nature medicine

Check for updates

Safety, tolerability and viral kinetics during SARS-CoV-2 human challenge in young adults

Ben Killingley[1,16], Alex J. Mann [2,16], Mariya Kalinova[2], Alison Boyers[2], Niluka Goonawardane[3], Jie Zhou[3], Kate Lindsell[4], Samanjit S. Hare[5], Jonathan Brown [3], Rebecca Frise[3], Emma Smith[6], Claire Hopkins[7], Nicolas Noulin[2], Brandon Löndt[2], Tom Wilkinson[8], Stephen Harden[9], Helen McShane [10], Mark Baillet[11], Anthony Gilbert[4], Michael Jacobs[12], Christine Charman[4], Priya Mande[4], Jonathan S. Nguyen-Van-Tam[13], Malcolm G. Semple [14], Robert C. Read [8], Neil M. Ferguson [15], Peter J. Openshaw [6], Garth Rapeport [6], Wendy S. Barclay [3], Andrew P. Catchpole[2] and Christopher Chiu [3] ✉

Since its emergence in 2019, severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) has caused hundreds of millions of cases and continues to circulate globally. To establish a novel SARS-CoV-2 human challenge model that enables controlled investigation of pathogenesis, correlates of protection and efficacy testing of forthcoming interventions, 36 volunteers aged 18–29 years without evidence of previous infection or vaccination were inoculated with 10 $TCID_{50}$ of a wild-type virus (SARS-CoV-2/human/GBR/484861/2020) intranasally in an open-label, non-randomized study (ClinicalTrials.gov identifier NCT04865237; funder, UK Vaccine Taskforce). After inoculation, participants were housed in a high-containment quarantine unit, with 24-hour close medical monitoring and full access to higher-level clinical care. The study's primary objective was to identify an inoculum dose that induced well-tolerated infection in more than 50% of participants, with secondary objectives to assess virus and symptom kinetics during infection. All pre-specified primary and secondary objectives were met. Two participants were excluded from the per-protocol analysis owing to seroconversion between screening and inoculation, identified post hoc. Eighteen (~53%) participants became infected, with viral load (VL) rising steeply and peaking at ~5 days after inoculation. Virus was first detected in the throat but rose to significantly higher levels in the nose, peaking at ~8.87 $\log_{10}$ copies per milliliter (median, 95% confidence interval (8.41, 9.53)). Viable virus was recoverable from the nose up to ~10 days after inoculation, on average. There were no serious adverse events. Mild-to-moderate symptoms were reported by 16 (89%) infected participants, beginning 2–4 days after inoculation, whereas two (11%) participants remained asymptomatic (no reportable symptoms). Anosmia or dysosmia developed more slowly in 15 (83%) participants. No quantitative correlation was noted between VL and symptoms, with high VLs present even in asymptomatic infection. All infected individuals developed serum spike-specific IgG and neutralizing antibodies. Results from lateral flow tests were strongly associated with viable virus, and modeling showed that twice-weekly rapid antigen tests could diagnose infection before 70–80% of viable virus had been generated. Thus, with detailed characterization and safety analysis of this first SARS-CoV-2 human challenge study in young adults, viral kinetics over the course of primary infection with SARS-CoV-2 were established, with implications for public health recommendations and strategies to affect SARS-CoV-2 transmission. Future studies will identify the immune factors associated with protection in those participants who did not develop infection or symptoms and define the effect of prior immunity and viral variation on clinical outcome.

C oronavirus disease 2019 (COVID-19) is a complex clinical syndrome caused by SARS-CoV-2. Despite extensive research into severe disease of hospitalized patients[1] and many large studies leading to approval of vaccines and antivirals[2–4], the global spread of SARS-CoV-2 continues and is, indeed, accelerating in many regions. Infections are typically mild or asymptomatic

in younger people, but these likely drive community transmission[5], and the detailed time course of infection and infectivity in this context has not been fully elucidated[6,7]. Deliberate human infection of low-risk volunteers enables the exact longitudinal measurement of viral kinetics, immunological responses, transmission dynamics and duration of infectious shedding after a fixed dose of

