nature

https://doi.org/10.1038/s41586-020-2488-1

Accelerated Article Preview

Suppression of a SARS-CoV-2 outbreak in the Italian municipality of Vo'

Received: 2 April 2020

Accepted: 23 June 2020

Accelerated Article Preview Published online 30 June 2020

Cite this article as: Lavezzo, E. et al. Suppression of a SARS-CoV-2 outbreak in the Italian municipality of Vo'. *Nature* https://doi.org/10.1038/s41586-020-2488-1 (2020).

Enrico Lavezzo, Elisa Franchin, Constanze Ciavarella, Gina Cuomo-Dannenburg, Luisa Barzon, Claudia Del Vecchio, Lucia Rossi, Riccardo Manganelli, Arianna Loregian, Nicolò Navarin, Davide Abate, Manuela Sciro, Stefano Merigliano, Ettore De Canale, Maria Cristina Vanuzzo, Valeria Besutti, Francesca Saluzzo, Francesco Onelia, Monia Pacenti, Saverio Parisi, Giovanni Carretta, Daniele Donato, Luciano Flor, Silvia Cocchio, Giulia Masi, Alessandro Sperduti, Lorenzo Cattarino, Renato Salvador, Michele Nicoletti, Federico Caldart, Gioele Castelli, Eleonora Nieddu, Beatrice Labella, Ludovico Fava, Matteo Drigo, Katy A. M. Gaythorpe, Imperial College COVID-19 Response Team, Alessandra R. Brazzale, Stefano Toppo, Marta Trevisan, Vincenzo Baldo, Christl A. Donnelly, Neil M. Ferguson, Ilaria Dorigatti & Andrea Crisanti

Article

Suppression of a SARS-CoV-2 outbreak in the Italian municipality of Vo'

https://doi.org/10.1038/s41586-020-2488-1

Received: 2 April 2020

Accepted: 23 June 2020

Published online: 30 June 2020

Enrico Lavezzo[1,12], Elisa Franchin[1,12], Constanze Ciavarella[3], Gina Cuomo-Dannenburg[3], Luisa Barzon[1], Claudia Del Vecchio[1], Lucia Rossi[2], Riccardo Manganelli[1], Arianna Loregian[1], Nicolò Navarin[6,11], Davide Abate[1], Manuela Sciro[2], Stefano Merigliano[7], Ettore De Canale[2], Maria Cristina Vanuzzo[2], Valeria Besutti[2], Francesca Saluzzo[1], Francesco Onelia[1], Monia Pacenti[2], Saverio Parisi[1], Giovanni Carretta[2], Daniele Donato[2], Luciano Flor[2], Silvia Cocchio[9], Giulia Masi[1], Alessandro Sperduti[6,11], Lorenzo Cattarino[3], Renato Salvador[7], Michele Nicoletti[10], Federico Caldart[10], Gioele Castelli[10], Eleonora Nieddu[10], Beatrice Labella[10], Ludovico Fava[10], Matteo Drigo[10], Katy A. M. Gaythorpe[3], Imperial College COVID-19 Response Team*, Alessandra R. Brazzale[8], Stefano Toppo[1,11], Marta Trevisan[1], Vincenzo Baldo[9], Christl A. Donnelly[3,4], Neil M. Ferguson[3], Ilaria Dorigatti[3,13]✉ & Andrea Crisanti[1,2,5,13]✉

On the 21st of February 2020 a resident of the municipality of Vo', a small town near Padua, died of pneumonia due to SARS-CoV-2 infection[1]. This was the first COVID-19 death detected in Italy since the emergence of SARS-CoV-2 in the Chinese city of Wuhan, Hubei province[2]. In response, the regional authorities imposed the lockdown of the whole municipality for 14 days[3]. We collected information on the demography, clinical presentation, hospitalization, contact network and presence of SARS-CoV-2 infection in nasopharyngeal swabs for 85.9% and 71.5% of the population of Vo' at two consecutive time points. On the first survey, which was conducted around the time the town lockdown started, we found a prevalence of infection of 2.6% (95% confidence interval (CI) 2.1-3.3%). On the second survey, which was conducted at the end of the lockdown, we found a prevalence of 1.2% (95% Confidence Interval (CI) 0.8-1.8%). Notably, 42.5% (95% CI 31.5-54.6%) of the confirmed SARS-CoV-2 infections detected across the two surveys were asymptomatic (i.e. did not have symptoms at the time of swab testing and did not develop symptoms afterwards). The mean serial interval was 7.2 days (95% CI 5.9-9.6). We found no statistically significant difference in the viral load of symptomatic versus asymptomatic infections (p-values 0.62 and 0.74 for *E* and *RdRp* genes, respectively, Exact Wilcoxon-Mann-Whitney test). This study sheds new light on the frequency of asymptomatic SARS-CoV-2 infection, their infectivity (as measured by the viral load) and provides new insights into its transmission dynamics and the efficacy of the implemented control measures.

As of 23rd May 2020, 5,105,881 confirmed cases and 333,446 deaths of a Novel Coronavirus Disease (COVID-19) have been reported worldwide[2]. In Italy, COVID-19 has caused over 32,616 confirmed deaths. The causative agent (SARS-CoV-2), a close relative of SARS-CoV[4], was introduced into the human population of Wuhan City, Hubei province (China) around the beginning of December 2019[5,6]. In Hubei province and in the rest of mainland China, recent reports suggest that strategies based on the isolation of cases and their contacts, along with drastic social distancing measures that include the quarantine of whole cities and regions, the closure of schools and workplaces and the cancellations of mass gatherings had a tremendous effect on the control of the epidemic[7,8]. However, the long-term effectiveness of these interventions remains unclear[9]. In Europe, similar interventions have been implemented to control the transmission of SARS-CoV-2. Recent analyses suggest that control is likely to be achieved across Europe[10]. In Italy, interventions have successfully controlled SARS-CoV-2 transmission in all regions, but uncertainties remain about the ability to avoid a resurgence of transmission as interventions are relaxed[11]. Effective long-term control of transmission in Europe and worldwide, depends on an improved understanding of the mechanisms of SARS-CoV-2 transmission, particularly on the relative contribution of asymptomatic, pre-symptomatic and symptomatic transmission[12]. This is particularly important given that, in the absence of a vaccine or effective treatment, alternative public health interventions are being evaluated to allow

[1]Department of Molecular Medicine, University of Padova, Padova, Italy. [2]Azienda Ospedale Padova, Via Giustiniani 1, Padova, Italy. [3]MRC Centre for Global Infectious Disease Analysis, Imperial College London, London, UK. [4]Department of Statistics, University of Oxford, Oxford, UK. [5]Imperial College London, Exhibition Road, London, UK. [6]Department of Mathematics "Tullio Levi-Civita", University of Padova, Padova, Italy. [7]Department of Surgery, Oncology and Gastroenterology, University of Padova, Padova, Italy. [8]Department of Statistical Sciences, University of Padova, Padova, Italy. [9]Department of Cardiac, Thoracic, Vascular Sciences and Public Health, University of Padova, Padova, Italy. [10]School of Medicine, University of Padova, Padova, Italy. [11]CRIBI Biotech Center, University of Padova, Padova, Italy. [12]These authors contributed equally: Enrico Lavezzo, Elisa Franchin. [13]These authors jointly supervised this work: Ilaria Dorigatti, Andrea Crisanti. *A list of authors and their affiliation appears at the end of the paper. ✉e-mail: i.dorigatti@imperial.ac.uk; a.drcrisanti@imperial.ac.uk

the population to maintain essential societal and economic activities while controlling the spread of SARS-CoV-2, limiting mortality and maintaining healthcare demand within capacity.

In this study we present the results of two surveys of the resident population of Vo', conducted less than two weeks apart, to investigate population exposure to SARS-CoV-2 before and after the lockdown. We present an analysis of population demography, prevalence of infection, viral load and frequency of symptomatic, asymptomatic and pre-symptomatic infections. We assessed the risk of SARS-CoV-2 infection associated with comorbidity and therapies for underlying conditions, characterised chains of transmission, studied the transmission dynamics of SARS-CoV-2 and assessed the impact of the lockdown. Our analyses show that viral transmission could be effectively and rapidly suppressed by combining the early isolation of infected people with community lockdown. The experience of Vo' shows that despite the silent and widespread transmission of SARS-CoV-2, transmission can be controlled and represents a model for the systematic suppression of viral outbreaks under similar epidemiological and demographic conditions.

Results

During the two surveys we collected nasopharyngeal swabs from 2,812 and 2,343 subjects, corresponding to 85.9% and 71.5% of the eligible study population (Figure 1). All age groups were homogeneously sampled with age-specific percentages ranging from 57.1% to 95.4% in the first survey and 40.1% to 80.4% in the second survey (Extended Data Table 1). Statistical analysis showed that while the recruited and non-recruited populations are different in terms of age distribution (p-values < 0.001 for the first and second surveys, Fisher's exact test), there was no statistically significant bias in the composition of the different age groups enrolled in the two surveys (p-value = 0.31, Exact Wilcoxon-Mann-Whitney Test) (Extended Data Figure 1). Notably, no additional infections were reported in Vo' in spite of the escalating epidemic in the surrounding regions.

