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

Check for updates

Enrico Lavezzo[1,13], Elisa Franchin[1,13], Constanze Ciavarella[2], Gina Cuomo-Dannenburg[2], Luisa Barzon[1], Claudia Del Vecchio[1], Lucia Rossi[3], Riccardo Manganelli[1], Arianna Loregian[1], Nicolò Navarin[4,5], Davide Abate[1], Manuela Sciro[3], Stefano Merigliano[6], Ettore De Canale[3], Maria Cristina Vanuzzo[3], Valeria Besutti[3], Francesca Saluzzo[1], Francesco Onelia[1], Monia Pacenti[3], Saverio G. Parisi[1], Giovanni Carretta[3], Daniele Donato[3], Luciano Flor[3], Silvia Cocchio[7], Giulia Masi[1], Alessandro Sperduti[4,5], Lorenzo Cattarino[2], Renato Salvador[6], Michele Nicoletti[8], Federico Caldart[8], Gioele Castelli[8], Eleonora Nieddu[8], Beatrice Labella[8], Ludovico Fava[8], Matteo Drigo[8], Katy A. M. Gaythorpe[2], Imperial College COVID-19 Response Team*, Alessandra R. Brazzale[9], Stefano Toppo[1,5], Marta Trevisan[1], Vincenzo Baldo[7], Christl A. Donnelly[2,10], Neil M. Ferguson[2], Ilaria Dorigatti[2,14] ✉ & Andrea Crisanti[1,11,14] ✉

On 21 February 2020, a resident of the municipality of Vo', a small town near Padua (Italy), died of pneumonia due to severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) infection[1]. This was the first coronavirus disease 19 (COVID-19)-related death detected in Italy since the detection 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]. Here we collected information on the demography, clinical presentation, hospitalization, contact network and the 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. From 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%). From the second survey, which was conducted at the end of the lockdown, we found a prevalence of 1.2% (95% 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 (that is, 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 ($P = 0.62$ and 0.74 for *E* and *RdRp* genes, respectively, exact Wilcoxon–Mann–Whitney test). This study sheds light on the frequency of asymptomatic SARS-CoV-2 infection, their infectivity (as measured by the viral load) and provides insights into its transmission dynamics and the efficacy of the implemented control measures.

As of 23 May 2020, 5,105,881 confirmed cases and 333,446 deaths of COVID-19 have been reported worldwide[2]. In Italy, COVID-19 has caused more than 32,616 confirmed deaths. The causative agent (SARS-CoV-2), a close relative of SARS-CoV[4], was detected in the human population in 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 considerable 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 the transmission of SARS-CoV-2 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

[1]Department of Molecular Medicine, University of Padova, Padua, Italy. [2]MRC Centre for Global Infectious Disease Analysis, Imperial College London, London, UK. [3]Azienda Ospedale Padova, Padua, Italy. [4]Department of Mathematics "Tullio Levi-Civita", University of Padova, Padua, Italy. [5]CRIBI Biotech Center, University of Padova, Padua, Italy. [6]Department of Surgery, Oncology and Gastroenterology, University of Padova, Padua, Italy. [7]Department of Cardiac, Thoracic, Vascular Sciences and Public Health, University of Padova, Padua, Italy. [8]School of Medicine, University of Padova, Padua, Italy. [9]Department of Statistical Sciences, University of Padova, Padua, Italy. [10]Department of Statistics, University of Oxford, Oxford, UK. [11]Department of Life Sciences, Imperial College London, London, UK. [13]These authors contributed equally: Enrico Lavezzo, Elisa Franchin. [14]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; andrea.crisanti@unipd.it

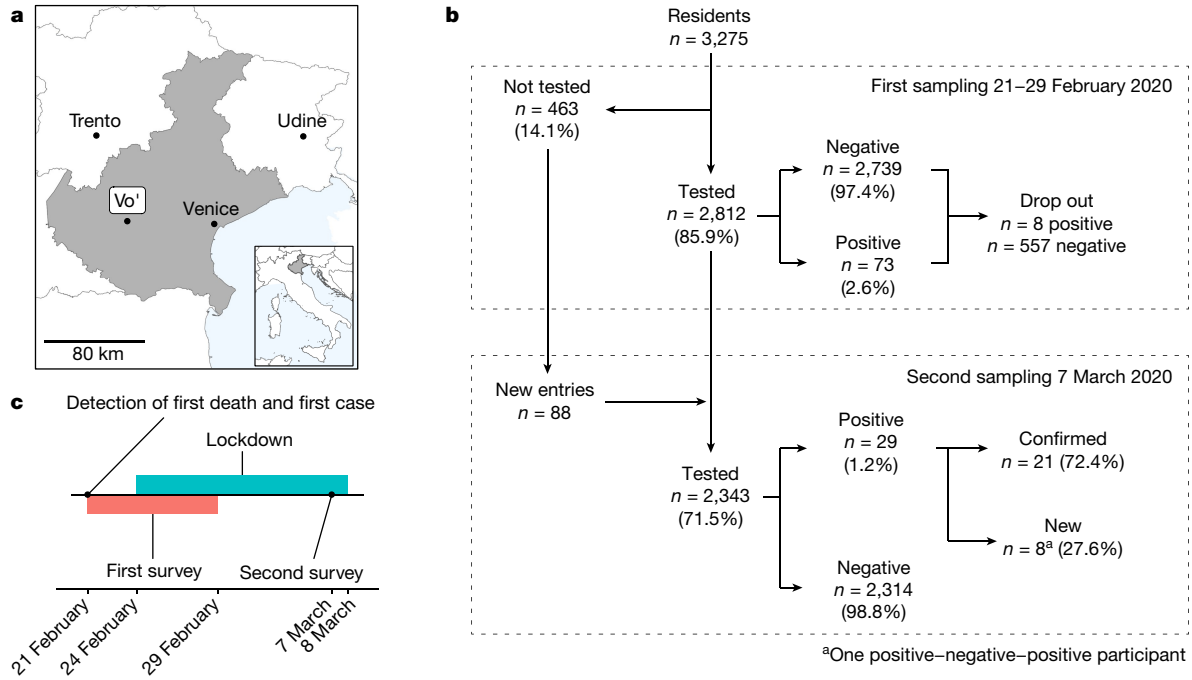


Fig. 1 | Study description. a, Map showing the location of Vo' and the Veneto region (grey area) within Italy, produced using shapefiles from GADM (https://gadm.org/) and Italian National Institute of Statistics (ISTAT; https://www.istat.it/it/archivio/222527 and https://www.istat.it/it/archivio/104317#accordions). **b**, Flow chart summarizing 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.

of asymptomatic, presymptomatic 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 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 2 weeks apart, to investigate population exposure to SARS-CoV-2 before and after the lockdown. We present an analysis of population demography, the prevalence of infection, viral load and the frequency of symptomatic, asymptomatic and presymptomatic infections. We assessed the risk of SARS-CoV-2 infection associated with comorbidity and therapies for underlying conditions, characterized 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.