[1]Department of Infectious Diseases, University College London Hospital, London, UK. [2]hVIVO Services Ltd., London, UK. [3]Department of Infectious Disease, Imperial College London, London, UK. [4]UK Vaccine Taskforce, Department for Business, Energy and Industrial Strategy, London, UK. [5]Department of Radiology, Royal Free London NHS Foundation Trust, London, UK. [6]National Heart and Lung Institute, Imperial College London, London, UK. [7]ENT Department, Guy's and St. Thomas' NHS Foundation Trust, London, UK. [8]Faculty of Medicine and Institute for Life Sciences, University of Southampton, and NIHR Southampton Biomedical Research Centre, University Hospital Southampton, Southampton, UK. [9]Department of Radiology, Southampton General Hospital, Southampton, UK. [10]Department of Paediatrics, University of Oxford, Oxford, UK. [11]S-cubed Biometrics, Abingdon, UK. [12]Department of Infectious Diseases, Royal Free London NHS Foundation Trust, London, UK. [13]Division of Epidemiology and Public Health, University of Nottingham School of Medicine, Nottingham, UK. [14]Health Protection Research Unit in Emerging and Zoonotic Infections, Institute of Infection, Veterinary and Ecological Sciences, Faculty of Health and Life Sciences, University of Liverpool, Liverpool; Respiratory Department, Alder Hey Children's Hospital, Liverpool, UK. [15]MRC Centre for Global Infectious Disease Analysis, School of Public Health, Imperial College London, London, UK. [16]These authors contributed equally: Ben Killingley, Alex Mann. ✉e-mail: c.chiu@imperial.ac.uk

a well-characterized virus[8–14]. Under these tightly controlled conditions, host factors leading to differences in clinical outcome can be robustly inferred. Although many human infection challenge models have been successfully established during inter-pandemic times, none has been successfully established during a pandemic[15], and no recent reports of coronavirus (including SARS-CoV-2) human challenge exist.

Experimental challenge with human pathogens requires careful ethical scrutiny and regulation but can deliver unparalleled information that may inform clinical policy and refinement of infection control measures, enabling the rapid evaluation of vaccines, therapeutics and diagnostics. Invaluable information, including unique assessments of immunity at the time of virus exposure, responses during the pre-symptomatic period and protective correlates in individuals who resist symptomatic infection, can be obtained using small numbers of participants under highly controlled settings, potentially leading to wider societal benefits that offset the personal risks undertaken by the volunteers[16]. Recognizing the potential benefits that might be derived from SARS-CoV-2 human challenge, the World Health Organization convened working groups early during the COVID-19 pandemic to consider the necessary ethical and practical frameworks[17]. The pros and cons of human infection challenge studies have been extensively reviewed elsewhere[18], but the key considerations underlying these studies during an active pandemic were to balance scientific and public health benefits with ensuring that any risks to study participants (known and as yet uncertain) were minimized and managed.

The unique strengths of SARS-CoV-2 human challenge experiments are the ability to standardize the viral inoculum, study conditions and exact timing of exposure, thus controlling for factors that unavoidably confound natural infection studies. This contrasts with even the most well-controlled field trials, including household contact studies. There, the viral quasi-species, inoculum dose, timing and conditions of exposure are unknown, and contacts are identified only after diagnosis of the index case, at which time secondary exposure has almost always already occurred[19], thus missing transmission as well as the early phase of infection. SARS-CoV-2 human challenge studies, therefore, fill a gap in the understanding of early factors involved in susceptibility to infection that cannot be addressed in other ways. With continuing infection and re-infection with SARS-CoV-2, such translational studies that can inform public health strategy, while accelerating access to improved interventions through improved mechanistic understanding and early proof-of-concept efficacy testing, remain a priority that justifies this approach[20].

Here we report results from the first volunteers inoculated with SARS-CoV-2 in a human challenge study, the primary objective of which was to identify an inoculum dose that induced well-tolerated infection in more than 50% of participants, with secondary objectives to assess virus and symptom kinetics during infection. Our findings demonstrate the feasibility of deliberate infection with SARS-CoV-2, with no evidence of major adverse events in these carefully selected, healthy, young adult volunteers, and provide insights into the early dynamics of infection.

Results

Thirty-six healthy volunteers aged 18–29 years were enrolled according to protocol-defined inclusion and exclusion criteria (see the Clinical Protocol in Supplementary Information). Screening included assessments for known risk factors for severe COVID-19, including comorbidities; low or high body mass index (BMI); abnormal blood tests, including full blood count, renal and liver function, clotting and peripheral blood viral serology; spirometry; echocardiography; and chest radiography (Fig. 1a and the Clinical Protocol in Supplementary Information). The protocol was given a favorable opinion by the UK Health Research Authority Ad Hoc Specialist Ethics Committee (reference 20/UK/2001 (screening