Analysis of infection prevalence

A total of 73 out of the 2,812 subjects tested at the first survey were positive, which gives a prevalence of 2.6% (95% CI 2.1-3.3%) (Table 1). The second survey identified 29 total positive cases (prevalence 1.2%; 95% CI 0.8-1.8%), 8 of which were new cases (0.3%; 95% CI 0.15-0.7%) (Figure 2). One of the 8 new infections detected in the second survey was a hospitalized subject who tested positive, then negative, then positive again. It is unclear whether this was a case of SARS-CoV-2 re-infection or the second test was a false negative. The frequency of the symptoms in the SARS-CoV-2 positive individuals was systematically recorded, with fever and cough being the most common (Extended Data Figure 1). Notably, a total of 29 out of the 73 individuals (39.7%; 95% CI 28.5-51.9%) who tested positive at the first survey were asymptomatic (i.e. did not show symptoms at the time of swab sampling nor afterwards, see definition of symptomatic in the Methods section). A similar proportion of asymptomatic infection was also recorded at the second survey (13 out of 29, 44.8%; 95% CI 26.5-64.3%); of the 8 new cases, 5 were asymptomatic (Table 2, Extended Data Figure 2). No infections were detected in either survey in 234 tested children ranging from 0 to 10 years, including those living in the same household as infected individuals (Extended Data Table 3). Up to the age of 50 years, the prevalence of infection oscillated between a central estimate of 1.2% to 1.7% (Extended Data Figure 1). Older individuals showed a three-fold increase in the infection prevalence (Table 2, Extended Data Figure 1). Of the 81 SARS-CoV-2 positive patients across the two surveys, 13 required hospitalization (16.0%). Their age distribution was as follows: 1 (7.7%) aged 41-50, 1 (7.7%) aged 51-60, 4 (30.8%) aged 61-70, 5 (38.5%) aged 71-80 and 2 (15.4%) aged 81-90.

A substantial fraction of infected individuals (58.9%; 95% CI 46.8%-70.3%, pre-symptomatic, symptomatic and asymptomatic combined over all ages) cleared the infection between the first and second surveys, i.e. had a negative test at the second survey after a positive test at the first survey (Extended Data Table 2). For all infections identified in the study, clearance was confirmed by an additional negative test conducted independently by the local health authority (data not shown). The time to viral clearance (time from the earliest positive sample for the subjects with more than one sample in the first survey and a negative sample in the second survey) ranged from 8 to 13 days and was on average 9.3 days, with standard deviation 2.0 days. The minimal duration of the positivity window (time from the earliest positive sample in the first survey and a positive sample in the second survey) ranged from 3 to 13 days and was on average 9.1 days, with standard deviation 2.3 days. In particular, 61.4% (95% CI 45.5-75.6%) of symptomatic and 55.2% (95% CI 35.7-73.6%) of asymptomatic SARS-CoV-2 infections cleared the virus during the study period (i.e. had a negative test after a positive result at the first survey); the highest rate (100%) was observed in the age groups of symptomatic 31-40 and 41-50 year olds (Extended Data Table 2). SARS-CoV-2 positivity overall (i.e. first and second survey combined) and at the first survey was more frequently associated with 71-80 year olds (compared to 21-30 year olds, p-value = 0.012 and p-value = 0.017 respectively) (Extended Data Figure 1). Being male was associated with COVID-19 positivity in the second survey (p-value = 0.04) (Table 2). Analyses of the association between common comorbidities such as diabetes, hypertension, vascular diseases, respiratory diseases in asymptomatic and symptomatic individuals and the use of treatment for a number of different conditions with symptomatic infection showed no significant association (Supplementary Table S3 and Supplementary Table S4).

Role of asymptomatic transmission

The analysis of viral genome equivalents inferred from Ct (cycle threshold) data from real-time reverse-transcription PCR (RT-PCR) assays indicated that asymptomatic and symptomatic individuals did not differ when compared for viral PCR template recovered in the nasopharyngeal swabs (p-values 0.62 and 0.74 for gene *E* and gene *RdRp*, respectively; Exact Wilcoxon-Mann-Whitney) (Extended Data Figure 3). We find that the viral load tends to peak around the day of symptom onset and for most of the subjects tends to decline after symptom onset (Extended Data Figure 3). The relative risk of contracting the infection for having close contacts with an infected relative, including those living in the same household gives an odd ratio of 84.5 (95% CI 16.8-425.4) (Supplementary Text S3 and Extended Data Table 4). Two out of the eight new infections detected in the second survey either shared household or had a contact history with asymptomatic individuals (Supplementary Table S1).

Reconstructing transmission chains

From the inferred transmission pairs, we estimated a serial interval distribution over the whole study period with mean 7.2 days (95% CI 5.9-9.6). We found that the lockdown reduced the serial interval from a mean of 7.6 days (95% CI: 6.4-8.7) before the lockdown to a mean of 6.2 days (95% CI: 2.6-10.7) after the lockdown. We also found that the lockdown substantially reduced transmission, with the reproduction number dropping from an initial value of 2.49 (95% CI 1.31-4.00) before the lockdown to 0.41 (95% CI 0.21-0.63) after the lockdown.

Modelling point prevalence data

We used the prevalence estimates obtained in Vo' at the first and second survey to calibrate a modified SEIR compartmental model of SARS-CoV-2 transmission that incorporates symptomatic, pre-symptomatic and asymptomatic infections, virus detectability (in swabs) before and after the infectious period and the impact of the

lockdown (Extended Data Figure 5). We assumed that pre-symptomatic, symptomatic and asymptomatic infections transmit the virus. We estimated that on average 41% of the infections are asymptomatic, that the mean infectious period is approximately 3.6 to 6.5 days, and that the lockdown reduced SARS-CoV-2 transmissibility, on average by between 82 to 98%, depending on the assumed initial value of R_0^1 and on the duration of virus detectability (Extended Data Table S5). The model suggests that on average up to 86.2% (range 82.2-91.6%)of the population would have been infected in the absence of interventions and that with the lockdown, 4.9% (range 2.9-8.1%) of the population of Vo' was infected by SARS-CoV-2 (Figure 3). These estimates are in line with the attack rates recently estimated for the Veneto region[11]. The model suggests that shorter values of the average duration of virus detectability beyond the infectious period better capture the central point prevalence estimates (Supplementary Table S5, Extended Data Figures 6). Our results suggest that SARS-CoV-2 was introduced into the Vo' population at the beginning of February 2020.

Discussion

The results of the two surveys carried out in Vo' provide important insights into the transmission dynamics of SARS-CoV-2. Our finding that 42.5% (95% CI 31.5-54.6%) of all confirmed SARS-CoV-2 infections across the two surveys were asymptomatic are in accordance with other population surveys[13]. Among confirmed SARS-CoV-2 infections, we did not observe significant differences in the frequency of asymptomatic infection between age groups (Figure S10, p-value = 0.96, Fisher's exact test). Among symptomatic individuals, older age groups tended to show higher frequencies of SARS-CoV-2 infection (Extended Data Table 2). Recent studies found that the clinical progression of infected children is generally milder than in adults[15–17]. We found that none of the children under 10 years of age who took part in the study tested positive for SARS-CoV-2 infection at either survey, despite at least 13 of them living together with infected family members (Extended Data Table 3). This agrees with a recent study conducted in Iceland[13] and is particularly intriguing given the very high observed odd ratio for adults to become infected when living together with SARS-CoV-2 positive family members. However, this result does not mean that children cannot be infected by SARS-CoV-2 but suggests that children may be less susceptible than adults. The pathogenesis of SARS-CoV-2 in young children is not well understood[16]. Notably, nasopharyngeal swabs are tested for the presence of SARS-CoV-2 and can only detect active infection, not exposure. A cross-sectional serological survey would clarify the actual infection rates of the whole population, including children's exposure, to SARS-CoV-2.

The contribution of asymptomatic infections to SARS-CoV-2 transmission is supported by the viral load data (Extended Data Figure 3), by the model fit to the observed prevalence data (Supplementary Table S5 and Extended Figure 6) and by the observation that 2 out of the 8 new infections detected in the second survey reported contacts with asymptomatic individuals (Supplementary Text S3). It remains to be determined the extent to which symptoms may promote viral shedding but the decreasing trend in viral load post symptom onset suggests that pre-symptomatic transmission may play an important role[18]. Asymptomatic and pre-symptomatic transmission pose clear challenges for the control of COVID-19 in the absence of strict social distancing measures or active epidemiological surveillance comprising, for instance, a test, trace and isolate strategy.

This study has informed the policy adopted by the Veneto Region, where swabs are available to all contacts of positive symptomatic cases. This testing and tracing approach has had a tremendous impact on the course of the epidemic in Veneto compared to other Italian regions. In this context, the control strategy applied to Vo' serves as a model to suppress SARS-CoV-2 transmission across spatial scales. Enhanced community surveillance, the early detection of SARS-CoV-2 transmission and the timely implementation of interventions are key to control COVID-19 and reduce its substantial public health, economic and societal burden worldwide.

Online content

Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at https://doi.org/10.1038/s41586-020-2488-1.