During the two surveys, we collected nasopharyngeal swabs from 2,812 and 2,343 study participants, which corresponded to 85.9% and 71.5% of the eligible study population, respectively (Fig. 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 ($P < 0.001$ 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 ($P = 0.31$, exact Wilcoxon–Mann–Whitney test) (Extended Data Fig. 1). Notably, no additional infections were reported in Vo' despite the escalating epidemic in the surrounding regions.

Analysis of infection prevalence

A total of 73 out of the 2,812 participants who were 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 of 1.2%; 95% CI: 0.8–1.8%), 8 of which were new cases (prevalence of 0.3%; 95% CI: 0.15–0.7%) (Fig. 2). One of the eight new infections detected in the second survey was a hospitalized participant who tested positive, then negative, then positive again. It is unclear whether this was a case of SARS-CoV-2 re-infection or whether the second test was a false negative. The frequency of the symptoms in the participants who were positive for SARS-CoV-2 infection was systematically recorded, with fever and cough being the most common (Extended Data Fig. 1). Notably, a total of 29 out of the 73 participants (39.7%; 95% CI: 28.5–51.9%) who tested positive at the first survey were asymptomatic (that is, did not show symptoms at the time of swab sampling nor afterwards; see the definition of symptomatic in the Methods). A similar

Table 1 | Participants positive for SARS-CoV-2 at the first and second surveys

	First survey		Second survey	
	Total positives	Percentage	Total positives	Percentage
Symptomatic at the time of sampling[a]	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	

[a]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, and loss of taste or smell.

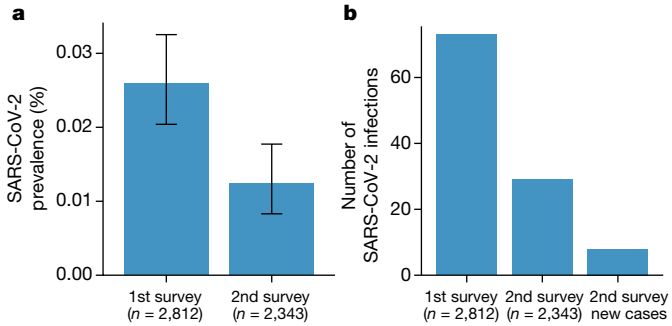


Fig. 2 | SARS-CoV-2 prevalence statistics. a, The prevalence of SARS-CoV-2 infection at the first survey ($x = 73$ positive out of $n = 2,812$ tested) and the second survey ($x = 29$ positive out of $n = 2,343$ tested). The error bars represent the 95% exact binomial CI. **b**, The number of SARS-CoV-2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

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 eight new cases, five were asymptomatic (Table 2, Extended Data Fig. 2). No infections were detected in either survey in 234 tested children ranging from 0 to 10 years of age, including those living in the same household as infected individuals (Extended Data Table 3). The prevalence of infection oscillated between a central estimate of 1.2% and 1.7% up to 50 years of age (Extended Data Fig. 1). Older participants showed a threefold increase in the infection prevalence (Table 2, Extended Data Fig. 1). Of the 81 participants who were positive for SARS-CoV-2 across the two surveys, 13 required hospitalization (16.0%). Their age distribution was as follows: 1 (7.7%) aged 41–50 years, 1 (7.7%) aged 51–60 years, 4 (30.8%) aged 61–70 years, 5 (38.5%) aged 71–80 years and 2 (15.4%) aged 81–90 years.

A substantial fraction of infected participants (58.9%; 95% CI: 46.8–70.3%, presymptomatic, symptomatic and asymptomatic combined over all ages) cleared the infection between the first and second surveys, that is, 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 that was conducted independently by the local health authority (data not shown). The time to viral clearance (the time from the earliest positive sample for the participants with more than one sample in the first survey to a negative sample in the second survey) ranged from 8 to 13 days and was on average 9.3 days, with a standard deviation of 2.0 days. The minimal duration of the positivity window (the time from the earliest positive sample in the first survey to a positive sample in the second survey) ranged from 3 to 13 days and was on average 9.1 days, with a standard deviation of 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 individuals with SARS-CoV-2 infections cleared the virus during the study period (that is, had a negative test after a positive result at the first survey); the highest rate (100%) was observed in the age groups of symptomatic individuals aged 31–40 and 41–50 years (Extended Data Table 2). SARS-CoV-2 positivity overall (that is, the first and second surveys combined) and at the first survey was more frequently associated with individuals who were 71–80 years of age (compared to those 21–30 years of age; $P = 0.012$ and $P = 0.017$, respectively) (Extended Data Fig. 1). Being male was associated with SARS-CoV-2 positivity in the second survey ($P = 0.04$) (Table 2). Analyses of the association between common comorbidities such as diabetes, hypertension, vascular diseases, respiratory diseases in asymptomatic and symptomatic people and the use of treatment for a number of different conditions with symptomatic infection showed no significant association (Supplementary Tables 3, 4).

Role of asymptomatic transmission

The analysis of viral genome equivalents inferred from cycle threshold data from real-time reverse transcription PCR (RT–PCR) assays indicated that asymptomatic and symptomatic participants did not differ when data from viral PCR templates recovered from the nasopharyngeal swabs of asymptomatic and symptomatic participants were compared ($P = 0.62$ and 0.74 for gene *E* and gene *RdRp*, respectively; exact Wilcoxon–Mann–Whitney test) (Extended Data Fig. 3). We found that the viral load tends to peak around the day of symptom onset and, for most of the participants, tends to decline after

Table 2 | Participants testing positive stratified by sex and age

	First survey			Second survey				
	n	Positive	Percentage	*n*	Positive	Percentage	New positive	Percentage
Sex								
Males	1,408	43	3.1	1,165	20	1.7	5	0.4
Females	1,404	30	2.1	1,178	9	0.8	3	0.3
P value			0.15			0.041		
Age group								
0–10	217	0	0.0	157	0	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	0.0
31–40	286	7	2.4	241	2	0.8	0	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	0.0
P value			<0.001[a]			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).
[a]Linear trend.