protocol dated 2 December 2020) and reference 20/UK/0002 (main protocol dated 16 February 2021)). Written informed consent was obtained from all volunteers before screening and study enrollment. The study was overseen by a trial steering committee with advice from an independent data and safety monitoring board. The study was discussed with the Medicines and Healthcare products Regulatory Agency (MHRA). Because no medicinal product was being investigated, the study was deemed not a clinical trial according to UK regulations; as such, a EudraCT number was not assigned, and the clinical study was subsequently registered with ClinicalTrials.gov (identifier NCT04865237). All participants were seronegative at screening by Quotient MosaiQ antibody microarray test and had no history of SARS-CoV-2 vaccination or infection However, two inoculated participants were subsequently found to have seroconverted between screening and inoculation, resulting in 34 individuals in the per-protocol analysis.

As this human challenge model was developed during the ongoing pandemic, with no directly comparable safety data and incomplete understanding of long-term effects after COVID-19, an adaptive protocol was designed with stepwise progression to ensure maximal risk mitigation during the early stages and progression only as data on the clinical features of human SARS-CoV-2 challenge were acquired. After extensive screening, participants were admitted to individual negative pressure rooms in an in-patient quarantine unit at the Royal Free London NHS Foundation Trust with 24-hour medical monitoring and access to higher-level clinical support. At admission and before inoculation, volunteers were screened for coincidental respiratory infection using the BioFire FilmArray. Initial cohorts comprised three sentinel individuals followed by seven additional participants. As per protocol, these first ten challenged participants were assigned to receive pre-emptive remdesivir (100 mg intravenously for 5 days) once two consecutive 12-hourly nose or throat swabs showed quantifiable SARS-CoV-2 detection by polymerase chain reaction (PCR), with the aim of mitigating any unexpected risk of progression to more severe disease. After review by the data and safety monitoring board and the trial steering committee, pre-emptive remdesivir was deemed unnecessary, and target recruitment of an additional 30 individuals under the same conditions but without remdesivir was advised. An additional sentinel cohort of three individuals was then challenged, with no pre-emptive remdesivir given. This was followed by three more groups of seven, seven and nine individuals, respectively. Once pre-emptive remdesivir was no longer used, clinical severity criteria (that is, persistent fever, persistent tachycardia, persistent severe cough, greater than mild computed tomography (CT) imaging changes or $SaO_2 \leq 94\%$) were defined for triggering of rescue treatment with monoclonal antibodies (REGEN-COV, Regeneron), but no such treatment was ultimately required. Participants were quarantined for at least 14 days after inoculation and until they met virological discharge criteria (Methods), with planned follow-up for 1 year to assess for prolonged symptoms, including smell disturbance and neurological dysfunction.

All participants were inoculated with 10 $TCID_{50}$ of SARS-CoV-2/human/GBR/484861/2020 (a D614G-containing pre-alpha wild-type virus; GenBank accession number OM294022; $TCID_{50}$ is the median tissue culture infectious dose) by intranasal drops (Fig. 1b). Eighteen participants, 53% according to the per-protocol analysis (95% confidence interval (CI) (35, 70)), subsequently developed PCR-confirmed infection. This infection rate met the protocol-specified primary endpoint target of 50–70%, and there was, therefore, no further dose escalation. Demographics between infected participants and those who remained uninfected were similar (Table 1).

