



OPEN

Unravelling the effect of New Year's Eve celebrations on SARS-CoV-2 transmission

Caspar Geenen^{1✉}, Jonathan Thibaut¹, Lies Laenen^{1,2}, Joren Raymenants¹, Lize Cuyper^{1,2}, Piet Maes³, Simon Dellicour^{3,4} & Emmanuel André^{1,2}

Public holidays have been associated with SARS-CoV-2 incidence surges, although a firm link remains to be established. This association is sometimes attributed to events where transmissions occur at a disproportionately high rate, known as superspreading events. Here, we describe a sudden surge in new cases with the Omicron BA.1 strain amongst higher education students in Belgium. Contact tracers classed most of these cases as likely or possibly infected on New Year's Eve, indicating a direct trigger by New Year celebrations. Using a combination of contact tracing and phylogenetic data, we show the limited role of superspreading events in this surge. Finally, the numerous simultaneous transmissions allowed a unique opportunity to determine the distribution of incubation periods of the Omicron strain. Overall, our results indicate that, even under social restrictions, a surge in transmissibility of SARS-CoV-2 can occur when holiday celebrations result in small social gatherings attended simultaneously and communitywide.

Surges in COVID-19 case numbers have been observed following public holidays and celebrations, although causal evidence remains limited^{1–7}. Modelling studies and empirical data have implicated superspreading events (SSE) as major contributors to the spread of SARS-CoV-2 and other pathogens^{8–12}. A SSE is an event where a disproportionately high number of transmissions take place. The exact role of SSEs in holiday-triggered SARS-CoV-2 incidence surges remains unclear.

Here, we investigate the effect of New Year's Eve celebrations in Belgium on the incidence of SARS-CoV-2, specifically within a local community of higher education students. Anticipating a rise in cases due to the emerging B.1.1.529 lineage – corresponding to the Omicron variant of concern – Belgian authorities implemented increasingly stringent restrictions on social interactions in November and December 2021. By the end of December, indoor mass gatherings were prohibited, face masks were required in public indoor areas and working from home was mandatory in many professional sectors. Schools were closed for the winter holidays from 25 December 2021 until 9 January 2022. Christmas and New Year celebrations were essentially limited to private gatherings in residential locations, although no limit was imposed on the number of attendees. Despite these measures, a surge in COVID-19 cases was detected after New Year's Eve 2022 (NYE) by our targeted test and trace program for higher education students in the city of Leuven¹³.

Estimates of the median and mean incubation periods have been reported for several variants of concern and are generally lower than for the ancestral SARS-CoV-2 strain¹⁴. However, accurate assessment of the shape of the incubation period distribution is, in practice, much harder, because it requires a high level of certainty about the moment of inoculation for a sufficiently large sample of cases¹⁵. Contact tracing can sometimes provide such information, by tracing individuals exposed to an infectious case (forward contact tracing), or to the suspected source case or event (backward contact tracing or source investigation). In this way, Tanaka et al. previously observed a mean incubation period of 3.03 days (SD: 1.35) for the Omicron BA.1 lineage¹⁶.

Using a combination of phylogenetic and contact tracing data, we aimed to investigate the relationship between NYE celebrations and Omicron case numbers in a population of higher education students in the Belgian city of Leuven. We also use the resulting dataset to determine the distribution of incubation periods of the Omicron BA.1 strain in this population.

¹Department of Microbiology, Immunology and Transplantation, Laboratory of Clinical Microbiology, KU Leuven, Herestraat 49, 3000 Leuven, Belgium. ²Department of Laboratory Medicine, National Reference Centre for Respiratory Pathogens, University Hospitals Leuven, Leuven, Belgium. ³Rega Institute, Department of Microbiology, Immunology and Transplantation, Laboratory of Clinical and Epidemiological Virology, KU Leuven, Leuven, Belgium. ⁴Spatial Epidemiology Lab, Université Libre de Bruxelles, Brussels, Belgium. ✉email: caspar.geenen@kuleuven.be

Results

Epidemic curve

We focused our analyses on the Omicron BA.1 strain, which became dominant over the Delta strain amongst students in Leuven in mid-December 2021 (Supplementary Fig. 1). Vaccination rates in this population were above 90% during the study period, with a mean of three months elapsed since the last dose (Supplementary Fig. 5).

Our test and trace program for the Leuven student population detected a short-lived surge in cases at the start of 2022, after winter holiday celebrations. Although the surge was associated with a rise in the number of PCR tests performed at the test centre, the test positivity rate did not decrease (Supplementary Fig. 1). In a target population of around 50,000 higher education students, 2,505 underwent a test in the first two weeks of the new year, of which 424 (16.9%) had a positive result. Twelve cases were excluded because the S-gene target of the PCR assay was positive and 35 because the signal was too weak (C_q value for N-gene target above 30) to confirm or exclude S-gene target failure. Another 13 were excluded because at the time of sampling they already had a recent diagnosis of COVID-19. Of the remaining cases, 286 (79%) were successfully interviewed by the contact tracing team and grouped according to the criteria described in the *Methods* section: possibly infected on NYE (36 cases, 13%), likely infected on NYE (54 cases, 19%), or other/unknown source (196 cases, 69%). Symptomatic cases were plotted by onset date to construct the epidemic curve (Fig. 1).