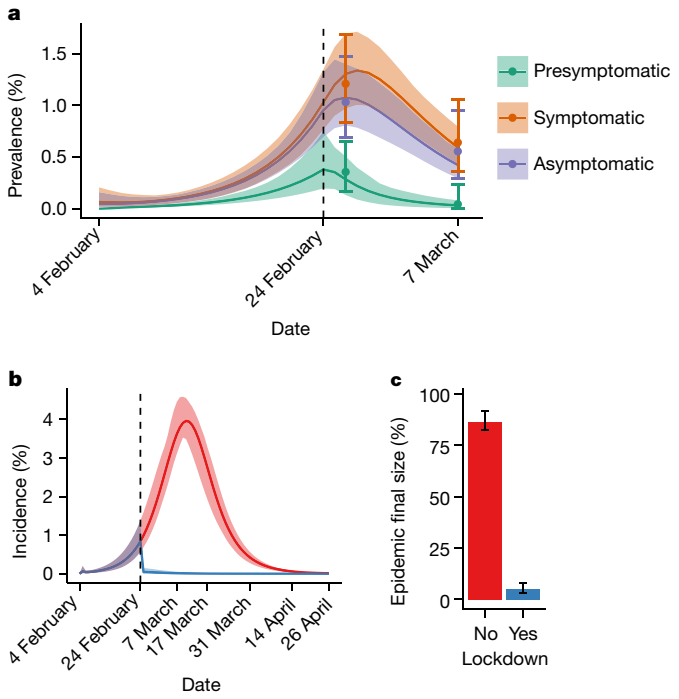


Fig. 3 | SARS-CoV-2 dynamics of the mitigated and counterfactual unmitigated epidemic in Vo' and the relative final size estimates. a, The prevalence of SARS-CoV-2 infection inferred from the observed prevalence data for symptomatic, presymptomatic and asymptomatic infections in the first and second surveys using R_0^1 (the reproduction number before the lockdown) = 2.4 and $1/\sigma$ (the average duration of positivity beyond the duration of the infectious period) = 4 days. The dashed vertical line represents the time that the lockdown started. The points represent the observed prevalence data, the 95% CI is the exact binomial CI. The solid lines represent the mean and the shading represents the 95% credible interval obtained with the model from 100 samples from the posterior distribution of the parameters. **b**, The 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 $1/\sigma = 4$ days. The dashed vertical line represents the time that the lockdown started. The solid lines represent the mean and the shading represents the 95% credible interval obtained with the model from 100 samples from the posterior distribution of the parameters. **c**, The mean epidemic final size (the proportion of the 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). The 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 (deviance information criterion) < 36.4.

symptom onset (Extended Data Fig. 3). The relative risk of contracting the infection from having close contacts with an infected relative, including those living in the same household, gives an odds ratio of 84.5 (95% CI: 16.8–425.4) (Extended Data Table 4, Supplementary Text 3). Two out of the eight participants with new infections that were detected in the second survey either shared a household or had a contact history with asymptomatic individuals (Supplementary Table 1).

Reconstructing transmission chains

From the inferred transmission pairs, we estimated a serial interval distribution over the whole study period with a mean of 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 (Extended Data Fig. 4). We also found that the lockdown substantially reduced transmission, with the reproduction number dropping from an initial value of 2.44 (95% CI: 1.30–3.91) before the lockdown to 0.41 (95% CI: 0.21–0.64) after the lockdown.

Modelling point prevalence data

We used the prevalence estimates obtained in Vo' at the first and second surveys to calibrate a modified susceptible–exposed–infectious–recovered compartmental model of SARS-CoV-2 transmission that incorporates symptomatic, presymptomatic and asymptomatic infections, virus detectability (in swabs) before and after the infectious period and the impact of the lockdown (Extended Data Fig. 5). We assumed that presymptomatic, 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 3.6–6.5 days, and that the lockdown reduced SARS-CoV-2 transmissibility on average by between 82% and 98%, depending on the assumed initial value of R_0^1 and on the duration of virus detectability (Supplementary Table 5). 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 (Fig. 3). These estimates are in line with the attack rates that were 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 (Extended Data Fig. 6, Supplementary Table 5). 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 is 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 (Supplementary Fig. 10; $P = 0.96$, Fisher's exact test). Among symptomatic participants, older age groups tended to show higher frequencies of SARS-CoV-2 infection (Extended Data Table 2). Recent studies have found that the clinical progression of infection in children is generally milder than in adults[14–16]. 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 odds ratio for adults to become infected when living together with family members who are positive for SARS-CoV-2. 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 infection in young children is not well understood[15]. 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 Fig. 3), by the model fit to the observed prevalence data (Extended Data Fig. 6, Supplementary Table 5) and by the observation that two out of the eight participants with new infections that were detected in the second survey reported contacts with asymptomatic individuals (Supplementary Text 3). The extent to which symptoms may promote viral