Dynamics of SARS-CoV-2 challenge infection. In the 18 infected individuals, viral shedding detected by quantitative PCR (qPCR) (a secondary endpoint) became quantifiable in throat swabs from
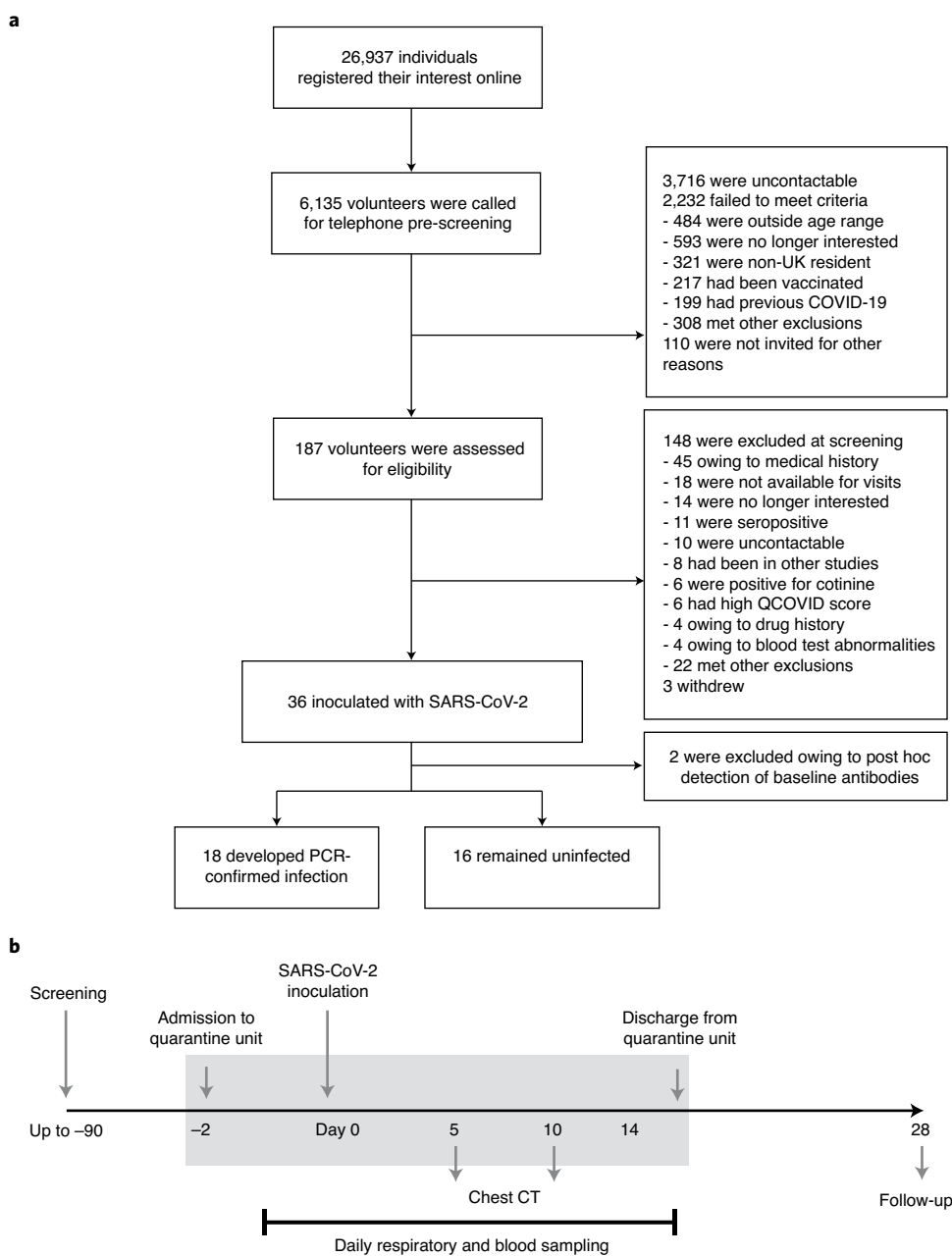
a

b

Fig. 1 | Screening, inoculation, assessments and sampling. Healthy adult volunteers aged 18-29 years were enrolled for SARS-CoV-2 challenge.
a, CONSORT diagram shows inclusions or exclusions and infection outcomes. **b**, Diagram showing the clinical study design up to day 28 after inoculation.

40 hours (median, 95% CI (40, 52)) (~1.67 days) after inoculation, significantly earlier than in the nose ($P = 0.0225$, where initial viral quantifiable detection occurred at 58 hours (95% CI (40, 76))) (~2.4 days) after inoculation (Fig. 2a,b). This was initially closely paralleled by viable virus measured by focus-forming assay (FFA), which was also quantifiably detected earlier in the throat than in the nose ($P = 0.0058$; Fig. 2b). VLs increased rapidly thereafter, with qPCR peaking in the throat at 112 hours (95% CI (76, 160)) (~4.7 days) after inoculation and later at 148 hours (95% CI (112, 184)) (~6.2 days) after inoculation in the nose (Fig. 2a,c). However, at its peak, VL was significantly higher in nasal samples at 8.87 (95% CI (8.41, 9.53)) $\log_{10}$ copies per milliliter and 3.9 (95% CI (3.34, 4.42)) $\log_{10}$ focus-forming units (FFU) per milliliter (FFU ml^{-1}) than in the throat at 7.65 (95% CI (7.39, 8.24)) $\log_{10}$ copies per

milliliter and 2.92 (95% CI (2.68, 3.56)) $\log_{10}$ FFU ml^{-1} ($P < 0.0001$ for qRT–PCR and $P = 0.0024$ for FFA) (Fig. 2d).