At the peak of the epidemic curve on 3 January 2022, those likely and possibly infected on NYE accounted for 59% of all symptomatic cases (26 cases out of 44), of which 65% (17 cases) were classed as likely infected on NYE.

Effective reproduction number

We estimated daily values of the reproduction number (R_{eff}) from the epidemic curve, using the method by Cori et al. with a one-day sliding window (Fig. 1, panel c)¹⁷. R_{eff} reached a peak of 3.6 (95% CI: 2.6–4.9) two days after the suspected peak in transmissions on NYE. This is consistent with a lag inherent to this method, approximately equal to the incubation period¹⁷. The peak and the lower bound of its credible interval were confirmed as significant outliers using a Grubbs test ($p=0.00001$ and $p=0.015$, respectively).

With the same method to estimate R_{eff} but using only case numbers before NYE as input, we estimated a forecast for the subsequent 7 days (Fig. 1, panel b). Actual case numbers rose considerably above the 95% prediction interval on 2 and 3 January, but not when excluding cases likely or possibly infected on NYE.

Case characteristics

The mean age of the 90 cases possibly or likely infected on NYE was 21.7 years (SD: 2.2 days) and 51% of cases were female. The vaccination rate amongst these cases was 95% (82 of 86 cases; missing data: 4 cases) and, for vaccinated individuals, the mean number of months elapsed since the last dose was 3.3 (95% CI: 2.9–3.7). The mean follow-up period of the 15 asymptomatic cases was 4.5 days after sampling of the positive test (range: one to nine days). Some cases (26 cases, 29%) attended a shared NYE event with one or more other cases possibly or likely infected on NYE. However, no single event was identified linking more than four cases.

Before our phylogenetic analysis, we used information from the contact tracing interviews to assess the role of superspreading events. 17 cases (19%) declined to share the total number of attendees at their NYE event. Out of the remaining cases, 30 (59%) reported being infected at an event with fewer than ten attendees. That means a superspreading event could be excluded, based on our definition of nine or more transmissions (see the *Methods* section for more detail). We also asked each case directly about known test results of other attendees, which included persons outside the region and scope of our contact tracing efforts. Negative test results thus excluded a superspreading source event for 48 cases (66%), whereas twelve cases (16%) originating at a SSE were infected at five separate events. Thus, the share of transmissions originating from a SSE was situated in the range of 16 to 34%.

Phylogenetic analysis

We performed a phylogenetic analysis to further clarify whether a small number of SSEs could have led to the observed surge in infections.

Out of the 90 cases possibly or likely infected on NYE, viral genomic sequencing was not attempted for five due to a weak positive PCR result. Whole genome sequencing failed for two samples, resulting in 83 successfully determined genomes (92%). A time-scaled phylogeny was constructed based on (a) these 83 sequences, (b) all 7,933 Belgian sequences belonging to the Omicron strain available on GISAID, collected between 1 November 2021 and 20 January 2022 and analysed for the purpose of baseline surveillance, and (c) a subset of 753 Omicron sequences collected during the same period in other European countries (see the *Methods* section for more detail)¹⁸.

We excluded 397 sequences, including 3 obtained from NYE samples, because they did not belong to the BA.1 lineage. Another 254 sequences, including one NYE sample, were excluded because the length was shorter than 27,000 bases. The geographic spread of all included samples is shown in Supplementary Fig. 4. The most common Pango lineages were BA.1.1 (20 cases), BA.1 (18 cases), BA.1.17.2 (13 cases), BA.1.14 (10 cases), and BA.1.17 (8 cases). Other lineages are shown in Supplementary Table 5.

In the resulting phylogenetic tree of 8,118 sequences, the NYE sequences spanned the range of strains circulating on a national scale (Fig. 2). Out of all 79 included sequences linked to NYE, 29 (37%) clustered together with one to three other NYE sequences, when using the cov2clusters clustering method¹⁹. However, most appeared isolated from other NYE sequences in the overall tree. No more than four NYE sequences could be grouped with 90% probability in a single cluster.