shedding remains to be determined, but the decreasing trend in viral load post-symptom onset suggests that presymptomatic transmission may play an important part[17]. Asymptomatic transmission and presymptomatic 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. & Cassone, A. In one Italian town, we showed mass testing could eradicate the coronavirus. *The Guardian* https://www.theguardian.com/commentisfree/2020/mar/20/eradicated-coronavirus-mass-testing-covid-19-italy-vo (2020).
2. World Health Organization. Novel coronavirus (COVID-19) situation. *WHO* https://covid19.who.int/ (2020).
3. Saini, V. Coronavirus: voices from a quarantined Italian town. *EU Observer* 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. *Imperial College London* https://www.imperial.ac.uk/media/imperial-college/medicine/mrc-gida/2020-02-15-COVID19-Report-5.pdf (2020).
6. Centre for Health Protection of the Hong Kong Special Administrative Region Government. CHP closely monitors cluster of pneumonia cases on Mainland. *The Government of the Hong Kong Special Administrative Region* 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. *Lancet Public Health* **5**, e261–e270 (2020).
8. Lai, S. et al. Effect of non-pharmaceutical interventions to contain COVID-19 in China. *Nature* https://doi.org/10.1038/s41586-020-2293-x (2020).
9. Anderson, R. M., Heesterbeek, H., Klinkenberg, D. & Hollingsworth, T. D. How will country-based mitigation measures influence the course of the COVID-19 epidemic? *Lancet* **395**, 931–934 (2020).
10. Flaxman, S. et al. Estimating the effects of non-pharmaceutical interventions on COVID-19 in Europe. *Nature*. https://doi.org/10.1038/s41586-020-2405-7 (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. *Imperial College London* 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. & Chowell, G. Estimating the asymptomatic proportion of coronavirus disease 2019 (COVID-19) cases on board the Diamond Princess cruise ship, Yokohama, Japan, 2020. *Euro Surveill*. **25**, 200018 (2020).
13. Gudbjartsson, D. F. et al. Spread of SARS-CoV-2 in the Icelandic population. *N. Engl. J. Med*. **382**, 2302–2315 (2020).
14. Cai, J. et al. A case series of children with 2019 novel coronavirus infection: clinical and epidemiological features. *Clin. Infect. Dis*. https://doi.org/10.1093/cid/ciaa198 (2020).
15. Zimmermann, P. & Curtis, N. Coronavirus infections in children including COVID-19: an overview of the epidemiology, clinical features, diagnosis, treatment and prevention options in children. *Pediatr. Infect. Dis. J*. **39**, 355–368 (2020).
16. Bi, Q. et al. Epidemiology and transmission of COVID-19 in 391 cases and 1286 of their close contacts in Shenzhen, China: a retrospective cohort study. *Lancet Infect. Dis*. https://doi.org/10.1016/S1473-3099(20)30287-5 (2020).
17. 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[2], Marc Baguelin[2], Samir Bhatt[2], Adhiratha Boonyasiri[2], Olivia Boyd[2], Lorenzo Cattarino[2], Constanze Ciavarella[2], Helen L. Coupland[2], Zulma Cucunubá[2], Gina Cuomo-Dannenburg[2], Bimandra A. Djafaara[2], Christl A. Donnelly[2,10], Ilaria Dorigatti[2], Sabine L. van Elsland[2], Rich FitzJohn[2], Seth Flaxman[12], Katy A. M. Gaythorpe[2], Will D. Green[2], Timothy Hallett[2], Arran Hamlet[2], David Haw[2], Natsuko Imai[2], Benjamin Jeffrey[2], Edward Knock[2], Daniel J. Laydon[2], Thomas Mellan[2], Swapnil Mishra[2], Gemma Nedjati-Gilani[2], Pierre Nouvellet[2], Lucy C. Okell[2], Kris V. Parag[2], Steven Riley[2], Hayley A. Thompson[2], H. Juliette T. Unwin[2], Robert Verity[2], Michaela A. C. Vollmer[2], Patrick G. T. Walker[2], Caroline E. Walters[2], Haowei Wang[2], Yuanrong Wang[2], Oliver J. Watson[2], Charles Whittaker[2], Lilith K. Whittles[2], Xiaoyue Xi[2] & Neil M. Ferguson[2]

[12]Department of Mathematics, Imperial College London, London, UK.

Methods

Study setting

The municipality of Vo', in the province of Padua, Veneto region, Italy, is located about 50 km west of Venice (Fig. 1a). According to the latest land registry, Vo' has a population of 3,275 individuals over an area of 20.4 km^2. Upon the detection of SARS-CoV-2 in a deceased resident of Vo' on 21 February 2020, the same day where the first COVID-19 case was detected in Vo' and 1 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 2-week lockdown (7 March 2020) (Fig. 1c). For each resident, we collected information on the sampling dates, the results of SARS-CoV-2 testing, demographics (for example, 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 data set spreadsheet is available at https://github.com/ncov-ic/SEIR_Covid_Vo. No statistical methods were used to predetermine sample size. The experiments were not randomized and the investigators were not blinded to allocation during experiments and outcome assessment.

The definition of symptomatic is as follows: a participant who required hospitalization and/or reported fever (yes/no or a temperature above 37 °C) 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, or 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). Sampling was performed according the US Centers for Disease Control and Prevention guidelines[18]. In brief, 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 h from collection. As a measure of the correct execution of the sampling, each PCR contains an internal control designed to amplify the human housekeeping gene encoding RNase P. 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). 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.[19] that targeted the genes encoding envelope (*E*) (E_Sarbeco_F, E_Sarbeco_R and E_Sarbeco_P1) and RNA-dependent RNA polymerase (*RdRp*: RdRp_SARSr-F, RdRp_SARSr-R, RdRP_SARSr-P1 and RdRp_SARSr-P2) 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) and run on ABI 7900HT Fast Sequence Detection Systems (Thermo Fisher Scientific). The sensitivity of the *E* 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, HCoV-NL63, HCoV-OC43 and HCoV-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

Cycle threshold (C_t) data from real-time RT–PCR assays were collected for *E* and *RdRp* genes. C_t data for the *E* gene were available for 30 symptomatic, 5 presymptomatic and 23 asymptomatic infections, and for the *RdRp* gene for 27 symptomatic, 9 presymptomatic and 26 asymptomatic infections. Genome equivalent copies per ml were inferred according to linear regression performed on calibration standard curves. The interpolated C_t 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[20]. Genome equivalent distributions from the two genes, for positive symptomatic, asymptomatic and presymptomatic participants were compared with the exact Wilcoxon–Mann–Whitney test. Both viral load genome equivalents and raw C_t data are provided in the data set.

Reconstructing transmission chains

We used data on contacts traced within the community and on household contacts derived from household composition data (reported in the data set) to impute chains of transmission and transmission clusters. We used the R package epicontacts[21,22] to visualize 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 1 and 2, respectively. In brief, we inferred the date of symptom onset for the participants who tested positive but with a missing onset date from the observed time-lags from symptom onset to confirmation (for the participants who tested 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 characterize the serial interval in the whole study period as well as during the pre-lockdown 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% CIs were estimated by bootstrapping. All details are provided in Supplementary Information Text 1 and 2.