In both nose and throat, viral detection continued at high levels for several days, and high cumulative VLs by area under the curve (AUC) were, therefore, seen, particularly in the nose (median 9.03, 95% CI (8.65, 9.43) copies per milliliter by qPCR)(Fig. 2e). In all infected participants, quantifiable virus by qPCR was still present at day 14 after inoculation, which necessitated prolonged quarantine of up to five extra days until qPCR cycle threshold (Ct) values had fallen to less than 33.5 in two consecutive nasal and throat swabs (as per-protocol-defined discharge criteria). At these later time points, VLs by qRT–PCR were more erratic, with low level qPCR positivity remaining in 15 of 18 (83%) participants at discharge. At day 28 after inoculation, six of 18 (33%) participants remained

Table 1 | Participant baseline physical and demographic characteristics, selected clinical features and adverse events

Group (serostatus at inoculation)[a]	Total	Infected (sero-negative)	Uninfected (sero-negative)	Uninfected (sero-positive)
	(n = 36)	(n = 18)	(n = 16)	(n = 2)
Characteristic				
Age (years)				
Mean (s.d.)	21.8 (2.9)	22.2 (2.9)	20.8 (2.2)	26.5 (3.5)
Min, Max	18, 29	18, 27	18, 25	25, 29
Gender, n (%)				
Male	26 (72)	12 (67)	14 (88)	0
Female	10 (28)	6 (33)	2 (12)	2 (100)
Race, n (%)				
White or Caucasian	33 (92)	17 (94)	14 (88)	2 (100)
Mixed ethnicity	3 (8)	1 (6) White/Latino	2 (12) 2× White/Asian	0
BMI (kg m^{-2})				
Mean (range, s.d.)	23.2 (19.6–29.7, 2.6)	22.8 (19.9–26.4, 2.2)	23.4 (19.6–29.7, 3.0)	25.2 (23.3–27.1, 2.7)
Symptoms				
Report of any symptoms on two consecutive days, n (%)	22 (61)	17 (94)	5 (31)	0
Fever				
>37.8 °C, n (%)		7 (39)	0	0
C-reactive protein				
C-reactive protein > 5 mg l^{-1}, n (%)		5 (28)	0	0
Antibody titers at 28 days after inoculation				
Neutralizing antibody titer (median)		863.5 (IQR 403)	Undetectable[b]	167.5
Spike-specific IgG titer (ELU ml^{-1}, median)		1,549 (IQR 1,865)	Undetectable[b]	178
Adverse events				
Any serious adverse event	0	0	0	
Clinically significant adverse events thought to be associated with viral infection that occurred or worsened during the observation period				
Smell disturbance				
During quarantine		15	0	0
Day 28		12[c]	1[d]	1[e]
Day 90		4	0	0
Day 180		5	0	0
Low white cell count (<2.0 × 10^9 per l)		1	0	0
Low lymphocytes (<0.75 × 10^9 per l)		9	0	0
Low neutrophils (<1.0 × 10^9 per l)		3	0	0
Low neutrophils (<0.5 × 10^9 per l)		1	0	0
Epididymal discomfort		1	0	0

[a]Serostatus was based on prospective (MosaiQ) and planned retrospective analysis (neutralizing and S-protein-binding antibodies) of samples before inoculation with challenge virus. [b]One participant was naturally infected with SARS-CoV-2 between discharge from quarantine and day 28 after inoculation, so this participant was excluded. Their neutralizing antibody titer at day 28 was 472, and their spike-specific IgG was 536.6. [c]Includes one participant who had not previously had smell disturbance. Participant reported a runny nose that disturbed smell at day 28, which quickly resolved (no significant change in UPSIT recorded). [d]Participant reported partial smell loss having had 'natural' COVID between discharge from quarantine and day 28 follow-up (no significant change in UPSIT recorded). [e]Participant reported smell disturbance only after performing the UPSIT; the score was not significantly different than baseline.

qPCR positive in the nose and two of 18 (11%) in the throat, but, by day 90, all participants were qPCR negative. Of the participants not meeting infection criteria and deemed uninfected, low-level non-consecutive viral detections were observed only by qPCR in the nose of three participants and in the throat of six participants (Extended Data Fig. 1a,b).