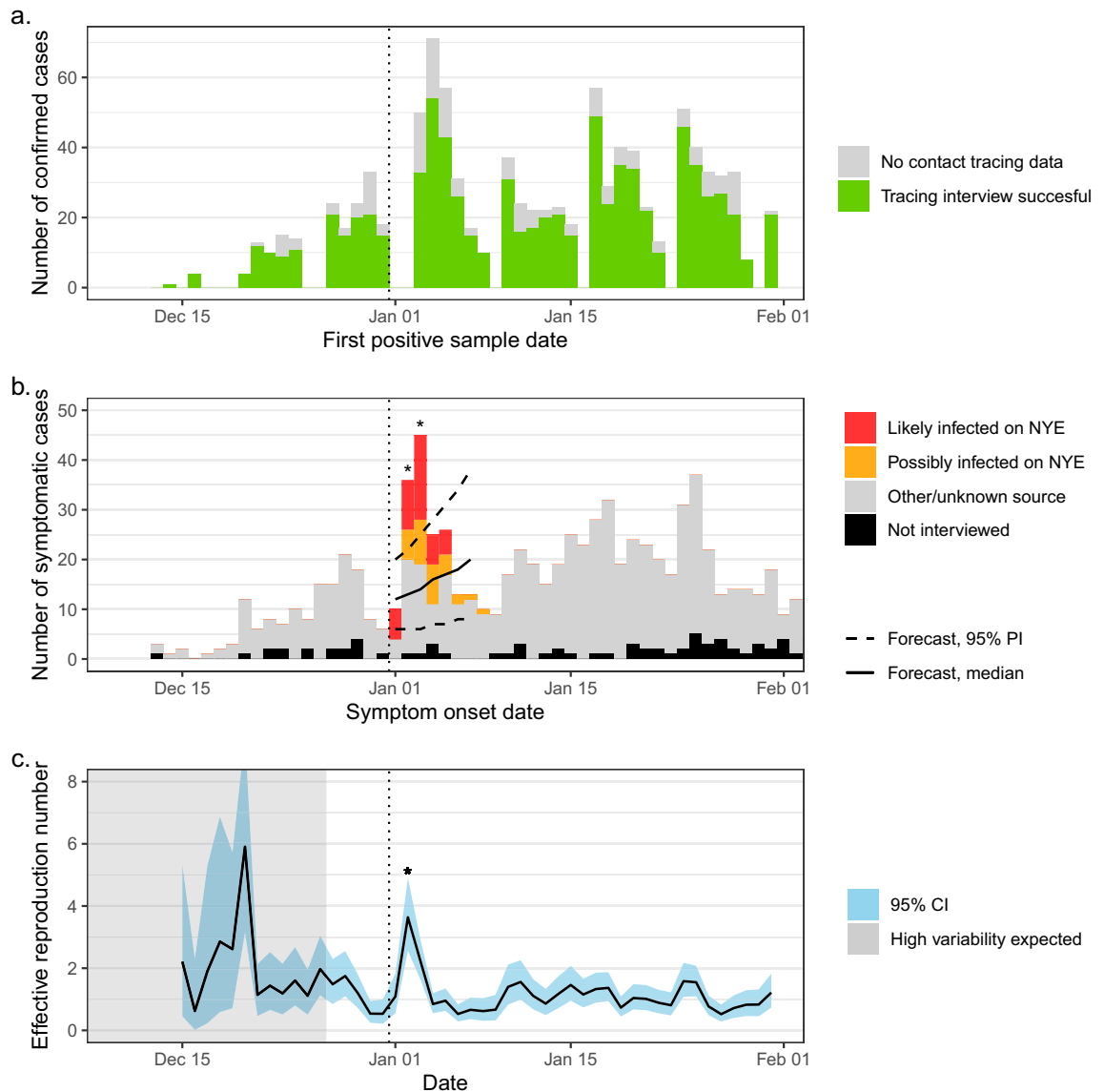


Figure 1. Evolution of Omicron BA.1 case numbers amongst higher education students at KU Leuven, Belgium. Panel a shows daily totals of Omicron BA.1 cases detected by the university test centre, stratified by sampling date. Successfully interviewed cases are shown in green. In panel b, symptomatic cases are shown by self-reported symptom onset date. Classifications of “possibly infected on NYE” (New Year’s Eve 2022) in yellow and “likely infected on NYE” in red were assessed by a contact tracer as described in the Methods section. Solid and dashed black lines indicate the median and 95% prediction interval (PI), respectively, of a 7-day forecast based only on case numbers before NYE. Asterisks indicate observations outside the 95% PI of the forecast. Panel c shows one-day sliding window estimates of the effective reproduction number (R_{eff}) for the student population in Leuven, derived from the epidemic curve in panel b using EpiEstim¹⁷. The blue band indicates the 95% credible interval (CI), while the grey area indicates the period when incidence was too low to reach an expected coefficient of variation above 0.3. An asterisk indicates that a significant outlier was detected using a one-sided Grubbs test. Vertical dotted black lines indicate NYE. In each panel, the graphs show only PCR-confirmed cases with a negative S-gene target on the TaqPath COVID-19 assay, corresponding to the Omicron BA.1 strain.

Incubation period

An assumption that source investigation accurately pinpointed the moment of infection on NYE allowed us to assess the distribution of incubation periods in this age group. We assumed that the infection occurred at midnight on NYE, which is the midpoint of our definition of NYE events and culturally is also the moment of closest interaction, as attendees hug and kiss to celebrate the new year. The median incubation period of Omicron was 2 to 3 days, when considering all symptomatic cases possibly or likely infected on NYE (mean: 2.8 and 2.4 days, N : 75 and 44, respectively). After fitting several distributions, we attained the best fit for possible transmissions on NYE with a Weibull distribution (shape and scale parameters $k=2.08$ and $\lambda=3.11$), with a median of 2.6 days (5th and 95th percentile 0.7 and 5.3 days, respectively) (Fig. 3).