Mathematical modelling

The first survey occurred between 21 and 29 February 2020 and the second survey occurred on 7 March 2020. In the model, we assumed that prevalence was taken on the weighted average of the first sample collection date, that is, on 26 February 2020 and on 7 March 2020. The flow diagram of the model is given in Extended Data Fig. 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 people incubate the virus (E compartment) and have undetectable viraemia for an average of $1/\nu$ days before entering a stage (TP compartment) that lasts an average of $1/\delta$ days, in which people 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 ($I_A + p$TP) and presymptomatic ($(1-p)$ TP) people contribute to the onward transmission of SARS-CoV-2 and that symptomatic, asymptomatic and presymptomatic people 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 people 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 people move into a test negative (TN) compartment. We assume a step change in the reproduction number on 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 per cent reduction in R_0^1 due to the intervention. We let T_i denote the number of participants 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, presymptomatic (that is, those showing no symptoms at the time of testing but develop symptoms afterwards) and symptomatic participants, respectively. We assume that the number of positive swabs among symptomatic, presymptomatic and asymptomatic infections on survey i follows a binomial distribution with parameters T_i and π_{Xi}, where π_{Xi} represents the probability of testing positive on survey i for X (where $X=$A, P, S). For symptomatic participants, π_{Si} is given by $\pi_{Si} = \frac{I_S(t_i)+\mathrm{TP_S}(t_i)}{N}$, for asymptomatic participants, π_{Ai} is given by $\pi_{Ai} = \frac{p\mathrm{TP}(t_i)+I_A(t_i)+\mathrm{TP_A}(t_i)}{N}$, and for presymptomatic participants, π_{Pi} is given by $\pi_{Pi} = \frac{(1-p)\mathrm{TP}(t_i)}{N}$, assuming perfect diagnostic sensitivity and specificity. The likelihood of the model is given by the product of the binomial distributions for symptomatic, presymptomatic and asymptomatic participants 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[23]. We fixed the average generation time (equal to $1/\nu+1/\delta+1/\gamma$) to 7 days[19] 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 per cent 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 surveys separately) with the age group (10 years of age bands, from 0 to >81 years of age) and sex (male and 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 who are resident in Vo' and the different types of comorbidities and treatments used. The analyses were repeated on the subset of patients who became negative at the second time point (results not shown). To increase the power of the data, we increased the sample size by including an additional 11 confirmed COVID-19 cases who were 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 programme 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 Padua. Study participation was by consent. For participants under 18 years of age, 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 data set is available at https://github.com/ncov-ic/SEIR_Covid_Vo. Queries can be addressed to A.C. (a.drcrisanti@imperial.ac.uk; andrea.crisanti@unipd.it) or I.D. (i.dorigatti@imperial.ac.uk).

Code availability

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

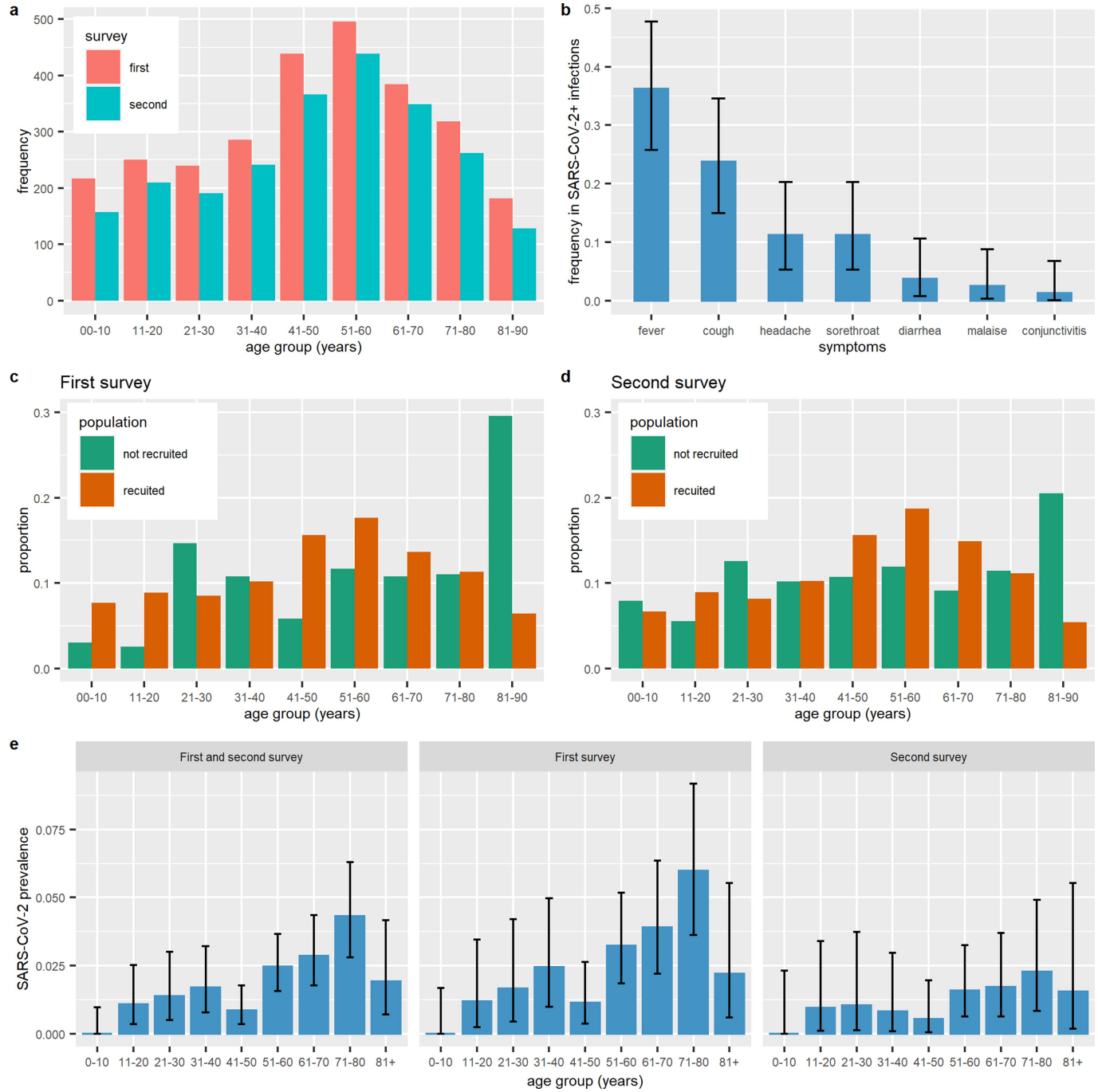
18. Centers for Disease Control and Prevention. Interim guidelines for collecting, handling, and testing clinical specimens from persons for coronavirus disease 2019 (COVID-19). *CDC* https://www.cdc.gov/coronavirus/2019-ncov/lab/guidelines-clinical-specimens.html (accessed 18 May 2020).
19. Corman, V. M. et al. Detection of 2019 novel coronavirus (2019-nCoV) by real-time RT-PCR. *Euro Surveill.* **25**, 2000045 (2020).
20. Python language reference, version 2.7 (Python Software Foundation, 2020).
21. R Core Team. R: A Language and Environment for Statistical Computing. http://www.R-project.org/ (R Foundation for Statistical Computing, 2020).
22. Nagraj, V. P. et al. epicontacts: handling, visualisation and analysis of epidemiological contacts. *F1000Res.* **7**, 566 (2018).
23. Robert, C. The Metropolis–Hastings algorithm. *Wiley StatsRef* https://doi.org/10.1002/9781118445112.stat07834 (2015).

Acknowledgements We 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; M. Perilli and S. Guglielmo for assistance in data collection and consistency check; and 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 the Wellcome Trust (grant 203851/Z/16/Z). L.C.O. from the Imperial College COVID-19 Response Team 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).

