensive whole-genome sequencing of SARS-CoV-2 using 1200 bp tiled amplicons and oxford nanopore rapid barcoding. *Biol. Methods Protoc.* 5, a014. <http://dx.doi.org/10.1093/biomethods/bpaa014>.
- Gorgels, K.M., van Alphen, L.B., van der Veer, B.M., Hackert, V.H., Hensels, A.Y., den Heijer, C.D., Dingemans, J., Savelkoul, P.H., Hoebe, C.J., 2022. Increased transmissibility of SARS-CoV-2 alpha variant (B. 1.1. 7) in children: three large primary school outbreaks revealed by whole genome sequencing in the netherlands. *BMC Infect. Dis.* 22 (1), 1–10.
- He, X., Lau, E.H., Wu, P., Deng, X., Wang, J., Hao, X., Lau, Y.C., Wong, J.Y., Guan, Y., Tan, X., et al., 2020. Temporal dynamics in viral shedding and transmissibility of COVID-19. *Nature Med.* 26 (5), 672–675.
- Hoang, D.T., Chernomor, O., von Haeseler, A., Minh, B.Q., Vinh, L.S., 2017. UFBoot2: Improving the Ultrafast Bootstrap Approximation. *Mol. Biol. Evol.* (ISSN: 0737-4038) 35 (2), 518–522. <http://dx.doi.org/10.1093/molbev/msx281>, arXiv:https://academic.oup.com/mbe/article-pdf/35/2/518/24367824/msx281.pdf.
- Hoang, T., Coletti, P., Melegaro, A., Wallinga, J., Grijalva, C.G., Edmunds, J.W., Beutels, P., Hens, N., 2019. A systematic review of social contact surveys to inform transmission models of close-contact infections. *Epidemiology (Cambridge, Mass.)* 30 (5), 723.
- Hufsky, F., Lamkiewicz, K., Almeida, A., Aouacheria, A., Arighi, C., Bateman, A., Baumbach, J., Beerenwinkel, N., Brandt, C., Cacciabue, M., Chuguransky, S., Drechsel, O., Finn, R.D., Fritz, A., Fuchs, S., Hattab, G., Hauschild, A.-C., Heider, D., Hoffmann, M., Hölzer, M., Hoops, S., Kaderali, L., Kalvari, I., von Kleist, M., Kmiecinski, R., Kühnert, D., Lasso, G., Libin, P., List, M., Löchel, H.F., Martin, M.J., Martin, R., Matschinske, J., McHardy, A.C., Mendes, P., Mistry, J., Navratil, V., Nawrocki, E.P., O'Toole, A.N., Ontiveros-Palacios, N., Petrov, A.I., Rangel-Pineros, G., Redaschi, N., Reimering, S., Reinert, K., Reyes, A., Richardson, L., Robertson, D.L., Sadegh, S., Singer, J.B., Theys, K., Upton, C., Welzel, M., Williams, L., Marz, M., 2020. Computational strategies to combat COVID-19: useful tools to accelerate SARS-CoV-2 and coronavirus research. *Brief. Bioinform.* (ISSN: 1477-4054) 22 (2), 642–663. <http://dx.doi.org/10.1093/bib/bbaa232>, arXiv:https://academic.oup.com/bib/article-pdf/22/2/642/36654876/bbaa232.pdf.
- Jiang, L., Tang, K., Levin, M., Irfan, O., Morris, S.K., Wilson, K., Klein, J.D., Bhutta, Z.A., 2020. COVID-19 and multisystem inflammatory syndrome in children and adolescents. *Lancet Infect. Dis.* 20 (11), e276–e288.
- Kalyaanamoorthy, S., Minh, B.Q., Wong, T.K.F., von Haeseler, A., Jermini, L.S., 2017. ModelFinder: fast model selection for accurate phylogenetic estimates. *Nature Methods* <http://dx.doi.org/10.1038/nmeth.4285>, URL <https://pubmed.ncbi.nlm.nih.gov/28481363/>.