In contrast, viable virus was detectable by FFA for a more limited duration: 156 hours (median, 95% CI (120, 192)) (6.5 days) in the nose and in the throat for 150 hours (95% CI (132, 180)) (6.25 days) (Fig. 2f). The average time after inoculation to clearance of viable

virus was 244 hours (95% CI (208, 256)) or 10.2 days from the nose and 208 hours (95% CI (172, 244)) or 8.7 days from the throat (Fig. 2g). VLs by qPCR and FFA were significantly correlated in both nose and throat (Extended Data Fig. 2). Although there was a striking degree of concordance between the shape and magnitude of individuals' VL curves (Fig. 2a) and between VLs in the nose and throat (Fig. 2i), greater inter-individual variability was observed in timing of VL between nose and throat. Despite relatively high levels of late qPCR detection, the latest that viable virus could be detected was day 12 after inoculation in the nose of one participant and day
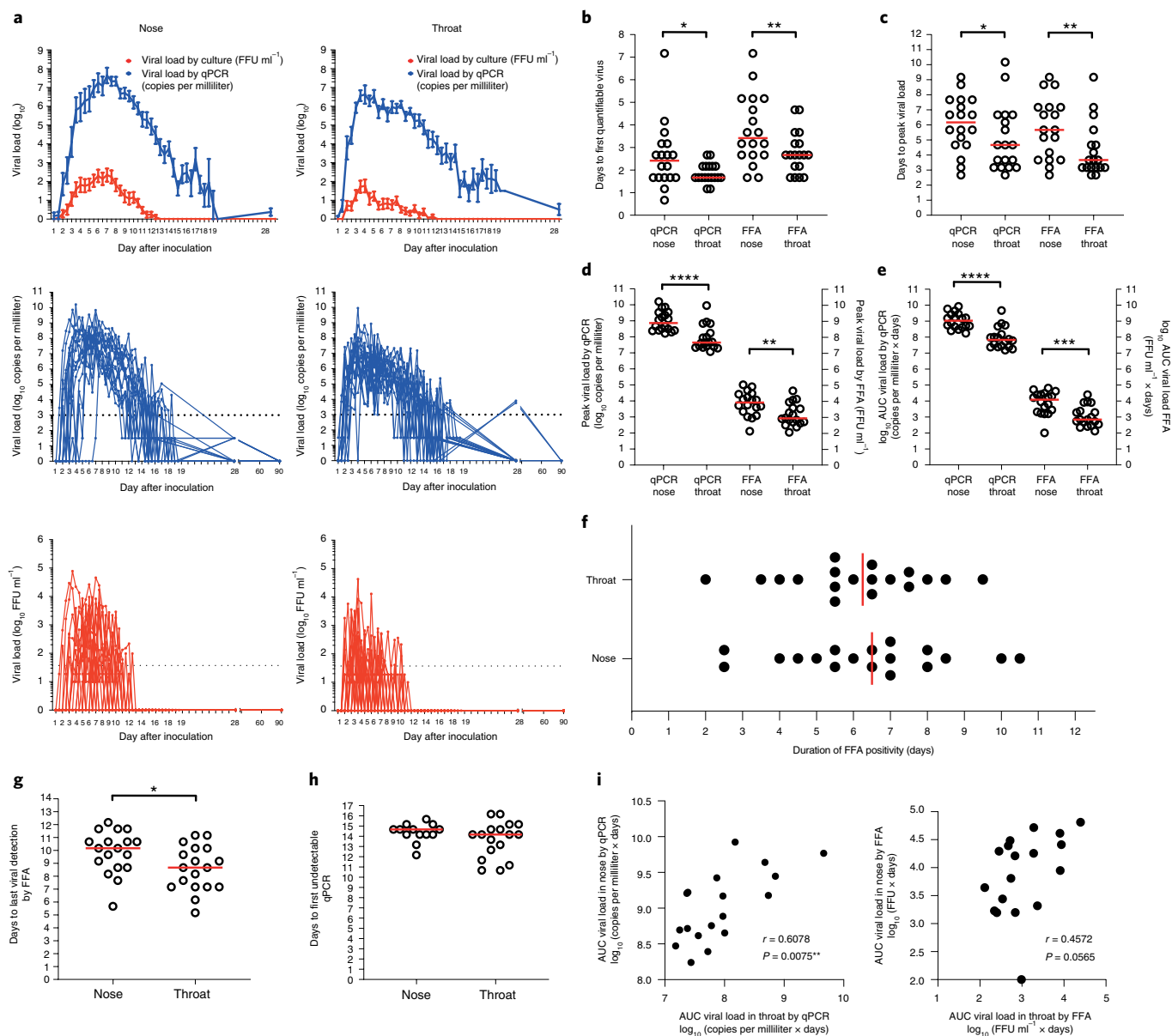
Fig. 2 | Viral shedding after a short incubation period peaks rapidly after human SARS-CoV-2 challenge. Healthy adult volunteers were challenged intranasally with SARS-CoV-2. In the infected individuals ($n = 18$ biologically independent participants), VL in twice-daily nose and throat swab samples was measured by qPCR (blue) and FFA (red) (**a**). Results are expressed as mean $\pm$ s.e.m. Dotted lines represent the LLOQ. Median time to first quantifiable virus (*$P = 0.0195$ and **$P = 0.0058$) (**b**) and peak VLs (*$P = 0.0214$ and **$P = 0.0053$) (**c**) are shown in red. Peak (****$P < 0.0001$ and **$P = 0.0023$) (**d**) and cumulative (AUC, ****$P < 0.0001$ and ***$P = 0.0007$) (**e**) VLs by qPCR and FFA in the nose and throat are compared. **f**, Total duration of viral detection by FFA in nose and throat are shown. Medians are shown in red. **g**, Median time to the last viral detection by FFA after inoculation is shown in red (*$P = 0.0112$). **h**, Median time to the first undetectable VL by qPCR in the individuals who became undetectable while in quarantine is shown in red. Two-sided Wilcoxon matched-pairs signed-rank tests are used to test significance. **i**, AUC VL by qPCR and FFA are correlated in nose versus throat. Spearman's r and P values are shown. *$P < 0.05$, **$P < 0.01$, ***$P < 0.0001$ and ****$P < 0.0001$.