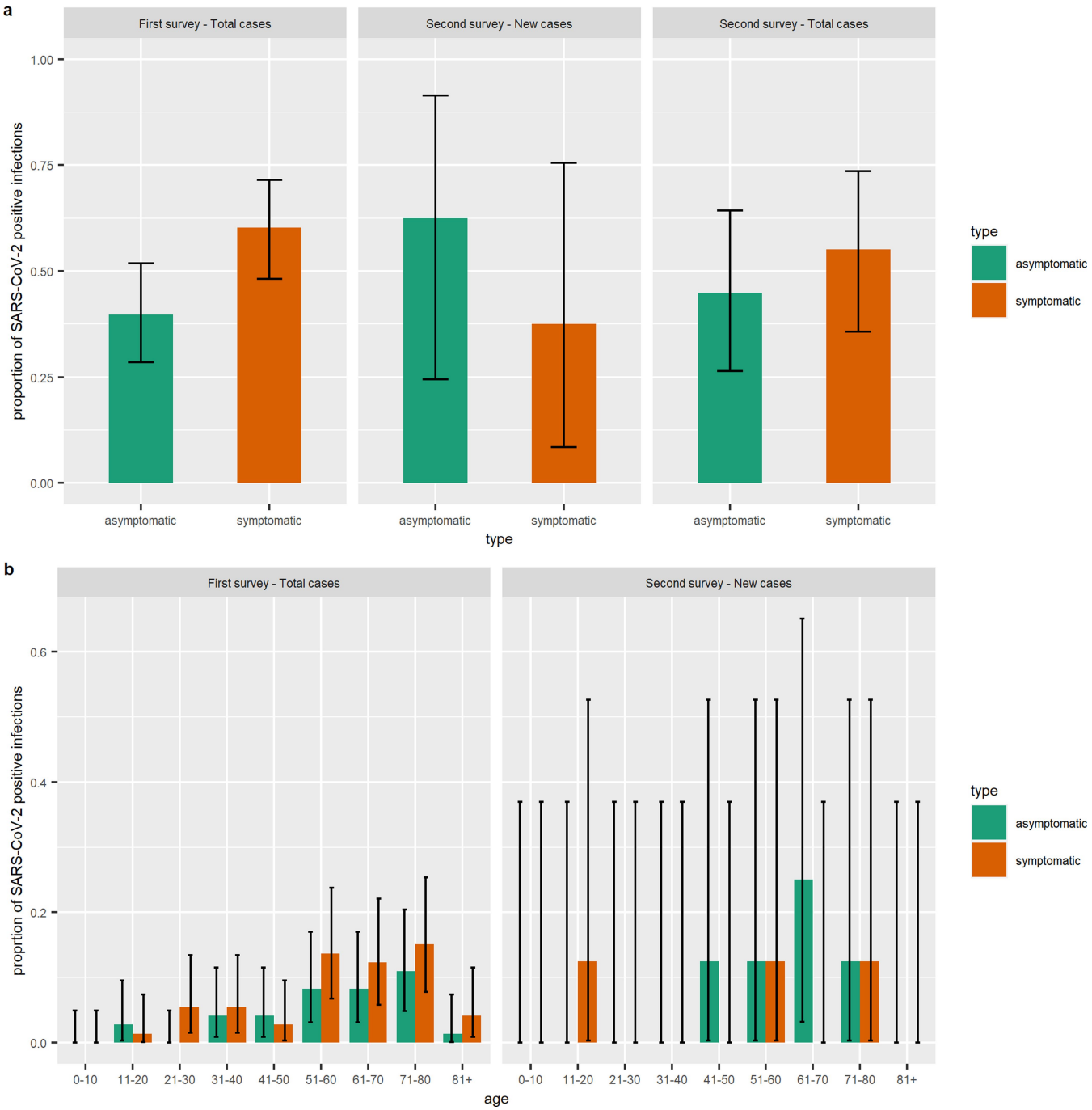
Author contributions A.C. conceived the project with input from E.L. and I.D. I.D. conceived the modelling with input from N.M.F. and C.A.D. E.L. coordinated data collection, curation and analyses. E.F. coordinated the diagnostic team and facilities. C.C. and G.C.-D. are joint second authors. E.F., L.B., C.D.V., L.R., R.M., A.L., D.A., M.S., E.D.C., M.C.V., F.S., F.O., V. Besutti, M.P., S.G.P., G.M. and M.T. performed laboratory testing on swabs and validated the results. E.L., S.T., V. Baldo, 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 the statistical analyses. C.C., L.C., N.M.F. and I.D. developed the mathematical model. G.C.-D., K.A.M.G., C.A.D. and I.D. performed cluster analysis. E.L., M.N., F.C., G. Castelli, E.N., B.L., L. Fava and M.D. performed data collection, direct contacting of study participants at follow up and consistency check on metadata. S.M., R.S., G. Carretta, D.D. and L. Flor organized sampling logistics. S.M. and R.S. performed swab samplings. The Imperial College 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. Baldo and C.A.D.

Competing interests The authors declare no competing interests.