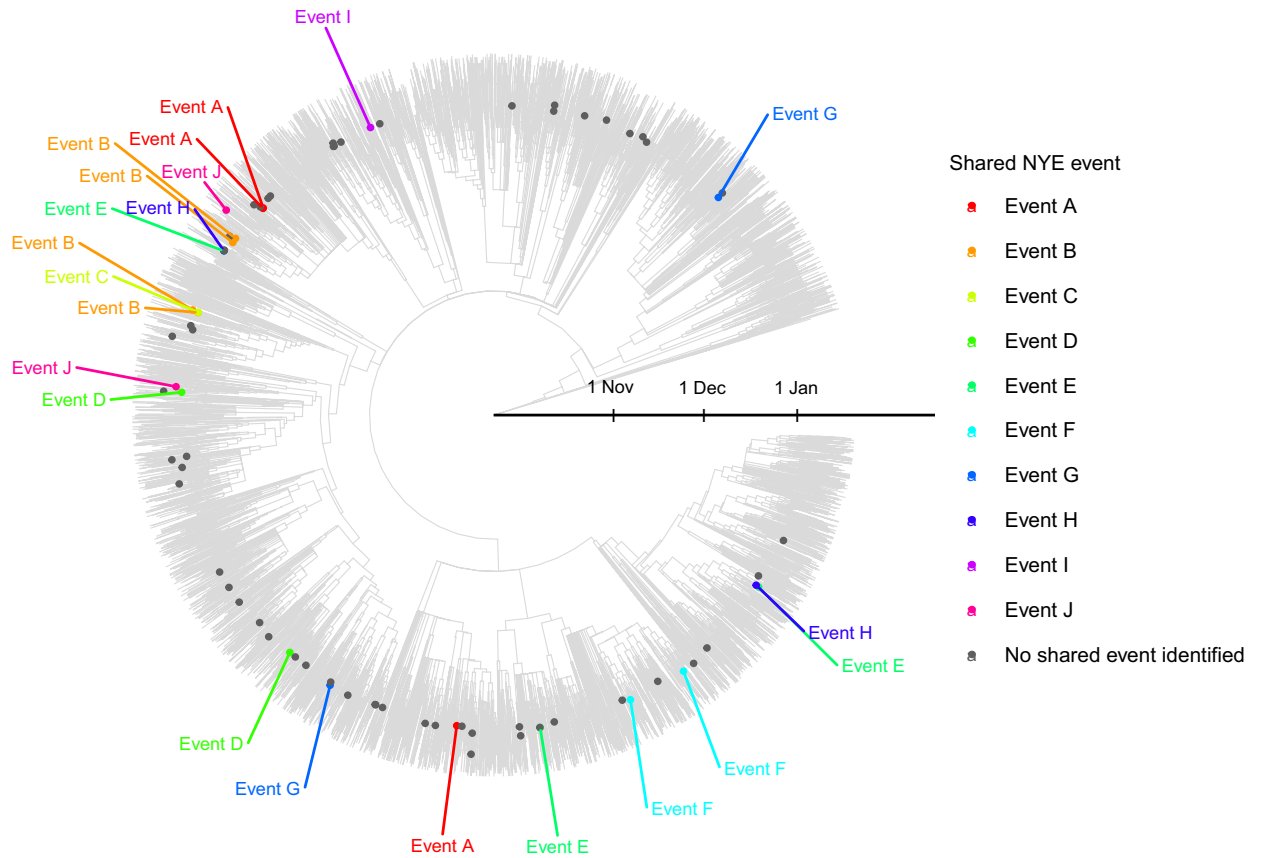


Figure 2. Phylogenetic context of NYE related infections in the student population. This figure shows a time-scaled phylogeny based on the maximum likelihood phylogenetic analysis of an alignment including all Belgian baseline surveillance sequences of the Omicron BA.1 lineage available on GISAID and sampled until 20 January 2022, as well as a selection of BA.1 sequences from other European countries. Dots show sequences obtained from Leuven students possibly or likely infected on NYE, coloured by which shared event they attended.

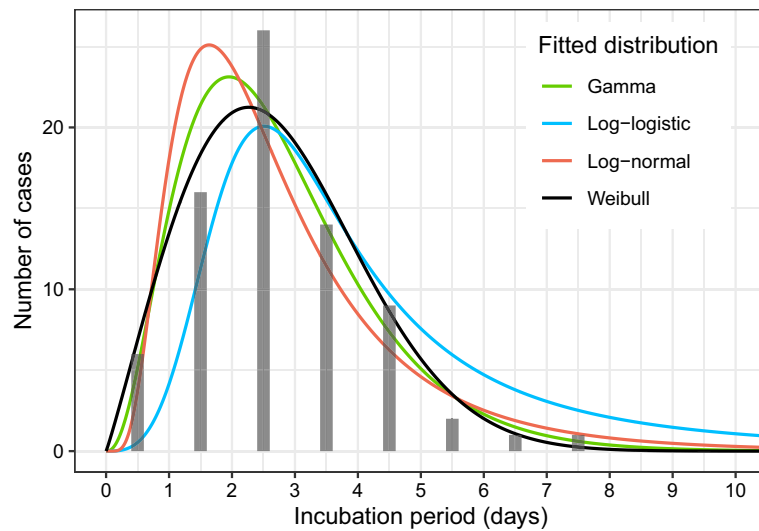


Figure 3. Observed incubation periods of 75 cases possibly or likely infected with the Omicron BA.1 strain on NYE. Various non-negative distributions are fitted using the maximum-likelihood method. The best fit is obtained with a Weibull distribution (shape and scale parameters 2.08 and 3.11, respectively) with a median of 2.6 days.

National reproduction number

We visualise the effect of Christmas and NYE celebrations on a national scale by estimating R_{eff} from publicly available daily case numbers (Fig. 4)²⁰. As symptom onset data were not available, we corrected for weekly reporting effects using the method by Alvarez *et al.*²¹. This method, which results in smoothing of the estimated reproduction number, shows a wide peak around Christmas and NYE. A second peak follows in the first week of schools reopening.