1. Crisanti, A. and Cassone, A. In one Italian town, we showed mass testing could eradicate the coronavirus. https://www.theguardian.com/commentisfree/2020/mar/20/eradicated-coronavirus-mass-testing-covid-19-italy-vo. (2020).
2. World Health Organisation. Novel Coronavirus (COVID-19) Situation. https://covid19.who.int/. (2020) [Accessed 23 May 2020].
3. Saini, V. Coronavirus: voices from a quarantined Italian town. https://euobserver.com/coronavirus/147552. (2020).
4. Zhou, P et al. A pneumonia outbreak associated with a new coronavirus of probable bat origin. *Nature*, **579**, 270-273 (2020).
5. Volz, E. et al. Report 5: Phylogenetic analysis of SARS-CoV-2. https://www.imperial.ac.uk/media/imperial-college/medicine/sph/ide/gida-fellowships/Imperial-College-COVID19-phylogenetics-15-02-2020.pdf (2020).
6. Centre for Health Protection of the Hong Kong Special Administrative Region Government. CHP closely monitors cluster of pneumonia cases on Mainland. https://www.info.gov.hk/gia/general/201912/31/P2019123100667.htm. (2020).
7. Prem, K., **et al**. The effect of control strategies to reduce social mixing on outcomes of the COVID-19 epidemic in Wuhan, China: a modelling study. *The Lancet Public Health*, **5**(5), 261-270 (2020).
8. Lai, S., et al. Effect of non-pharmaceutical interventions for containing the COVID-19 outbreak: an observational and modelling study. Preprint at https://www.medrxiv.org/content/10.1101/2020.03.03.20029843v3 (2020).
9. Anderson, R.M., Heesterbeek, H., Klinkenberg, D. and Hollingsworth, T.D. How will country-based mitigation measures influence the course of the COVID-19 epidemic? *The Lancet*, **395**, 931-934 (2020).
10. Flaxman, S. et al. Estimating the number of infections and the impact of nonpharmaceutical interventions on COVID-19 in 11 European countries. https://www.imperial.ac.uk/media/imperial-college/medicine/sph/ide/gida-fellowships/Imperial-College-COVID19-Europe-estimates-and-NPI-impact-30-03-2020.pdf(2020).
11. Vollmer, M.A.C. et al. Report 20: Using mobility to estimate the transmission intensity of COVID-19 in Italy: A subnational analysis with future scenarios. https://www.imperial.ac.uk/media/imperial-college/medicine/mrc-gida/2020-05-04-COVID19-Report-20.pdf (2020).
12. Mizumoto, K., Kagaya, K., Zarebski, A. and Chowell, G. Estimating the asymptomatic proportion of coronavirus disease 2019 (COVID-19) cases on board the Diamond Princess cruise ship, Yokohama, Japan. *Eurosurveillance*, **25** (10): pii=200018 (2020).
13. Gudbjartsson, D.F., et al. Spread of SARS-CoC-2 in the Icelandic population. *New England Journal of Medicine*, **382**: 2302-231 (2020).
14. Protezione Civile & Civile. Press Release. http://www.protezionecivile.gov.it/media-communication/press-release (2020).
15. Cai, J. et al. A Case Series of children with 2019 novel coronavirus infection: clinical and epidemiological features. *Clinical Infectious Diseases*, **28**: 28 (2020).
16. Zimmermann, P. and Curtis, N. Coronavirus Infections in Children Including COVID-19. *The Pediatric Infectious Disease Journal*, **39**(5): 355-368 (2020).
17. Bi, Q et al. Epidemiology and Transmission of COVID-19 in Shenzhen China: Analysis of 391 cases and 1,286 of their close contacts. Preprint at https://www.medrxiv.org/content/10.1101/2020.03.03.20028423v3, 10.1101/ 2020.03.03.20028423 (2020).
18. He, X et al. Temporal dynamics in viral shedding and transmissibility of COVID-19. *Nat Med*, **26**: 672–675 (2020).

Imperial College COVID-19 Response Team

Kylie E. C. Ainslie[3], Marc Baguelin[3], Samir Bhatt[3], Adhiratha Boonyasiri[3], Olivia Boyd[3], Lorenzo Cattarino[3], Constanze Ciavarella[3], Helen L. Coupland[3], Zulma Cucunubá[3], Gina Cuomo-Dannenburg[3], Bimandra A. Djafaara[3], Christl A. Donnelly[3,4], Ilaria Dorigatti[3], Sabine L. van Elsland[3], Rich FitzJohn[3], Seth Flaxman[3], Katy A.M. Gaythorpe[3], Will D. Green[3], Timothy Hallett[3], Arran Hamlet[3], David Haw[3], Natsuko Imai[3], Benjamin Jeffrey[3], Edward Knock[3], Daniel J. Laydon[3], Thomas Mellan[3], Swapnil Mishra[3], Gemma Nedjati-Gilani[3], Pierre Nouvellet[3], Lucy C. Okell[3], Kris V. Parag[3], Steven Riley[3], Hayley A. Thompson[3], H. Juliette T. Unwin[3], Robert Verity[3], Michaela A. C. Vollmer[3], Patrick G. T. Walker[3], Caroline E. Walters[3], Haowei Wang[3], Yuanrong Wang[3], Oliver J. Watson[3], Charles Whittaker[3], Lilith K. Whittles[3], Xiaoyue Xi[3] & Neil M. Ferguson[3]

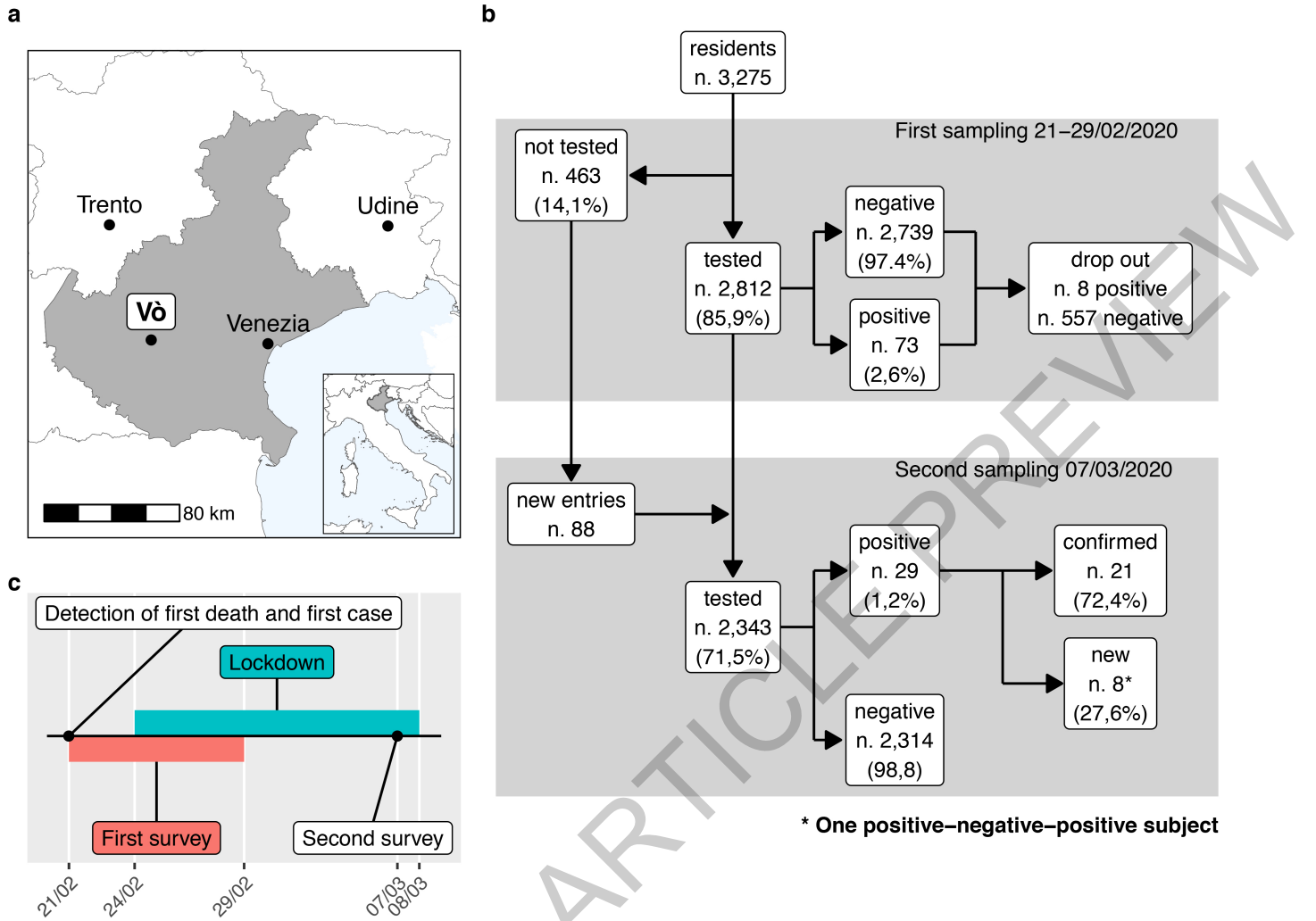


Fig. 1 | Study description. (a) Map showing the location of Vo′ and of the Veneto region (grey area) within Italy, produced using shapefiles from GADM (https://gadm.org/) and ISTAT (https://www.istat.it/it/archivio/222527 and https://www.istat.it/it/archivio/104317#accordions). (b) Flow chart summarising the key statistics on the two sequential nasopharyngeal swab surveys conducted in Vo′ to assess the transmission of SARS-CoV-2 before and after the implementation of interventions. (c) Summary of the key events in the study period.

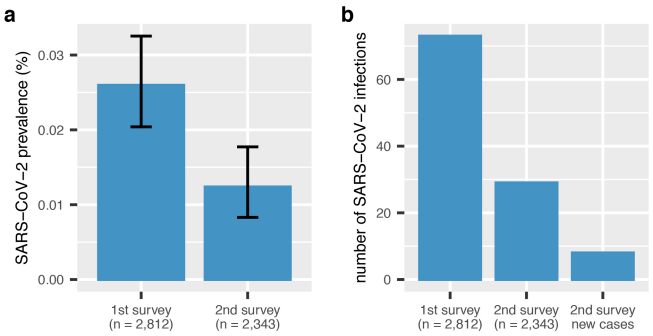


Fig. 2 | SARS-CoV-2 prevalence statistics. (a) Prevalence of SARS-CoV-2 infection at the first (x = 73 positive out of n = 2,812 tested) and second survey (x = 29 positive out of n = 2,343 tested). The error bars represent the 95% exact binomial confidence interval. (b) Number of SARS-CoV-2 infections detected in the sampled population of the residents of Vo′ in the first (x = 73) and in the second survey (x = 29, of which 8 were new infections).