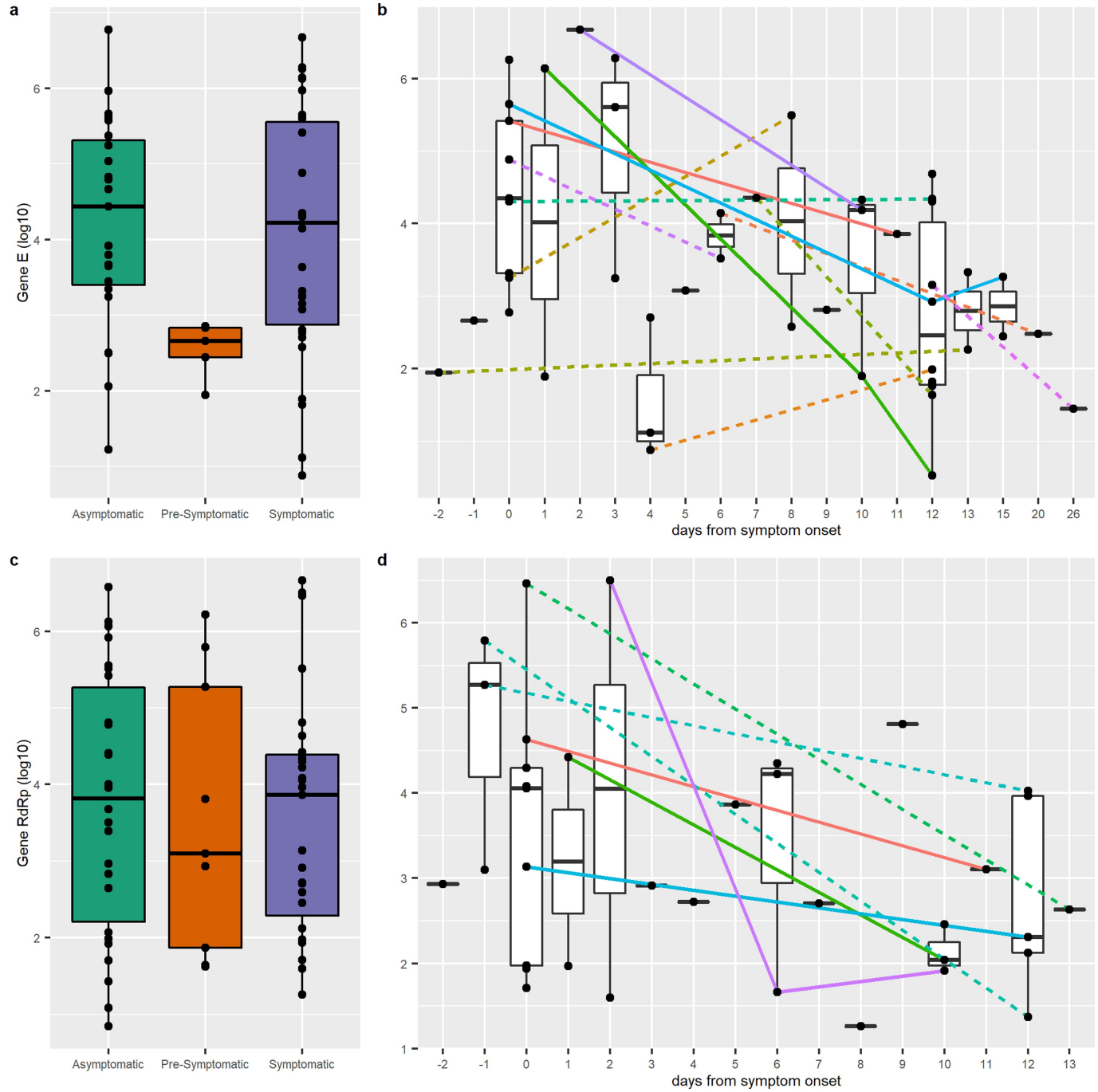
Additional information
Supplementary information is available for this paper at https://doi.org/10.1038/s41586-020-2488-1.
Correspondence and requests for materials should be addressed to I.D. or A.C.
Peer review information *Nature* thanks Gabriel Leung, Malik Peiris and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.
Reprints and permissions information is available at http://www.nature.com/reprints.



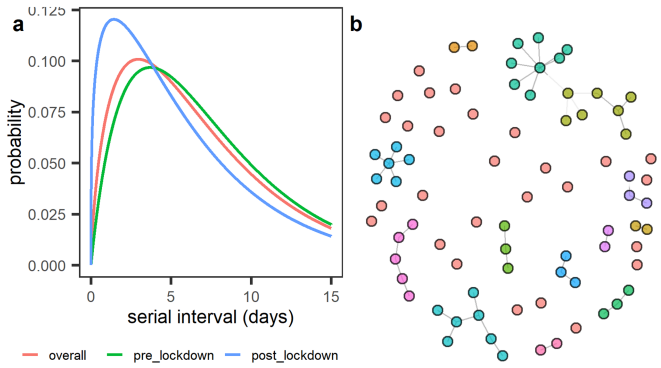
Extended Data Fig. 1 | Summary statistics, frequency of symptoms and prevalence by age. a, Age distributions (in years) of the participants enrolled in the first and second surveys. **b**, Frequency of individual symptoms (fever $x = 29$, cough $x = 19$, sore throat $x = 9$, headache $x = 9$, diarrhoea $x = 3$, malaise $x = 2$ and conjunctivitis $x = 1$) among participants with confirmed SARS-CoV-2 infection across the whole study period (that is, the first and second surveys aggregated; $n = 80$ participants). The error bars represent the 95% exact binomial CI. **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 $x = 0, 5, 6, 9, 7, 23, 21, 25$ and 6, tested $n = 374, 460, 431, 527, 805, 935, 733, 580$ and 310, respectively, in age groups 0–10, 11–20, 21–30, 31–40, 41–50, 51–60, 61–70, 71–80 and 81+ years) and at the first (positive $x = 0, 3, 4, 7, 5, 16, 15, 19$ and 4, tested $n = 217, 250, 240, 286, 439, 496, 384, 318$ and 182, respectively, in age groups 0–10, 11–20, 21–30, 31–40, 41–50, 51–60, 61–70, 71–80 and 81+ years) and second (positive $x = 0, 2, 2, 2, 2, 7, 6, 6$ and 2, tested $n = 157, 210, 191, 241, 366, 439, 389, 262$ and 128, respectively, in age groups 0–10, 11–20, 21–30, 31–40, 41–50, 51–60, 61–70, 71–80 and 81+ years) surveys separately. The error bars represent the 95% exact binomial CI.



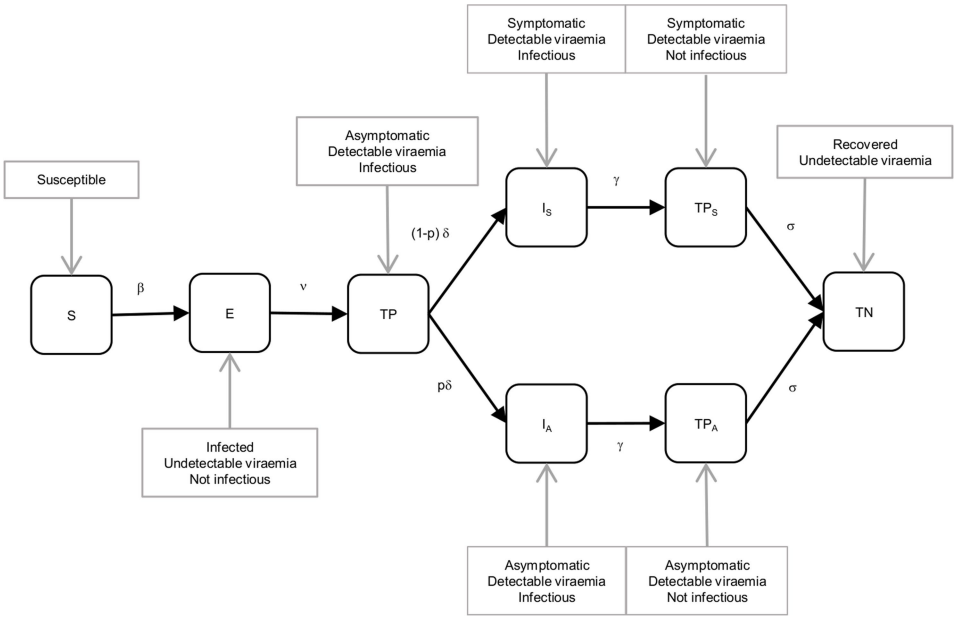
Extended Data Fig. 2 | Symptomatic and asymptomatic infection statistics. **a**, Relative proportion of asymptomatic and symptomatic SARS-CoV-2 infections among the total number of positive swabs in the first survey (first survey – total cases; asymptomatic $x = 29$, symptomatic $x = 44$, tested $n =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