11 in the throat of two participants (Fig. 2g). In contrast, swabs by qPCR that became undetectable in quarantine during the resolution phase first occurred at 352 hours (95% CI (340, 364)) (~14.6 days) in the nose and 340 hours (95% CI (304, 352)) (~14.7 days) in the throat, although some later continued to fluctuate around the limits of quantification and detection (Fig. 2h).

Of the first ten participants prospectively assigned to receive pre-emptive remdesivir on PCR-confirmed infection, six became infected. No apparent differences were seen in VL by qPCR (Extended Data Fig. 3a) or FFA (Extended Data Fig. 3b) between remdesivir-treated and untreated infected participants, and

cumulative virus (AUC) was similar (Extended Data Fig. 3c). Although there was an apparent trend toward lower mean nasal VL during the treatment period and delayed VL peak in the six remdesivir-treated participants (Extended Data Fig. 3d), this was not observed in the throat, was primarily driven by one participant (Extended Data Fig. 4), and was not statistically significant. With no significant differences between remdesivir-treated and untreated participants, infected participants were, therefore, analyzed together.

Thus, after SARS-CoV-2 human challenge, viral shedding begins within 2 days of exposure, rapidly reaching high levels with viable
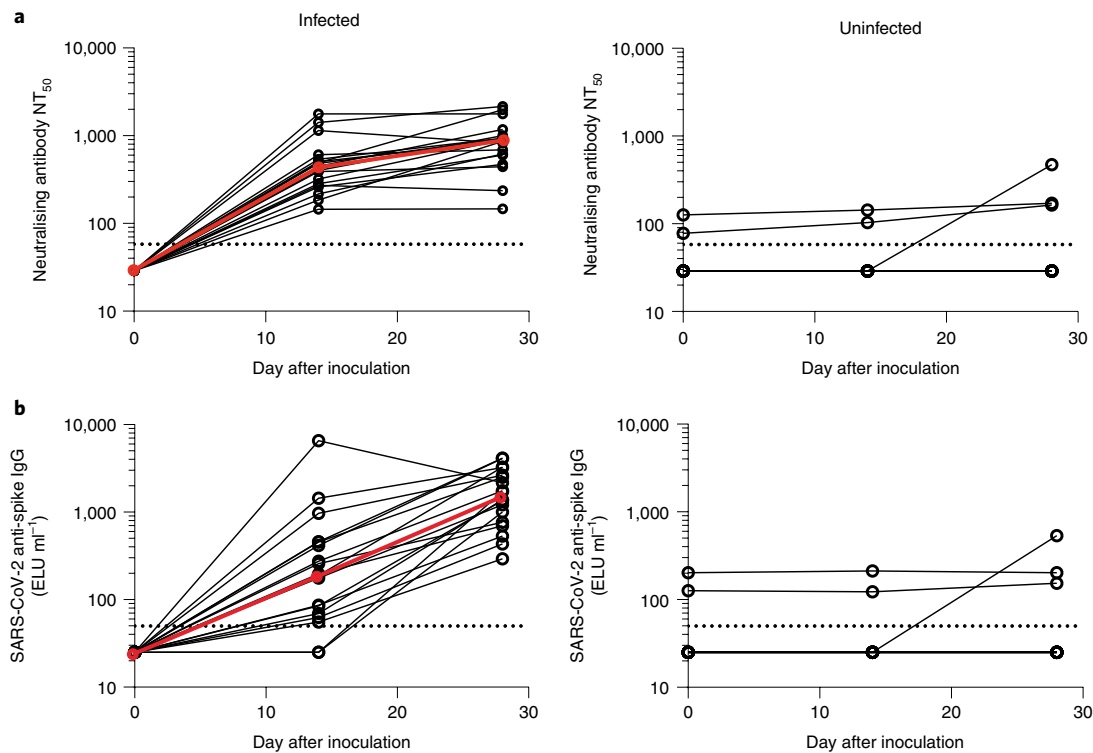
Fig. 3 | Neutralizing antibodies are generated more rapidly than anti-spike protein IgG after human SARS-CoV-2 challenge. Serum was collected before inoculation and at days 14 and 28 after inoculation with SARS-CoV-2. **a**, Serum neutralizing antibodies were measured by microneutralization assay in participants who became infected ($n = 18$ biologically independent participants) and those who remained uninfected ($n = 18$ biologically independent participants). **b**, Serum SARS-CoV-2 anti-spike IgG was measured by ELISA and expressed as ELU ml^{-1}. Individual data points and median (red) are shown. The LLOD of each assay is shown by the dotted line. Undetectable samples were assigned a value of half the LLOD.