Similar peaks in transmissivity around the holiday celebrations can be observed in many surrounding European countries (Supplementary Fig. 3).

Discussion

Taken together, these results suggest a strong link between simultaneous social gatherings on NYE and a subsequent surge of detected COVID-19 cases amongst students in Leuven, Belgium. The combined phylogenetic and contact tracing data imply that small-scale social gatherings on NYE, rather than superspreading events, were the main contributor to a surge in new cases.

Mandated restrictions on social interactions resulted in limited event sizes, taking place mainly in private residential settings. As a result, despite a high secondary attack rate, superspreading events were responsible for only 16 to 34% of cases infected on NYE. Both contact tracing and phylogenetic data indicate that the surge of infections was not caused by one or a few superspreading events.

To our knowledge, this study shows the first example in the literature of a strong link between public holiday celebrations and a surge in cases of SARS-CoV-2. Interestingly, simultaneous celebrations in small circles, rather than mass gatherings, appeared as the main driver of the surge.

The observed wave of cases consisted largely of first-generation infections directly originating from these social gatherings. Incidence in the targeted student population quickly stabilised thereafter, likely due to a combination of government-mandated contact restrictions, low threshold testing, contact tracing and reduced social interaction due to the students' approaching January exams.

However, in the absence of these limiting factors, such a simultaneous increase in transmissivity could logically accelerate an exponential rise in cases. In fact, nationally reported case numbers and their derived effective reproduction numbers suggest that this may be exactly what happened on a national and even international scale with the spread of the Omicron BA.1 strain.

Estimates of the national effective reproduction number suggest a possible “triple whammy” of transmissibility effects around the winter holidays. First, Christmas and New Year celebrations provide two occasions for large scale inter-generational and inter-regional transmission, as friends and family gather with a range of risk factors: enclosed spaces, extended contact, close physical interaction, inter-household contact and strong vocalisation. The resulting cases have relatively little opportunity for onward transmission until schools and many workplaces reopen simultaneously, leading to further transmissions within regions and generations. Further research might confirm the relative contributions of transmission across regions and generations.

The incubation period is an important parameter for epidemiological modelling. We attain similar values for the mean and median incubation period of the Omicron strain (2.7 and 2.6 days, respectively) as reported previously, while providing further data on the incubation period distribution^{14,16}. Values ranged from less than one to eight days and followed a broad unimodal skewed-right distribution. The best fit was obtained with a Weibull distribution with 5th and 95th percentiles of 0.7 and 5.3 days, respectively (Fig. 3).

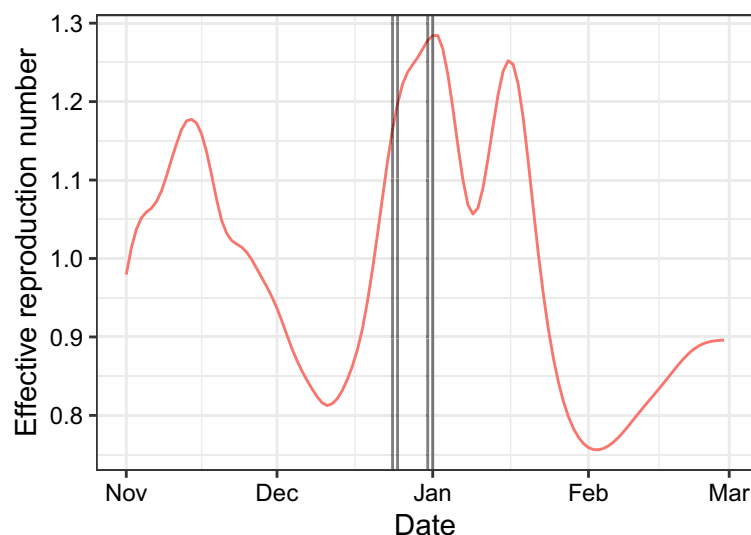


Figure 4. Estimate of the effective reproduction number (R_{eff}) on a national scale. Reported case numbers in Belgium were used to estimate R_{eff} with EpiInvert. Vertical black lines indicate the following dates, which typically correspond with holiday celebrations: 24 December, 25 December, 31 December, and 1 January.

This study has several limitations. First, its focus on a population of higher education students limits generalisability to other age and socio-economic groups.

Second, determining the source event of an infection through contact tracing has inherent limitations. Our definition of “likely infected on NYE” excludes cases who reported multiple plausible source events, but self-reporting bias may still reduce the reliability of the data. Indeed, the phylogenetic analysis indicated that some cases linked to the same event were not closely related. This observation could further support a limited role of superspreading, but could also indicate that the NYE event was not the source of their infection. However, the association with NYE is further supported by a peak in the effective reproduction number and a surge in cases above the forecast.

Third, the study design and the lack of exposure data from negative controls preclude definite evidence for a causal link in the epidemiological sense.

Finally, we were unable to determine the number of secondary infections outside of the target population, and thus the direct effect of NYE on national case numbers.

In summary, a simultaneous increase in social interactions, as can occur during holiday celebrations, can lead to a surge of SARS-CoV-2 infections. This can happen even in the presence of social restrictions, through an increase of transmissions at small-scale gatherings.