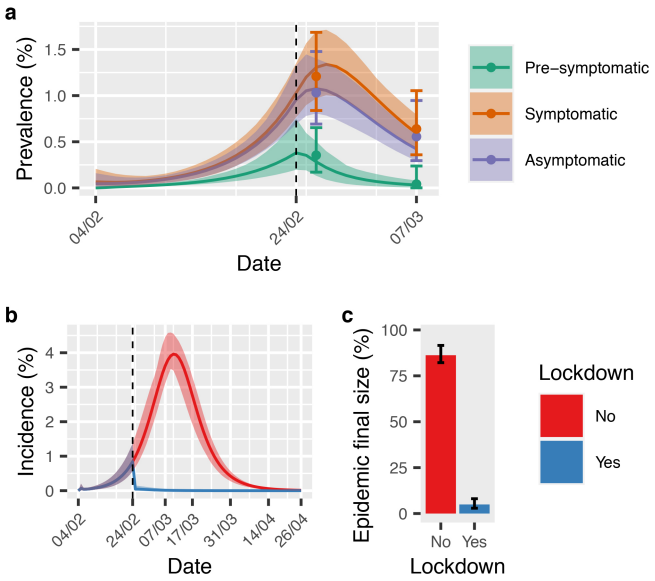


Fig. 3 | SARS-CoV-2 dynamics of the mitigated and counterfactual unmitigated epidemic in Vo' and relative final size estimates. (a) Prevalence of SARS-CoV-2 infection inferred from the observed prevalence data for symptomatic, pre-symptomatic and asymptomatic infections in the first and second survey using R_0^1 (the reproduction number before the lockdown) = 2.4 and 1/σ (the average duration of positivity beyond the duration of the infectious period) = 4 days. The dashed vertical line represents the time the lockdown started. The points represent the observed prevalence data, the 95% CI is the exact binomial confidence interval. The solid lines represent the mean and the shading represents the 95% Credible Interval (CrI) obtained from 100 samples from the posterior distribution of the parameters. (b) Incidence of the epidemic fitted to the prevalence data (blue) and of the unmitigated epidemic (red), obtained assuming the same initial reproduction number value $R_0^1 = 2.4$ throughout the whole epidemic and 1/σ = 4 days. The dashed vertical line represents the time the lockdown started. The solid lines represent the mean and the shading represent the 95% CrI obtained from 100 samples from the posterior distribution of the parameters. (c) Mean epidemic final size (the proportion of population infected at the end of the epidemic) of the counterfactual unmitigated epidemic (red) and of the epidemic fitted from the prevalence data with the lockdown (blue). Error bars represent the range (minimum to maximum) of the mean final size obtained from n = 100 independent samples drawn from the posterior distribution of the parameters, calculated over the models with DIC < 36.4.

Table 1 | Individuals positive for SARS-CoV-2 at the first and second survey

	First survey		Second survey	
	Total positives	**(%)**	**Total positives**	**(%)**
Symptomatic at the time of sampling*	34	(46.6)	15	(51.7)
Presymptomatic at the time of sampling	10	(13.7)	1	(3.4)
Asymptomatic§	29	(39.7)	13	(44.8)
Total	73		29	

*Defined as the presence of hospitalization and/or fever and/or cough and/or at least two of the following symptoms: sore throat, headache, diarrhoea, vomit, asthenia, muscle pain, joint pain, loss of taste or smell

^individuals testing negative for SARS-CoV-2 at the first survey.

Table 2 | Individuals tested and testing positive for SARS-CoV-2 at the first and second survey stratified by sex and by age groups

	First survey			Second survey				
	n	Positive	(%)	n	Positive	(%)	New positive	(%)
Sex								
Males	1408	43	(3.1)	1165	20	(1.7)	5	(0.4)
Females	1404	30	(2.1)	1178	9	(0.8)	3	(0.3)
p-value			0.15			0.041		
Age group								
00-10	217	0	(0.0)	157	0	(0.0)		(0.0)
11-20	250	3	(1.2)	210	2	(1.0)	1	(0.5)
21-30	240	4	(1.7)	191	2	(1.0)		(0.0)
31-40	286	7	(2.4)	241	2	(0.8)		(0.0)
41-50	439	5	(1.1)	366	2	(0.5)	1	(0.3)
51-60	496	16	(3.2)	439	7	(1.6)	2	(0.5)
61-70	384	15	(3.9)	349	6	(1.7)	2	(0.6)
71-80	318	19	(6.0)	262	6	(2.3)	2	(0.8)
81+	182	4	(2.2)	128	2	(1.6)		(0.0)
p-value			< 0.001*			0.48		
Total	2,812	73	(2.6)	2,343	29	(1.2)	8	(0.3)

P-values (two-sided) were computed using Fisher's exact test (for sex) and the likelihood ratio test (for age-group)
*Linear trend.

Methods

Study setting

The municipality of Vo', in the province of Padua, Veneto region, Italy, is located about 50 kilometers west of Venice (Figure 1a). The map shown in Figure 1 was produced using shapefiles from GADM (https://gadm.org/) and ISTAT (https://www.istat.it/it/archivio/222527 and https://www.istat.it/it/archivio/104317#accordions). According to the latest land registry, Vo' has a population of 3,275 individuals over an area of 20.4 square kilometers. Upon the detection of SARS-CoV-2 in a deceased resident of Vo' on 21 February, the same day where the first COVID-19 case was detected in Vo' and one day after the first locally acquired COVID-19 infection was identified in Italy, we conducted an epidemiological study to investigate the prevalence of SARS-CoV-2 infection in the population. Sampling was conducted on the majority of the Vo' population at two time points, the first during the days immediately after the detection of the first cases (21 – 29 February 2020) and the second one at the end of the two-weeks lockdown (07 March 2020) (Figure 1c). For each resident we collected information on the sampling dates, the results of SARS-CoV-2 testing, demographics (e.g. age and sex), residence, health record (including symptoms and COVID-19 related hospitalization dates, previous conditions and therapy taken for other illnesses), household size and contact network. The data were collated using Microsoft Excel and the dataset spreadsheet is available at https://github.com/ncov-ic/SEIR_Covid_Vo. **Definition of symptomatic**: a subject who required hospitalization and/or reported fever (yes/no or a temperature above 37 degrees Celsius) and/or cough and/or at least two of the following symptoms: sore throat, headache, diarrhoea, vomit, asthenia, muscle pain, joint pain, loss of taste or smell, shortness of breath.

Laboratory Methods

Upper respiratory tract samples were collected by healthcare professionals with a single flocked tapered swab used for the oropharynx then nasal mid-turbinates and immediately put into a sterile tube containing transport medium (eSwab®, Copan Italia Spa, Brescia, Italy). Sampling was performed according Centers for Disease Control and Prevention (CDC) guidelines[19]. Briefly, for oropharyngeal sampling, the swab was inserted into the posterior pharynx and tonsillar areas and rubbed over both tonsillar pillars and posterior oropharynx, avoiding touching the tongue, teeth, and gums; for deep nasal sampling, the swab was inserted into both nostrils for about 2 cm while gently rotating against the nasal wall several times. Samples were stored at 2-8 °C until testing, which was performed within 72 hours from collection. As a measure of the correct execution of the sampling, each PCR contains an internal control designed to amplify the human RNase P housekeeping gene. Reactions that failed to show the internal positive control were classified as invalid and repeated. Total nucleic acids were purified from 200 µL of nasopharyngeal swab samples and eluted in a final volume of 100 µL by using a MagNA Pure 96 System (Roche Applied Sciences, Basel, Switzerland). Detection of SARS-CoV-2 RNA was performed by an in house real-time RT-PCR method, which was developed according the protocol and the primers and probes designed by Corman *et al*.[20] targeting the envelope (*E*) (*E_Sarbeco_F*, *E_Sarbeco_R*, *E_Sarbeco_P1*) and RNA-dependent RNA-polymerase (*RdRp:RdRp_SARSr-F*, *RdRp_SARSr-R*, *RdRP_SARSr-P1*, and *RdRp_SARSr-P2*) genes of SARS-CoV-2. Real-time RT-PCR assays were performed in a final volume of 25 µL, containing 5 µL of purified nucleic acids, using One Step Real Time kit (Thermo Fisher Scientific, Waltham, MA, USA) and run on ABI 7900HT Fast Sequence Detection Systems (Thermo Fisher Scientific). The sensitivity of the *E* gene and *RdRp* gene assays was 5.0 and 50 genome equivalent copies per reaction at 95% detection probability, respectively. Both assays had no cross-reactivity with the endemic human coronaviruses HCoV-229E, -NL63, -OC43 and -HKU1 and with MERS-CoV. All tests were performed at the Clinical Microbiology and Virology Unit of Padova University Hospital, which is the Regional Reference Laboratory for emerging viral infections. After an initial period of dual testing by the National Reference Laboratory at the Italian Institute of Health (Istituto Superiore di Sanità), which demonstrated 100% agreement of results, the Regional Reference Laboratory received accreditation as Reference Laboratory for COVID-19 testing.

Assessment of genome equivalents

Ct (cycle threshold) data from real time RT-PCR assays were collected for *E* and *RdRp* genes. Ct data for gene *E* were available for 30 symptomatic, 5 pre-symptomatic and 23 asymptomatic infections and for gene *RdRp* for 27 symptomatic, 9 pre-symptomatic and 26 asymptomatic infections. Genome equivalent copies per ml were inferred according to linear regression performed on calibration standard curves. The interpolated Ct values were further multiplied by 100, according to the final dilution factor (1:100). Linear regression was calculated in Python3.7.3 using modules scipy 1.4.1, numpy 1.18.1, and matplotlib 3.2.1[21]. Genome equivalents distributions from the two genes, for positive symptomatic, asymptomatic and pre-symptomatic subjects were compared with the Exact Wilcoxon-Mann-Whitney test. Both viral load genome equivalents and raw Ct data are provided in the dataset.

Reconstructing transmission chains

We used data on contacts traced within the community and on household contacts derived from household composition data (reported in the dataset) to impute chains of transmission and transmission clusters. We used the R package epicontacts[22,23] to visualise the reconstructed transmission chains. We provide the algorithms used to infer the serial interval (the time from symptom onset of the infector to symptom onset of the infectee) distribution and the effective reproduction number (the average number of secondary infections generated by the identified infectors) in Supplementary Information Text S1 and S2, respectively. Briefly, we inferred the date of symptom onset for the subjects testing positive but with missing onset date from the observed time-lags from symptoms onset to confirmation (for the subjects testing positive at multiple sampling times, we used the first sampling time). We then used the observed and inferred dates of symptom onset alongside the contact information to infer transmission pairs within the sampled population. In turn, reconstructed transmission pairs were used to characterise the serial interval in the whole study period as well as during the pre- and post- lockdown periods. Central effective reproduction number estimates were calculated as the average number of secondary infections generated by observed or imputed infectors, having assigned the infector stochastically when more than one or no potential infectors were identified. The 95% confidence intervals were estimated by bootstrapping. All details are provided in Supplementary Information Text S1 and S2.