- Kistler, K.E., Huddleston, J., Bedford, T., 2021. Rapid and parallel adaptive mutations in spike S1 drive clade success in SARS-CoV-2. *BioRxiv* <http://dx.doi.org/10.1101/2021.09.11.459844>, arXiv:https://www.biorxiv.org/content/early/2021/10/15/2021.09.11.459844.full.pdf, URL <https://www.biorxiv.org/content/early/2021/10/15/2021.09.11.459844>.
- Lessler, J., Grabowski, M.K., Grantz, K.H., Badillo-Goicoechea, E., Metcalf, C.J.E., Lupton-Smith, C., Azman, A.S., Stuart, E.A., 2021. Household COVID-19 risk and in-person schooling. *Science* 372 (6546), 1092–1097.
- Libin, P.J.K., Deforche, K., Abecasis, A.B., Theys, K., 2018. VIRULIGN: fast codon-correct alignment and annotation of viral genomes. *Bioinformatics* (ISSN: 1367-4803) 35 (10), 1763–1765. <http://dx.doi.org/10.1093/bioinformatics/bty851>, arXiv:https://academic.oup.com/bioinformatics/article-pdf/35/10/1763/28604627/bty851.pdf.
- Libin, P.J., Willem, L., Verstraeten, T., Torneri, A., Vanderlocht, J., Hens, N., 2021. Assessing the feasibility and effectiveness of household-pooled universal testing to control COVID-19 epidemics. *PLoS Comput. Biol.* 17 (3), e1008688.
- Liu, Q.-H., Zhang, J., Peng, C., Litvinova, M., Huang, S., Poletti, P., Trentini, F., Guzzetta, G., Marziano, V., Zhou, T., et al., 2022. Model-based evaluation of alternative reactive class closure strategies against COVID-19. *Nature Commun.* 13 (1), 1–10.
- Lopez-Leon, S., Wegman-Ostrosky, T., del Valle, N.C.A., Perelman, C., Sepulveda, R., Rebolledo, P.A., Cuapio, A., Villapol, S., 2022. Long COVID in children and adolescents: A systematic review and meta-analyses. *MedRxiv*.
- Lorthe, E., Bellon, M., Michielin, G., Berthelot, J., Zaballa, M.-E., Pennacchio, F., Bekliz, M., Laubscher, F., Arefi, F., Perez-Saez, J., et al., 2022. Epidemiological, virological and serological investigation of a SARS-CoV-2 outbreak (alpha variant) in a primary school: A prospective longitudinal study. *Plos One* 17 (8), e0272663.

- Manica, M., Poletti, P., Deandrea, S., Mosconi, G., Ancarani, C., Lodola, S., Guzzetta, G., d'Andrea, V., Marziano, V., Zardini, A., Trentini, F., Odone, A., Tirani, M., Ajelli, M., Merler, S., 2022. Estimating SARS-CoV-2 transmission in educational settings: A retrospective cohort study. *Influenza Other Respiratory Viruses* 1–7. <http://dx.doi.org/10.1111/irv.13049>.
- Mensah, A.A., Sinnathamby, M., Zaidi, A., Coughlan, L., Simmons, R., Ismail, S.A., Ramsay, M.E., Saliba, V., Ladhani, S.N., 2021. SARS-CoV-2 infections in children following the full re-opening of schools and the impact of national lockdown: prospective, national observational cohort surveillance, July–December 2020, England. *J. Infect.* 82 (4), 67–74.
- Meuris, C., Kremer, C., Geerinck, A., Locquet, M., Bruyère, O., Defêche, J., Meex, C., Hayette, M.-P., Duchene, L., Dellot, P., Azarar, S., Maréchal, N., Sauvage, A.-S., Fripiat, F., Giot, J.-B., Léonard, P., Fombellida, K., Moutschen, M., Durkin, K., Artesi, M., Bours, V., Faes, C., Hens, N., Darcis, G., 2021. Transmission of SARS-CoV-2 After COVID-19 Screening and Mitigation Measures for Primary School Children Attending School in Liège, Belgium. *JAMA Netw. Open* (ISSN: 2574-3805) 4 (10), e2128757. <http://dx.doi.org/10.1001/jamanetworkopen.2021.28757>, arXiv:https://jamanetwork.com/journals/jamanetworkopen/articlepdf/2784812/meuris_2021_oi_210838_1633035195.09961.pdf.