Methods

Setting

The study was conducted in the setting of a dedicated testing and contact tracing program targeting around 50,000 higher education students at KU Leuven Association in Leuven, Belgium. A test and trace program was implemented, specifically targeting this student population. Students could attend for a PCR test at no cost and even in the absence of specific test criteria. The test centre closely cooperated with the contact tracing team to improve the flow of information.

New Year’s Eve 2022 (NYE) encouraged many students to briefly meet with friends and family, before returning to a generally solitary period of studying for their January tests. On a national level, the Omicron BA.1 strain was dominant and rising at the time²². Government-mandated restrictions resulted in limited event sizes (Supplementary Fig. 2)²³.

Inclusion and exclusion

In the days after NYE, we observed a surge in case numbers and initiated further investigation. We attempted to interview each case with a positive result on the TaqPath COVID-19 assay at the university test centre between 25 December 2021 and 14 January 2022. Cases with a previous positive test result in the 60 days preceding their sample collection were excluded, as well as samples with only a weak positive signal (C_q value for SARS-CoV-2 N-gene target above 30). To construct the epidemic curve, we plotted symptomatic cases according to self-reported disease onset date. Based on national genomic monitoring, S-gene target failure of the TaqPath COVID-19 assay was used as a marker of the Omicron BA.1 lineage²². S-gene positive samples were considered as Delta before and Omicron BA.2 from 17 January 2022 onwards, corresponding to the first week when the share of new Omicron BA.2 cases overtook the share of Delta cases on a national level (Supplementary Fig. 1)²².

Effective reproduction number

For the local student population, R_{eff} was estimated from the epidemic curve as described by Cori et al. with a one-day sliding window¹⁷. A one-day window was expected to result in broad credible intervals. However, this window was chosen because we were investigating a suspected change in transmissibility occurring during a very short time (NYE). Therefore, estimates of daily variability were considered useful. We used a parametric gamma distribution of the serial interval with a mean of 2.75 days (SD: 2.54 days)²⁴. A Grubbs test was used to determine whether the main peak of the R_{eff} estimate was an outlier or a stochastic effect.

We inferred R_{eff} on a national scale from publicly reported case numbers²⁰. As their symptom onset dates were not available to us, we inferred R_{eff} from sampling dates using a 7-day sliding window in the R package EpiInvert. This method corrects for reporting biases such as the “weekend effect” but smooths the estimated value of R_{eff} ²¹. We did not use the option to ignore the effect of public holidays on R_{eff} , because they are the subject of this study. The serial interval was assumed normally distributed with a mean of 2.75 days (SD: 2.54 days). We did not correct for delays between symptom onset and sampling.

Forecast

To construct a forecast of symptomatic case numbers amongst students, we first estimated R_{eff} with the above method (Cori et al., one-day window). However, here we only included cases with symptom onset before NYE. We used the mean R_{eff} estimate in the seven days before NYE to calculate the forecast, assuming that R_{eff} was constant throughout this period.

Definition of superspreading event

A SSE was defined as an event with more transmissions than would originate from 99% of infectious cases, when assuming a simple Poisson distribution of infections²⁵. Using our estimated R_{eff} of 3.6 for Leuven students on NYE, events with at least nine transmissions were considered superspreading events.

An alternative would have been to base our definition of a SSE on the estimated R_{eff} in the days before NYE. However, in that case, any increase in transmissibility on NYE – even homogeneously distributed across the population – would have resulted in a higher observed proportion of transmissions occurring at superspreading

events. Our aim was instead to determine whether individual variation in transmissibility (i.e., superspreading) increased on NYE. Therefore, in our definition of a SSE, the threshold number of transmissions varies with R_{eff} .

Source investigation

At least one attempt was made to contact each student with a positive PCR test for a telephone interview, in an effort to trace the source of their infection and their close contacts during the contagious period. Close contact was defined as direct contact, or an interaction with at least 2 of the following characteristics: duration longer than 15 min, distance shorter than 1.5 m and absence of face masks. The contact elicitation window was defined as starting 7 days before diagnosis or symptom onset of the case, whichever was earlier²⁶.

Particular attention was paid to NYE events, even if this meant extending the contact elicitation window. A NYE event was defined as any social gathering involving close contact (within or outside the household) occurring between noon on 31 December 2021 and noon on 1 January 2022. The contact tracers determined which cases were possibly and likely infected on NYE using the following criteria. Cases were classed as “possibly infected on NYE” if they had attended an NYE event (with or without other known cases in attendance), developed symptoms 0 to 14 days afterward, and – in the period from 1 week before NYE to symptom onset – had no close contact with a potential source case who was not present at the NYE event. Cases fulfilling these criteria were additionally considered “likely infected on NYE” if they were in contact with another known case during the NYE event, but not on any of the preceding or subsequent seven days.

Phylogenetic analysis

Our phylogenetic analysis included, first of all, the genomic sequences obtained from students possibly or likely infected on NYE.

In addition, we included all Belgian Omicron sequences available on GISAID, sampled between 1 November 2021 and 20 January 2022, and analysed for the purpose of baseline genomic surveillance. This was determined by searching for the word “baseline” in the “sampling strategy” metadata field. If this field was blank, the sequence was excluded.

Finally, a selection was included of other European Omicron sequences from the same period, labelled as “high coverage” by GISAID. “High coverage” here indicates less than 1% of undefined bases, less than 0.05% unique amino acid mutations, and no insertions or deletions unless verified by the submitter. The selection was obtained through proximity sampling – genetic similarity to the included Belgian sequences – and limited to 20 genomic sequences per country and per month.