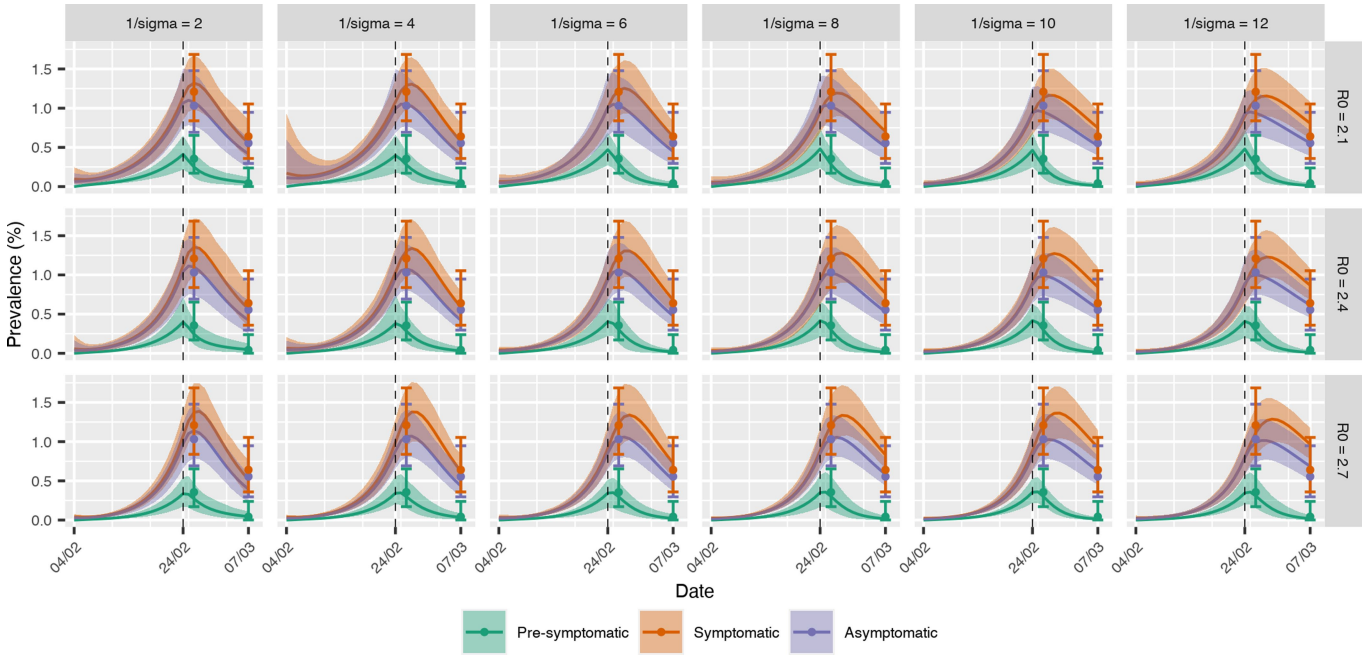
Extended Data Fig. 3 | Viral load for asymptomatic, pre-symptomatic and symptomatic infections and viral load dynamics relative to the number of days from symptom onset. a, The median (solid line), the interquartile range (that is, 25th to 75th percentiles (box)) and the range (that is, minimum to maximum (whiskers)) of gene *E* genome equivalent copies per ml ($\log_{10}$ scale, *y* axis) calculated from RT–PCR interpolated values (asymptomatic $n = 23$, pre-symptomatic $n = 5$ and symptomatic $n = 30$). The raw C_t data and the derived values of the genome equivalent copies are provided in the data set. **b**, The median (solid line), the interquartile range (that is, 25th to 75th percentiles (box)) and the range (that is, minimum to maximum (whiskers)) of gene *E* genome equivalent copies per ml ($\log_{10}$ scale, *y* axis) versus the number of days from symptom onset (days, *x* axis); $n = 34$ participants. The lines in colour join measurements from the same participant. The solid lines identify the four participants with sequential viral load measurements for both gene *E* and gene *RdRp*. **c**, The median (solid line), the interquartile range (that is, 25th to 75th percentiles (box)) and the range (that is, minimum to maximum (whiskers)) of *RdRp* genome equivalent copies per ml ($\log_{10}$ scale, *y* axis) calculated from RT–PCR interpolated values (asymptomatic $n = 26$, pre-symptomatic $n = 9$ and symptomatic $n = 27$). The raw C_t data and the derived values of genome equivalent copies are provided in the data set. **d**, The median (solid line), the interquartile range (that is, 25th to 75th percentiles (box)) and the range (that is, minimum to maximum (whiskers)) of *RdRp* genome equivalent copies per ml ($\log_{10}$ scale, *y* axis) versus the number of days from symptom onset (days, *x* axis); $n = 28$ participants. The lines in colour join measurements from the same participant. The solid lines identify the four participants with sequential viral load measurements for both 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 node (circle) represents a positive infection, and the edges (the line connecting the nodes) connect positive infections that reported contacts or are household members. The 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 surveys. Further details are provided in the Methods.



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 surveys. 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 that lockdown started. The points represent the observed prevalence data; the 95% CI is the exact binomial CI. The solid lines represent the mean and the shading represents the 95% credible interval obtained from 100 samples from the posterior distribution of the parameters.

Extended Data Table 1 | Age distribution of Vo' residents and the number of tested participants 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 (Symp) category includes both symptomatic and pre-symptomatic participants. The percentages represent the proportions positive among those tested; that is, the probability of testing positive given symptomatic or asymptomatic (Asymp) infection.

^Participants not available at the second survey are reported within parentheses.

Extended Data Table 3 | Children negative for SARS-CoV-2 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

*Five participants are residents outside Vo' and are not included in the released data set.
§Both participants did not reside in Vo' and were not included in the released data set.

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

		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.