Mathematical modelling

The first survey occurred between 21st and 29th February 2020 and the second survey occurred on 7th March 2020. In the model we assumed that prevalence was taken on the weighted average of the first sample collection date, i.e. on 26th February 2020 and on 7th March 2020. The flow diagram of the model is given in Extended Data Figure 5. We assumed that the population of Vo' was fully susceptible to SARS-CoV-2 (S compartment) at the start of the epidemic. Upon infection, infected subjects incubate the virus (E compartment) and have undetectable viraemia for an average of 1/v days before entering a stage (TP compartment) that lasts an average of 1/δ days, in which subjects show no symptoms and have detectable viraemia. We assume that a proportion p of the infected population remains asymptomatic throughout the whole course of the infection (I_A compartment) and that the remaining proportion $1-p$ develops symptoms (I_S compartment). We assume that symptomatic (I_S), asymptomatic (I_A+pTP) and pre-symptomatic ((1−p) TP) subjects contribute to the onward transmission of SARS-CoV-2 and

that symptomatic, asymptomatic and pre-symptomatic subjects transmit the virus for an average of $1/\delta + 1/\gamma$ days. We further assume that the virus can be detected by swab testing beyond the duration of the infectious period; this assumption is compatible with the hypothesis that transmission occurs for viral loads above a certain threshold but the diagnostic test can detect the presence of virus below the threshold for transmission. Compartments TP_S and TP_A respectively represent symptomatic and asymptomatic subjects who are no longer infectious but have a detectable viral load, and hence test positive. Eventually, the viral load of all infections decreases below detection and subjects move into a test negative (TN) compartment. We assume a step change in the reproduction numberon the day that lockdown started. Before the implementation of quarantine the reproduction number is given by $R_0^1 = \beta\left(\frac{1}{\gamma} + \frac{1}{\delta}\right)$ and we assume that it drops to $R^2 = wR_0^1$ after the start of the lockdown, where $1 - w$ represents the percent reduction in R_0^1 due to the intervention. We let T_i denote the number of subjects swabbed on survey i ($i = 1, 2$) and let P_{Ai}, P_{Pi} and P_{Si} respectively denote the number of swabs testing positive among asymptomatic, pre-symptomatic (i.e. those showing no symptoms at the time of testing but developing symptoms afterwards) and symptomatic subjects, respectively. We assume that the number of positive swabs among symptomatic, pre-symptomatic and asymptomatic infections on survey i follows a binomial distribution with parameters T_i and Xi, where π_{Xi} represents the probability of testing positive on survey i for class X (= A,S). For symptomatic subjects, π_{Si} is given by $\pi_{Si} = \frac{I_S(t_i) + TP_S(t_i)}{N}$, for asymptomatic subjects π_{Ai} it is given by $\pi_{Ai} = \frac{pTP(t_i) + I_A(t_i) + TP_A(t_i)}{N}$ and for pre-symptomatic subjects π_{Pi} is given by $\pi_{Pi} = \frac{(1-p)TP(t_i)}{N}$, assuming perfect diagnostic sensitivity and specificity. The likelihood of the model is given by the product of the binomial distributions for symptomatic, pre-symptomatic and asymptomatic subjects at times t_i, $i = 1, 2$. Inference was conducted in a Bayesian framework, using the Metropolis-Hastings Markov Chain Monte Carlo (MCMC) method with uniform prior distributions[24]. We fixed the average generation time (equal to $1/\nu + 1/\delta + 1/\gamma$) to 7 days[20] and let the model infer $1/\nu$ and $1/\delta$. We explored the following values of R_0^1: 2.1, 2.4, 2.7, which are compatible with a doubling time of 3-4 days, as observed in Vo' and elsewhere in the Veneto region. We assumed that seeding of the infection occurred on 4 February 2020. We explored different scenarios on the average duration of viral detectability beyond the infectious period and fixed $1/\sigma$ to be 2, 4, 6, 8, 10 and 12 days. We estimate the number of infections introduced in the population from elsewhere at time t_0 (4 February 2020), the proportion of asymptomatic infections p, the average durations $1/\nu$, $1/\delta$ and $1/\gamma$ and the percent reduction in R_0^1 due to the interventions $(1 - w)100\%$.

Analysis of associations

We applied logistic regression to test the association between SARS-CoV-2 positivity (overall and at the first and second survey separately) with the age-group (10 years age bands, from 0 to 81+) and sex (male, female). We used Fisher's exact test for comparing two binomial proportions to assess whether there is an association between the presence of symptoms for 41 confirmed COVID-19 cases resident in Vo' and different types of comorbidities and treatments used. The analyses were repeated on the subset of patients who became negative at the second timepoint (results not shown). To increase the power of the data, we increased the sample size by including additional 11 confirmed COVID-19 cases resident in other villages close to Vo'. None of these scenarios provided significant associations at the 5% level.

Ethical approval statement

The first sampling of the Vo' population was conducted within the surveillance program established by the Veneto Region and did not require ethical approval; the second sampling was approved by the Ethics Committee for Clinical Research of the province of Padova. Study participation was by consent. For subjects under the age of 18 years, consent was provided by a parent or legal guardian.

Reporting summary

Further information on research design is available in the Nature Research Reporting Summary linked to this paper.

Data availability

The dataset is available at https://github.com/ncov-ic/SEIR_Covid_Vo.

Code availability

The code is available at https://github.com/ncov-ic/SEIR_Covid_Vo.

19. Centers for Disease Control and Prevention. Interim Guidelines for Collecting, Handling, and Testing Clinical Specimens from Persons for Coronavirus Disease 2019 (COVID-19). https://www.cdc.gov/coronavirus/2019-ncov/lab/guidelines-clinical-specimens.html [Accessed 18 May 2020]
20. Corman, V.M. et al. Detection of 2019 novel coronavirus (2019-nCoV) by real-time RT-PCR. *Eurosurveillance*, **25**(3):pii=2000045 (2020).
21. Python Software Foundation. Python Language Reference, version 2.7, http://www.python.org, (2020).
22. R Core Team. R: A language and environment for statisitical computing, https://www.R-project.org, (2020).
23. Nagraj, V. P. et al. epicontacts: Handling, Visualisation and Analysis of Epidemiological Contacts, (2020).
24. Robert C. The Metropolis-Hastings Algorithm. Wiley Stats Ref: Statistics Reference Online https://onlinelibrary.wiley.com/doi/full/10.1002/9781118445112.stat07834 (2015).

Acknowledgements We wish to thank the mayor of the municipality of Vo', G. Martini for his unreserved support throughout the study. A special thanks to the population of Vo' who volunteered en masse to this study. We thank M. Perilli and S. Guglielmo, M.D., for assistance in data collection and consistency check, F. Bosa and G. Rupolo from the Italian Red Cross for the support in patient samplings. This work was supported by the Veneto Region and was jointly funded by the UK Medical Research Council (MRC) [grant MR/R015600/1], the UK Department for International Development (DFID) under the MRC/DFID Concordat agreement, the Abdul Latif Jameel Foundation and is also part of the EDCTP2 programme supported by the European Union. I.D. acknowledges research funding from a Sir Henry Dale Fellowship funded by the Royal Society and Wellcome Trust [grant 213494/Z/18/Z]. C.C. acknowledges funding from Wellcome Trust [grant 203851/Z/16/Z]. L.O. and G.C.D. acknowledge research funding from The Royal Society. L.B., E.L. and S.T. acknowledge research funding from the European Union's Horizon 2020 research and innovation programme, under grant agreement No 874735 (VEO).

Author contributions A.C. conceived the project with input from E.L. and I.D. I.D. conceived the modelling with input from N.M.F. and C.A.D. E.L. coordinated data collection, curation and analyses. E.F. coordinated the diagnostic team and facilities. C.C. and G.C.D. are joint second authors. E.F., L.B., C.D.V., L.R., R.M., A.L., D.A.A., M.S., E.D., M.C.V., F.S., F.O., V.B., G.M., and M.T. performed laboratory testing on swabs and validated the results. E.L., S.T., V.B., A.S., N.N., and S.C. analysed the data, contributed to the discussion and commented on the manuscript. A.R.B., I.D. and C.A.D. performed statistical analyses. C.C, L.C., N.M.F. and I.D. developed the mathematical model. G.C.D., K.M.G., C.A.D and I.D. performed cluster analysis. E.L., M.N., F.C., G.C., E.N., B.L., L.F. and M.D. performed data collection, direct contacting of subjects at follow up and consistency check on metadata. S.M., R.S., G.C., D.D., and L.F. organized sampling logistics, S.M. and R.S. performed swab samplings. Imperial College London COVID-19 Response Team contributed to the discussion and background understanding of COVID-19 epidemiology. A.C. and I.D. wrote the manuscript, with contribution from E.L., L.B., V.B. and C.A.D. A.C. and I.D. are joint senior authors.