Sequences with fewer than 27,000 known bases or not belonging to the BA.1 lineage were excluded.

The Nextstrain pipeline with its default parameters was used to construct a time-scaled phylogenetic tree²⁷. Multiple sequence alignment was accomplished using MAFFT, with the WIV04 genome (GISAID accession number EPI_ISL_402124) as the reference sequence²⁸. A phylogenetic tree was constructed using Augur, calling IQ-TREE with a general time reversible model, and time-scaled using TreeTime^{29–31}. Phylogenetic clustering was performed with the cov2clusters method, using previously reported coefficients ($\beta_0 = 3$, $\beta_1 = -19,735.98$, $\beta_2 = -0.075$) and a pairwise probability threshold of 90%¹⁹.

As our analysis was focused on the Omicron BA.1 lineage, only the corresponding Nextstrain clade (21 K) was visualised using the R package ggtree³².

Ethical approval

The study protocol was approved by the Ethics Committee Research UZ / KU Leuven (reference no. S64919). The planning, conduct and reporting of the study were performed in line with the Declaration of Helsinki, as revised in 2013. The Ethics Committee Research UZ / KU Leuven waived the need for individual informed consent, since the data gathered did not exceed what was required for the purpose of safeguarding public health.

Data availability

The findings of this study are based on metadata associated with 8,770 sequences available on GISAID up to 10 October 2022, via gisaid.org/EPI_SET_221010sa (NYE samples, Supplementary Table 4) and gisaid.org/EPI_SET_221128mt (Belgian and European context, Supplementary Table 3). All other data generated or analysed during this study are included in this published article and its supplementary information files (Supplementary Tables 1, 2).

Received: 20 December 2022; Accepted: 11 December 2023

Published online: 14 December 2023

References

1. Rocha, I. C. N., Pelayo, M. G. A. & Rackimuthu, S. Kumbh mela religious gathering as a massive superspreading event: Potential culprit for the exponential surge of COVID-19 cases in India. *Am. J. Trop. Med. Hygiene* <https://doi.org/10.4269/ajtmh.21-0601> (2021).
2. Rocha, I. C. N., Pelayo, M. G. A. & Sammatid, C. Traveling and celebrating during songkran as super spreading events: a potential triggering factor of the surge of COVID-19 cases in Thailand. *Int. J. Travel Med. Glob. Health* **9**, (2021).
3. Klausner, Z., Fattal, E., Hirsch, E. & Shapira, S. C. A single holiday was the turning point of the COVID-19 policy of Israel. *Int. J. Infect. Dis.* **101**, 368–373 (2020).
4. Aragão, D. P., dos Santos, D. H., Mondini, A. & Gonçalves, L. M. G. National holidays and social mobility behaviors: Alternatives for forecasting COVID-19 deaths in Brazil. *Int. J. Environ. Res. Public Health* **18**, 11595 (2021).