- Mohammad, T., Choudhury, A., Habib, I., Asrani, P., Mathur, Y., Umair, M., Anjum, F., Shafiq, A., Yadav, D.K., Hassan, M., et al., 2021. Genomic variations in the structural proteins of SARS-CoV-2 and their deleterious impact on pathogenesis: A comparative genomics approach. *Front. Cell. Infect. Microbiol.* 951.
- Nguyen, L.-T., Schmidt, H.A., von Haeseler, A., Minh, B.Q., 2014. IQ-TREE: A Fast and Effective Stochastic Algorithm for Estimating Maximum-Likelihood Phylogenies. *Mol. Biol. Evol.* (ISSN: 0737-4038) 32 (1), 268–274. <http://dx.doi.org/10.1093/molbev/msu300>, arXiv:<https://academic.oup.com/mbe/article-pdf/32/1/268/13171186/msu300.pdf>.
- Omotoso, O.E., Olugbami, J.O., Gbadegesin, M.A., 2021. Assessment of intercontinent mutation hotspots and conserved domains within SARS-CoV-2 genome. *Infect. Genet. Evol.* 96, 105097.
- Pierce, C.A., Herold, K.C., Herold, B.C., Chou, J., Randolph, A., Kane, B., McFarland, S., Gurdasani, D., Pagel, C., Hotez, P., et al., 2022. COVID-19 and children. *Science* 377 (6611), 1144–1149.
- Poletti, P., Tirani, M., Cereda, D., Trentini, F., Guzzetta, G., Sabatino, G., Marziano, V., Castrofino, A., Grosso, F., Del Castillo, G., et al., 2021. Association of age with likelihood of developing symptoms and critical disease among close contacts exposed to patients with confirmed SARS-CoV-2 infection in Italy. *JAMA Netw. Open* 4 (3), e211085.
- Sagulenko, P., Puller, V., Neher, R.A., 2018. TreeTime: Maximum-likelihood phylodynamic analysis. *Virus Evol.* (ISSN: 2057-1577) 4 (1), <http://dx.doi.org/10.1093/ve/vex042>, vex042, arXiv:<https://academic.oup.com/ve/article-pdf/4/1/vex042/23361340/vex042.pdf>.
- Sciensano, 2021. COVID-19 wekelijks epidemiologisch bulletin (29 oktober 2021). https://covid-19.sciensano.be/sites/default/files/Covid19/COVID-19_Weekly%20report_20211029%20-%20NL.pdf.
- Sciensano, 2022. Prevalence and incidence of antibodies against SARS-CoV-2 in children and school staff an observational sero-prevalence prospective cohort study. Accessed June 24, 2022, https://www.sciensano.be/sites/default/files/report_seroprev_sars-cov-2_schools_t4_oct2021.pdf.
- Shannon, C.E., 1997. The mathematical theory of communication. 1963. *MD Comput.: Comput. Med. Pract.* 14 (4), 306–317.
- Sinha, I.P., Harwood, R., Semple, M.G., Hawcutt, D.B., Thursfield, R., Narayan, O., Kenny, S.E., Viner, R., Hewer, S.L., Southern, K.W., 2020. COVID-19 infection in children. *Lancet Respiratory Med.* 8 (5), 446–447.
- Southall, E., Holmes, A., Hill, E.M., Atkins, B.D., Leng, T., Thompson, R.N., Dyson, L., Keeling, M.J., Tildesley, M.J., 2021. An analysis of school absences in England during the COVID-19 pandemic. *BMC Med.* 19 (1), 1–14.