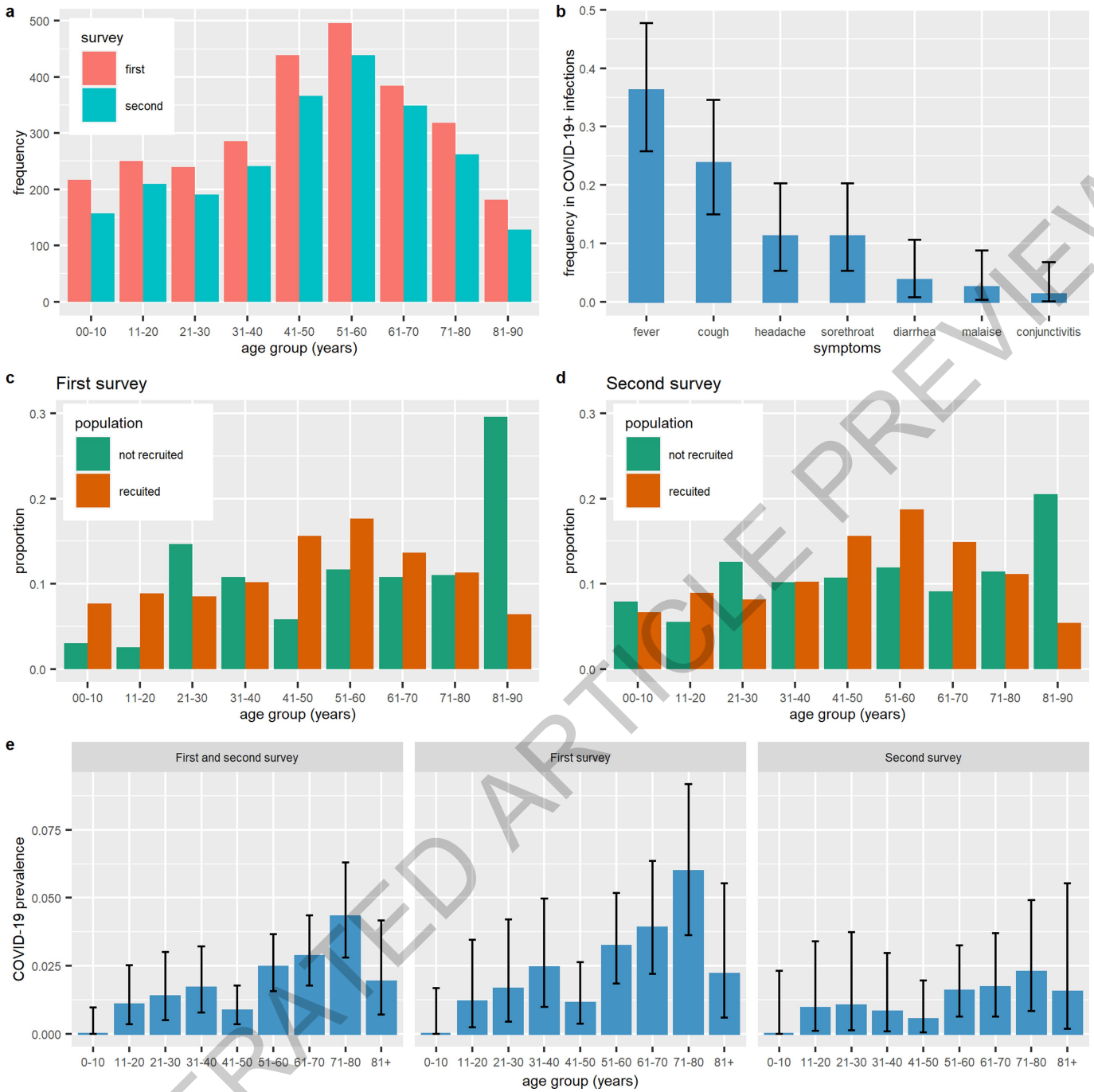
Competing interests The authors declare no competing interests.

Additional information

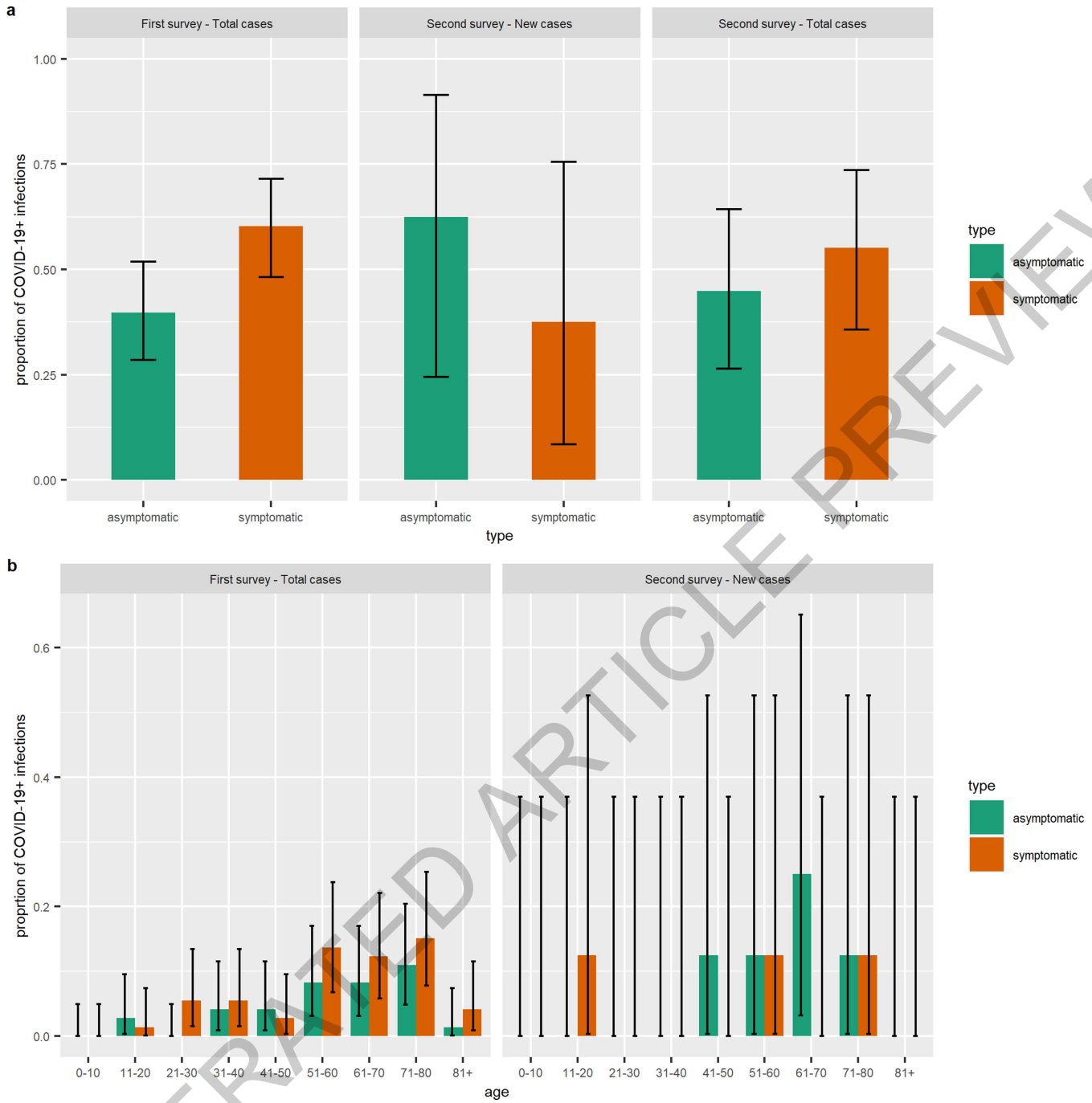
Supplementary information is available for this paper at https://doi.org/10.1038/s41586-020-2488-1.

Correspondence and requests for materials should be addressed to A.C. or I.D.

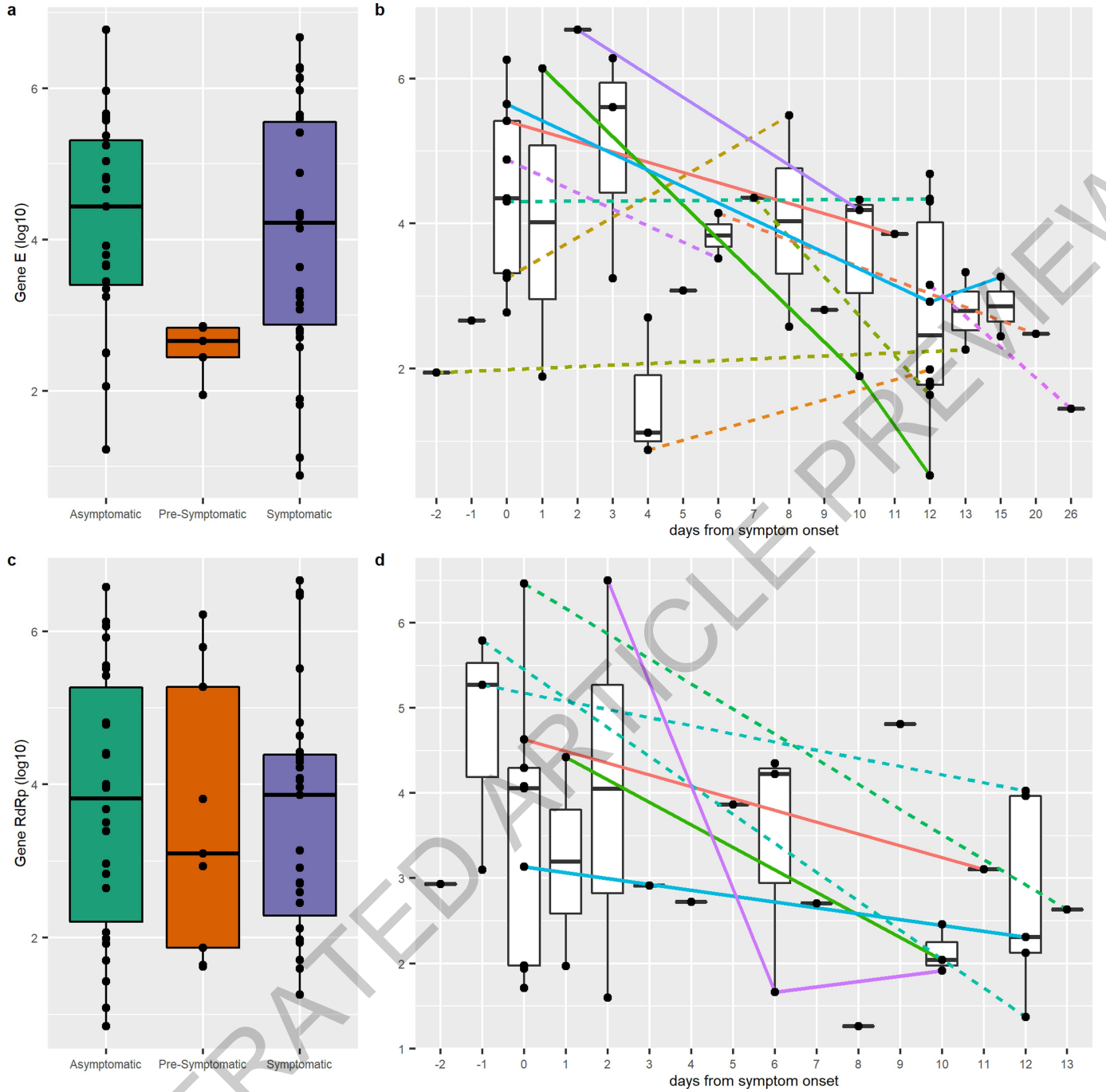
Peer review information *Nature* thanks Gabriel Leung, Malik Peiris and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.