5. Rahman, F. N., Rahman, A. K. M. F., Iwuagwu, A. O. & Dalal, K. COVID-19 transmission due to mass mobility before and after the largest festival in Bangladesh: An epidemiologic study. *Inquiry J. Health Care Organ. Provision Financ.* **58**, 004695802110234 (2021).
6. Chen, J. *et al.* Travel rush during Chinese spring festival and the 2019-nCoV. *Travel Med. Infect. Dis.* **37**, 101686 (2020).
7. Mehta, S. H., Clipman, S. J., Wesolowski, A. & Solomon, S. S. Holiday gatherings, mobility and SARS-CoV-2 transmission: results from 10 US states following Thanksgiving. *Sci. Rep.* **11**, 17328 (2021).
8. Koopsen, J. *et al.* Epidemiologic and genomic analysis of SARS-CoV-2 delta variant superspreading event in nightclub, The Netherlands, June 2021. *Emerg. Infect. Dis.* **28**, 1012–1016 (2022).
9. Kang, C. R. *et al.* Coronavirus disease exposure and spread from nightclubs, South Korea. *Emerg. Infect. Dis.* **26**, (2020).
10. Endo, A., Abbott, S., Kucharski, A. J. & Funk, S. Estimating the overdispersion in COVID-19 transmission using outbreak sizes outside China. *Wellcome Open Res* **5**, (2020).
11. Goyal, A., Reeves, D. B. & Schiffer, J. T. Multi-scale modelling reveals that early super-spreader events are a likely contributor to novel variant predominance. *J. R. Soc. Interface* **19**, (2022).
12. Wong, G. *et al.* MERS, SARS, and Ebola: The role of super-spreaders in infectious disease. *Cell Host Microbe* <https://doi.org/10.1016/j.chom.2015.09.013> (2015).
13. Vanbesien, M. *et al.* Risk factors for SARS-CoV-2 transmission in student residences: a case-ascertained study. *Arch. Public Health* **80**, 212 (2022).
14. Du, Z. *et al.* Shorter serial intervals and incubation periods in SARS-CoV-2 variants than the SARS-CoV-2 ancestral strain. *J. Travel Med.* <https://doi.org/10.1093/jtm/taac052> (2022).
15. Xin, H. *et al.* The incubation period distribution of Coronavirus Disease 2019: A systematic review and meta-analysis. *Clin. Infect. Dis.* **73**, 2344–2352 (2021).
16. Tanaka, H. *et al.* Shorter incubation period among COVID-19 cases with the BA.1 omicron variant. *Int. J. Environ. Res. Public Health* **19**, 6330 (2022).
17. Cori, A., Ferguson, N. M., Fraser, C. & Cauchemez, S. A new framework and software to estimate time-varying reproduction numbers during epidemics. *Am. J. Epidemiol.* **178**, 1505–1512 (2013).
18. Shu, Y. & McCauley, J. GISAID: Global initiative on sharing all influenza data – from vision to reality. *Eurosurveillance* **22**, (2017).
19. Sobkowiak, B. *et al.* Cov2clusters: genomic clustering of SARS-CoV-2 sequences. *BMC Genomics* **23**, 1–11 (2022).
20. Sciensano. Epistat - COVID-19 Monitoring. <https://epistat.sciensano.be/covid/>.
21. Alvarez, L., Colom, M., Morel, J.-D. & Morel, J.-M. Computing the daily reproduction number of COVID-19 by inverting the renewal equation using a variational technique. *Proc. Natl. Acad. Sci.* **118**, (2021).
22. Cuypers, L., Baele, G., Dellicour, S., Maes, P. & André, E. Genomic surveillance report - Update for Belgium, 26/04/2022. <https://www.uzleuven.be/nl/laboratoriumgeneeskunde/genomic-surveillance-sars-cov-2-belgium> (2022).
23. Consultative Committee tightens preventative health measures against omicron variant. https://www.belgium.be/en/news/2021/consultative_committee_tightens_preventative_health_measures_against_omicron_variant (2021).
24. Kremer, C. *et al.* Serial intervals for SARS-CoV-2 omicron and delta variants, Belgium, November 19–December 31, 2021. *Emerg. Infect. Dis.* **28**, 1699–1702 (2022).
25. Lloyd-Smith, J. O., Schreiber, S. J., Kopp, P. E. & Getz, W. M. Superspreading and the effect of individual variation on disease emergence. *Nature* **438**, 355–359 (2005).
26. Raymenants, J. *et al.* Empirical evidence on the efficiency of backward contact tracing in COVID-19. *Nat. Commun.* **13**, 4750 (2022).
27. Hadfield, J. *et al.* Nextstrain: real-time tracking of pathogen evolution. *Bioinformatics* **34**, 4121–4123 (2018).
28. Katoh, K. & Standley, D. M. MAFFT multiple sequence alignment software version 7: Improvements in performance and usability. *Mol. Biol. Evol.* **30**, 772–780 (2013).
29. Huddleston, J. *et al.* Augur: a bioinformatics toolkit for phylogenetic analyses of human pathogens. *J. Open Source Softw.* **6**, 2906 (2021).
30. Nguyen, L.-T., Schmidt, H. A., von Haeseler, A. & Minh, B. Q. IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Mol. Biol. Evol.* **32**, 268–274 (2015).
31. Sagulenko, P., Puller, V. & Neher, R. A. TreeTime: Maximum-likelihood phylodynamic analysis. *Virus Evol.* **4**, (2018).
32. Xu, S. *et al.* Gtree: A serialized data object for visualization of a phylogenetic tree and annotation data. *iMeta* (2022) <https://doi.org/10.1002/imt2.56>.

Acknowledgements

We gratefully acknowledge all data contributors, i.e., the Authors and their Originating laboratories responsible for obtaining the specimens, and their Submitting laboratories for generating the genetic sequence and metadata and sharing via the GISAID Initiative, on which this research is based. All members of the Belgian SARS-CoV-2 genomic consortium are gratefully acknowledged. We thank everyone involved in the set-up and running of the KU Leuven test and trace program, especially the many student workers who formed the bulk of the team. The test centre was funded by both the National Institute for Health and Disability Insurance (RIZIV/INAMI) and the regional Flemish government's Agentschap voor Zorg & Gezondheid. The contact tracing program benefited from the generous contributions of KU Leuven. UZ Leuven, as a national reference centre for respiratory pathogens, is supported by Sciensano, which is gratefully acknowledged. JT is funded through a Research Foundation Flanders (FWO) fundamental research fellowship (grant no. 1130423N). JR is funded through a FWO strategic basic research fellowship (grant no. 1S88721N). SD acknowledges support from the *Fonds National de la Recherche Scientifique* (F.R.S.-FNRS, Belgium; grant no. F.4515.22), from FWO (grant no. G098321N), and from the European Union Horizon 2020 project MOOD (grant agreement no. 874850). CG and EA acknowledge support from the DURABLE project. The DURABLE project has been co-funded by the European Union, under the EU4Health Programme (EU4H), project no. 101102733. Co-funded by the European Union. Views and opinions expressed are however those of the authors only and do not necessarily reflect those of the European Union or the European Health and Digital Executive Agency. Neither the European Union nor the granting authority can be held responsible for them.

Author contributions

C.G., E.A., and S.D. conceived and designed the analysis. C.G., J.T., and L.L. collected the data. C.G. and J.T. performed the analysis. C.G. wrote the paper. J.T., L.L., J.R., L.C., P.M., S.D., and E.A. critically reviewed the paper.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1038/s41598-023-49678-x>.

Correspondence and requests for materials should be addressed to C.G.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2023