- Suchard, M.A., Lemey, P., Baele, G., Ayres, D.L., Drummond, A.J., Rambaut, A., 2018. Bayesian phylogenetic and phylodynamic data integration using BEAST 1.10. *Virus Evol.* 4 (1), vey016.
- Sun, K., Wang, W., Gao, L., Wang, Y., Luo, K., Ren, L., Zhan, Z., Chen, X., Zhao, S., Huang, Y., et al., 2021. Transmission heterogeneities, kinetics, and controllability of SARS-CoV-2. *Science* 371 (6526), eabe2424.
- Tamura, K., Nei, M., 1993. Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees. *Mol. Biol. Evol.* (ISSN: 0737-4038) 10 (3), 512–526. <http://dx.doi.org/10.1093/oxfordjournals.molbev.a040023>, arXiv:<https://academic.oup.com/mbe/article-pdf/10/3/512/11168171/2TAMU.PDF>.
- Torneri, A., Libin, P., Vanderlocht, J., Vandamme, A.-M., Neyts, J., Hens, N., 2020. A prospect on the use of antiviral drugs to control local outbreaks of COVID-19. *BMC Med.* 18 (1), 1–9.
- Torneri, A., Willem, L., Colizza, V., Kremer, C., Meuris, C., Darcis, G., Hens, N., Libin, P.J., 2022. Controlling SARS-CoV-2 in schools using repetitive testing strategies. *Elife* 11, e75593.
- Trunfio, M., Richiardi, L., Alladio, F., Staffilano, E., Longo, B., Venuti, F., Ghisetti, V., Burdino, E., Bonors, S., Vineis, P., et al., 2022. Determinants of SARS-CoV-2 contagiousness in household contacts of symptomatic adult index cases. *Front. Microbiol.* 694.
- Ulyte, A., Radtke, T., Abela, I.A., Haile, S.R., Ammann, P., Berger, C., Trkola, A., Fehr, J., Puhon, M.A., Kriemler, S., 2021. Evolution of SARS-CoV-2 seroprevalence and clusters in school children from June 2020 to April 2021: prospective cohort study Ciao Corona. *Swiss Med. Wkly.* 151, w30092.
- Vlachos, J., Hertegård, E., Svaleryd, H.B., 2021. The effects of school closures on SARS-CoV-2 among parents and teachers. *Proc. Natl. Acad. Sci.* 118 (9).
- Walker, A.S., Pritchard, E., House, T., Robotham, J.V., Birrell, P.J., Bell, I., Bell, J.I., Newton, J.N., Farrar, J., Diamond, I., et al., 2021. Ct threshold values, a proxy for viral load in community SARS-CoV-2 cases, demonstrate wide variation across populations and over time. *Elife* 10.
- Wang, S., Xu, X., Wei, C., Li, S., Zhao, J., Zheng, Y., Liu, X., Zeng, X., Yuan, W., Peng, S., 2022. Molecular evolutionary characteristics of SARS-CoV-2 emerging in the United States. *J. Med. Virol.* 94 (1), 310–317.
- Zhou, R., Li, F., Chen, F., Liu, H., Zheng, J., Lei, C., Wu, X., 2020. Viral dynamics in asymptomatic patients with COVID-19. *Int. J. Infect. Dis.* 96, 288–290.
- Zimmerman, K.O., Akinboyo, I.C., Brookhart, M.A., Boutzoukas, A.E., McGann, K.A., Smith, M.J., Maradiaga Panayotti, G., Armstrong, S.C., Bristow, H., Parker, D., et al., 2021. Incidence and secondary transmission of SARS-CoV-2 infections in schools. *Pediatrics* 147 (4).
- Zou, L., Ruan, F., Huang, M., Liang, L., Huang, H., Hong, Z., Yu, J., Kang, M., Song, Y., Xia, J., et al., 2020. SARS-CoV-2 viral load in upper respiratory specimens of infected patients. *N. Engl. J. Med.* 382 (12), 1177–1179.