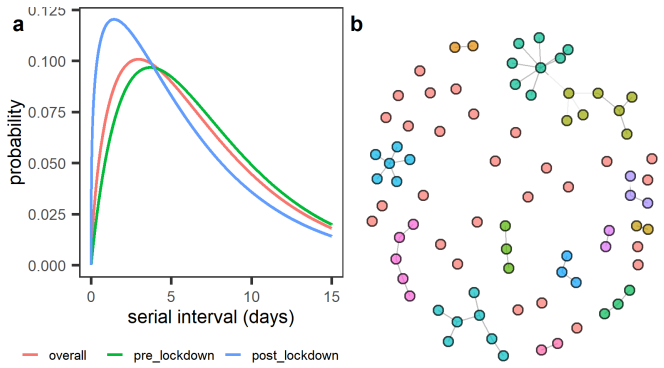
Extended Data Fig. 1 | Summary statistics, frequency of symptoms and prevalence by age. (a) age distributions (in years) of the subjects enrolled in the first and second survey; (b) frequency of individual symptoms (fever x = 29, cough x = 19, sore throat x = 9, headache x = 9, diarrhoea x = 3, malaise x = 2, conjunctivitis x = 1) among confirmed COVID-19 infected subjects across the whole study period (i.e. first and second survey aggregated, n = 80 subjects) with error bars representing the 95% exact binomial confidence interval; (c) age distribution of the population recruited and not recruited in the first survey; (d) age distribution of the population recruited and not recruited in the second survey; (e) SARS-CoV-2 prevalence by age at the first and second surveys combined (positive x = 0, 5, 6, 9, 7, 23, 21, 25, 6 tested n = 374, 460, 431, 527, 805, 935, 733, 580, 310 respectively in age groups 0-10, 11-20, 21-30, 31-40, 41-50, 51-60, 61-70, 71-80, 81+ years) and at the first (positive x = 0, 3, 4, 7, 5, 16, 15, 19, 4 tested n = 217, 250, 240, 286, 439, 496, 384, 318, 182 respectively in age groups 0-10, 11-20, 21-30, 31-40, 41-50, 51-60, 61-70, 71-80, 81+ years) and second (positive x = 0, 2, 2, 2, 2, 7, 6, 6, 2 tested n = 157, 210, 191, 241, 366, 439, 389, 262, 128 respectively in age groups 0-10, 11-20, 21-30, 31-40, 41-50, 51-60, 61-70, 71-80, 81+ years) surveys separately with error bars representing the 95% exact binomial confidence interval.



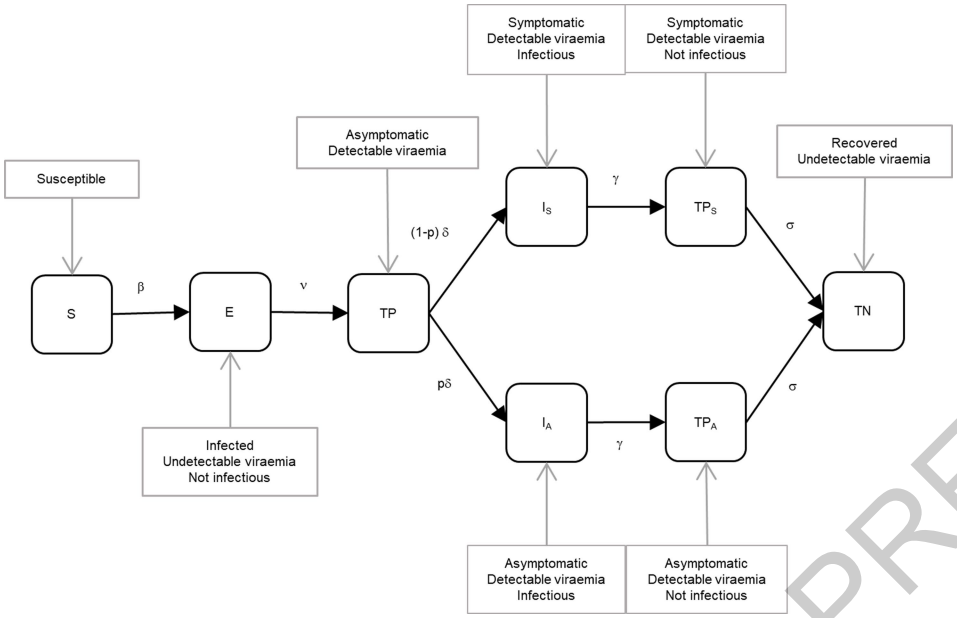
Extended Data Fig. 2 | Symptomatic and asymptomatic infection statistics. (a) Relative proportion of asymptomatic and symptomatic SARS-CoV-2 infections among the total number of positive swabs in the first survey (First survey – Total cases, asymptomatic x = 29, symptomatic x = 44, tested n = 73), second survey (Second survey – Total cases, asymptomatic x = 13, symptomatic x = 16, tested n = 29) and among the number of new positive swabs in the second survey (Second survey – New cases, asymptomatic x = 5, symptomatic x = 3, tested n = 8). Error bars represent the 95% exact binomial confidence interval; (b) Age distribution and relative proportion of asymptomatic and symptomatic COVID-19 positive infections among the total number of positive swabs in the first survey (First survey – Total cases, asymptomatic x = 0, 2, 0, 3, 3, 6, 6, 8, 1 symptomatic x = 0, 1, 4, 4, 2, 10, 9, 11, 3 respectively in age groups 0-10, 11-20, 21-30, 31-40, 41-50, 51-60, 61-70, 71-80, 81+ years; tested n = 73) and among the number of new positive swabs in the second survey (Second survey – New cases, asymptomatic x = 0, 0, 0, 0, 1, 1, 2, 1, 0, symptomatic x = 0, 1, 0, 0, 0, 1, 0, 1, 0 respectively in age groups 0-10, 11-20, 21-30, 31-40, 41-50, 51-60, 61-70, 71-80, 81+ years; tested n = 8). Error bars represent the 95% exact binomial confidence interval.



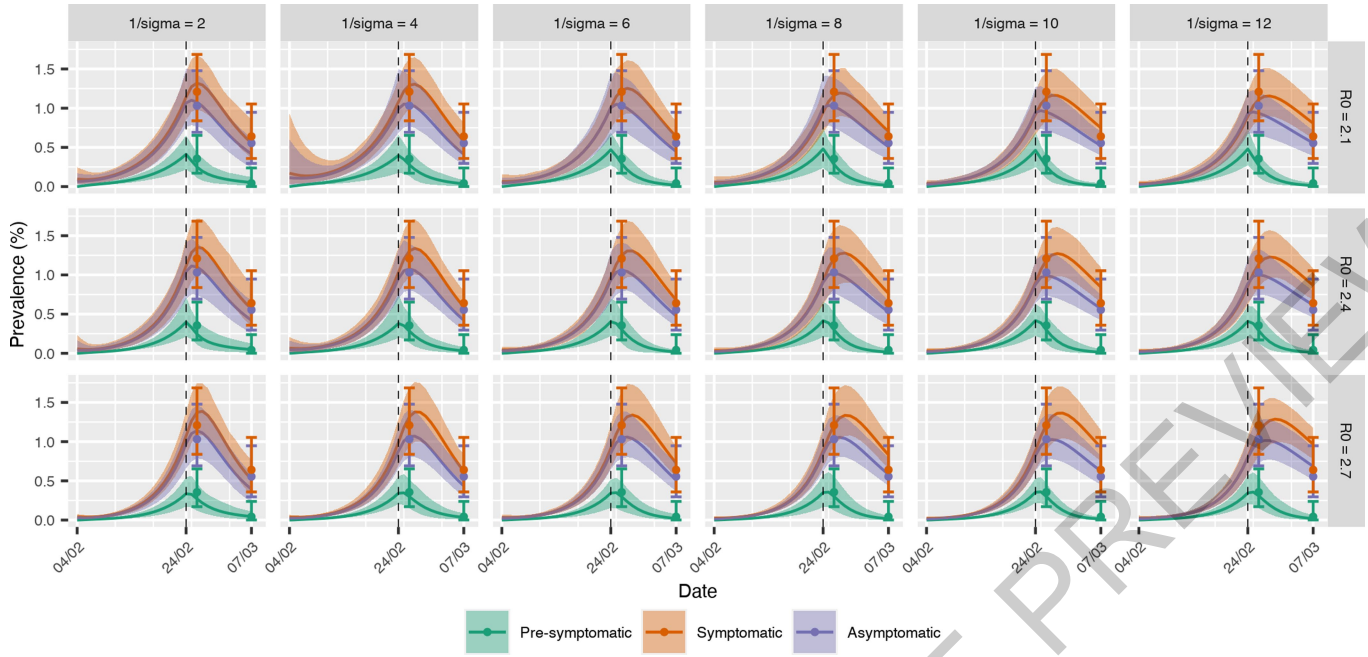
Extended Data Fig. 3 | Viral load for asymptomatic, pre-symptomatic and symptomatic infections and viral load dynamics relative to the number of days from symptom onset. (a) Median (solid line), interquartile (i.e. 25th–75th percentiles, box) and range (i.e. minimum – maximum, whiskers) of Gene *E* genome equivalent copies per ml (log10 scale, y axis) calculated from RT-PCR interpolated values (asymptomatic n = 23, pre-symptomatic n = 5 and symptomatic n = 30). The raw Ct data and the derived values of genome equivalent copies are provided in the dataset. (b) Median (solid line), interquartile (i.e. 25th–75th percentiles, box) and range (i.e. minimum – maximum, whiskers) of Gene *E* genome equivalent copies per ml (log10 scale, y axis) versus the number of days from symptom onset (days, x-axis); n = 34 subjects, lines in colour join measurements from the same subject; solid lines identify the 4 subjects with sequential viral load measurements both for Gene *E* and Gene *RdRp*. (c) Median (solid line), interquartile (i.e. 25th–75th percentiles, box) and range (i.e. minimum – maximum, whiskers) of Gene *RdRp* genome equivalent copies per ml (log10 scale, y axis) calculated from RT-PCR interpolated values (asymptomatic n = 26, pre-symptomatic n = 9 and symptomatic n = 27). The raw Ct data and the derived values of genome equivalent copies are provided in the dataset. (d) Median (solid line), interquartile (i.e. 25th–75th percentiles, box) and range (i.e. minimum – maximum, whiskers) of Gene *RdRp* genome equivalent copies per ml (log10 scale, y axis) versus the number of days from symptom onset (days, x-axis); n = 28 subjects, lines in colour join measurements from the same subject; solid lines identify the 4 subjects with sequential viral load measurements both for Gene *E* and Gene *RdRp*.



Extended Data Fig. 4 | Serial interval distribution and transmission chains. (a) Estimated serial interval distributions for the whole study period (overall) and for the pre-lockdown (before 24 February 2020) and post-lockdown (after 24 February 2020) periods. (b) Observed transmission clusters from reported and household contacts. Each nodes (circle) represents a positive infection, edges (line connecting nodes) connect positive infections that reported contacts or are household members; different colours represent different clusters of infection.



Extended Data Fig. 5 | Flow chart of the mathematical model fitted to the point prevalence data observed in Vo' at the first and second survey.



Extended Data Fig. 6 | SARS-CoV-2 dynamics in Vo' inferred from the fit of the dynamical model to the observed prevalence of symptomatic, pre-symptomatic and asymptomatic infections in the first and second survey. Each sub-panel represents the model fit using the specified values of R_0^1 (the reproduction number before the lockdown) and $1/\sigma$ (the average duration of positivity beyond the duration of the infectious period). The dashed vertical line represents the time lockdown started. The points represent the observed prevalence data, the 95% CI is the exact binomial confidence interval. The solid lines represent the mean and the shading represent the 95% CrI obtained from 100 samples from the posterior distribution of the parameters.

Extended Data Table 1 | Age distribution of Vo' residents and number of tested subjects at the two time-points across different age groups

Age group (years)	Resident subjects	First survey		Second survey	
		n.	(%)	n.	(%)
00-10	231	217	93,9	157	68,0
11-20	262	250	95,4	210	80,2
21-30	308	240	77,9	191	62,0
31-40	336	286	85,1	241	71,7
41-50	466	439	94,2	366	78,5
51-60	550	496	90,2	439	79,8
61-70	434	384	88,5	349	80,4
71-80	369	318	86,2	262	71,0
81+	319	182	57,1	128	40,1
total	3275	2812	85,9	2343	71,5

Extended Data Table 2 | Age distribution of symptomatic and asymptomatic individuals at the first and second surveys

Age group	Tested at first survey		Positive at first survey				Tested at second survey		Positive at second survey							
									Total cases				New cases only			
	Symp	Asymp	Symp^	(%)	Asymp^	(%)	Symp	Asymp	Symp	(%)	Asymp	(%)	Symp	(%)	Asymp	(%)
00-10	28	189	-	(-)	-	(-)	15	142	-	(-)	-	(-)	-	(-)	-	(-)
11-20	24	226	1	(4.2)	2 (1)	(0.9)	22	188	2	(9.1)	-	(-)	1	(4.5)	-	(-)
21-30	14	226	4 (2)	(28.6)	0	(-)	10	181	2	(20.0)	-	(-)	-	(-)	-	(-)
31-40	23	263	4	(17.4)	3	(1.1)	20	221	-	(-)	2	(0.9)	-	(-)	-	(-)
41-50	27	412	2	(7.4)	3 (1)	(0.7)	27	339	-	(-)	2	(0.6)	-	(-)	1	(0.3)
51-60	32	464	10	(31.3)	6 (1)	(1.3)	28	411	5	(17.9)	2	(0.5)	1	(3.6)	1	(0.2)
61-70	16	368	9	(56.3)	6	(1.6)	16	333	2	(12.5)	4	(1.2)	-	(-)	2	(0.6)
71-80	21	297	11 (1)	(52.4)	8 (1)	(2.7)	15	247	3	(20.0)	3	(1.2)	1	(6.7)	1	(0.4)
81+	8	174	3	(37.5)	1 (1)	(0.6)	8	120	2	(25.0)	-	(-)	-	(-)	-	(-)
Total	193	2619	44	(22.8)	29	(1.1)	161	2182	16	(9.9)	13	(0.6)	3	(1.9)	5	(0.2)

The symptomatic category includes both symptomatic and pre-symptomatic subjects. The percentages represent the proportions positives among those tested, i.e. the probability of testing positive given symptomatic or asymptomatic infection. Symp = symptomatic; Asymp = asymptomatic. ^Subjects not available at second survey are reported within parentheses.

Extended Data Table 3 | SARS-CoV-2 negative children living in households with infected relatives

	first survey	second survey
n (age group 0-10)	217	157
with positive cohabitant*	10	3
with positive relative not cohabitant§	2	0

*5 subjects are resident outside Vo', not included in the released dataset. §both subjects did not reside in Vo' and were not included in the released dataset.

Extended Data Table 4 | Results of the second survey for subjects living with or reporting close contacts with SARS-CoV-2 infected relatives

		Second survey			
		New cases		Negative	
		n.	(%)	n.	(%)
Subjects living with or reporting close contacts with infected relatives	Yes	6	(75.0)	78	(3.4)
	No	2	(25.0)	2197	(96.6)
Total		8		2275	

nature research

Corresponding author(s): Ilaria Dorigatti and Andrea Crisanti

Last updated by author(s): Jun 6, 2020

Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see our Editorial Policies and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed	
☐	☒	The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒	A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒	The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☐	☒	A description of all covariates tested
☐	☒	A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒	A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☐	☒	For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☐	☒	For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☐	☒	For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☐	☒	Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection	Data were collated using Microsoft Excel (version 16.16.6).
Data analysis	We developed code in the R programming language (version 3.6.0) using packages tidyr, dplyr, readr, janitor, lubridate, magrittr, ggplot2, plotly, htmlwidgets, epicontacts, igraph, GGally, epitrix, ggpubr, flextable, network, sna, scales, intergraph, cowplot, readxl, odin, stringr, ggrepel, gridExtra, sf, ggspatial and in Python (version 3.7.3) using modules scipy 1.4.1, numpy 1.18.1, and matplotlib 3.2.1. The code is available at https://github.com/ncov-ic/SEIR_Covid_Vo.

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

The dataset is available at https://github.com/ncov-ic/SEIR_Covid_Vo.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	Vo' has a resident population of 3,275 inhabitants. We collected nasopharyngeal swabs from 2,812 and 2,343 subjects in the first and second screening respectively, corresponding to 85.9% and 71.5% of the eligible population. No sample size calculation was performed, we aimed to recruit as many residents as possible.
Data exclusions	We excluded from the analysis the data collected on a small number of subjects, including 11 confirmed COVID-19 infections, who did not reside in Vo'.
Replication	Detection of SARS-CoV-2 RNA was performed by an in-house real-time RT-PCR method performed at the Clinical Microbiology and Virology Unit of Padova University Hospital, which is the Regional Reference Laboratory for emerging viral infections. The samples collected in the initial phase of the survey were validated by the National Reference Laboratory at the Italian Institute of Health (Istituto Superiore di Sanità) and demonstrated 100% agreement with the in-house assay. Given the 100% agreement on the samples collected in the initial phase and due to the large number of samples analyzed by the laboratory during the epidemic, we did not validate all samples collected in Vo' across the two surveys.
Randomization	Randomization is not relevant in our study, we aimed to enroll as many study participants as possible.
Blinding	Blinding is not relevant in our study, it was an observational study.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☒	☐ Antibodies
☒	☐ Eukaryotic cell lines
☒	☐ Palaeontology and archaeology
☒	☐ Animals and other organisms
☐	☒ Human research participants
☒	☐ Clinical data
☒	☐ Dual use research of concern

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☒	☐ Flow cytometry
☒	☐ MRI-based neuroimaging

Human research participants

Policy information about studies involving human research participants

Population characteristics	We collected information on sampling dates, results of SARS-CoV-2 testing, age, sex, symptoms, underlying health conditions, pharmacological therapy, hospitalization, household composition and contact network. The recruited subjects were between 1 month and 100 years of age and 49.9% were male and 50.1% were female. The underlying health conditions and pharmacological therapies of the recruited population at the time of the study are described in Supplementary Tables S3 and S4.
Recruitment	Study participation was by consent. For subjects under the age of 18 years, consent was provided by a parent or legal guardian. Participation in the study was publicized through local authorities. The age distribution of the recruited versus not recruited population was statistically different, as described in the main text and in Extended Data Figure 1.
Ethics oversight	The Ethics Committee for Clinical Research of the province of Padova approved the study.

Note that full information on the approval of the study protocol must also be provided in the manuscript.