virus detectable up to 12 days after inoculation and significantly higher VL in the nose than the throat despite its later onset.

Detection of serum neutralizing antibodies. The rapid onset of infection was reflected in serum antibody responses, analyzed as exploratory endpoints. No increase in serum antibodies by microneutralization or anti-spike protein IgG enzyme-linked immunosorbent assay (ELISA) was observed in those deemed uninfected, even where isolated viral detections had occurred, except for one participant who acquired natural COVID-19 after discharge from quarantine and seroconverted between days 14 and 28 after inoculation (Fig. 3a,b). In contrast, serum antibodies were generated in all infected participants with neutralizing antibody titers of 425 (median, interquartile range (IQR) 269) at 14 days after inoculation and a further rise to 863.5 (IQR 403) at 28 days (Fig. 3a). A slower rise was seen in spike-protein-binding IgG measured by ELISA, with a median increase to 192.5 (IQR 393.1) ELISA laboratory units (ELU) per milliliter (ELU ml^{-1}) at day 14, followed by an increment by day 28 to 1,549 (IQR 1,865) ELU ml^{-1} (Fig. 3b). Of note, in the two participants who seroconverted between screening and inoculation, both neutralizing and S-protein-binding antibodies were detectable at admission to the quarantine unit on day −2 before inoculation. Both participants were excluded from the per-protocol infection rate analysis but remained uninfected, with no change in their serum antibody levels after inoculation.

Symptoms and safety analysis. After infection, as part of the secondary endpoint analyses, symptoms by self-reported diary (Supplementary Table 1) became apparent from 2–4 days after inoculation (Fig. 4a) when symptoms started diverging from challenged but uninfected participants, who reported both fewer and milder

symptoms with no consistent pattern (Fig. 4a and Extended Data Fig. 1c). Symptom scores exhibited greater variability than VLs, with inconsistent onset and peak cumulative daily scores ranging from 0 to 29 (Extended Data Figs. 4 and 5). Symptoms were most frequent in the upper respiratory tract and included nasal stuffiness, rhinitis, sneezing and sore throat (Fig. 4b,c and Extended Data Figs. 4 and 5). Systemic symptoms of headache, muscle/joint aches, malaise and feverishness were also recorded. There was no difference in symptoms between remdesivir-treated and untreated participants (Extended Data Fig. 6). All symptoms were mild to moderate (Extended Data Figs. 7 and 8), with peak symptoms (at 112 hours after inoculation (95% CI (88, 208))) aligning closely with peak VL in the nose, which was significantly later than peak VL in the throat by FFA (88 hours, 95% CI (76, 112), $P = 0.0114$) (Fig. 4d). However, despite the temporal association between nasal VL and symptoms, there was no correlation between the amount of viral shedding by qPCR or FFA and symptom score AUC (Fig. 4e,f).

Seven participants (39% of those infected) had temperatures of >37.8 °C. Otherwise, there were no notable disturbances in any clinical assessments used to partly determine the primary endpoint, including daily spirometry and thoracic CT scans. No serious adverse events were reported, and no criteria for commencing rescue therapy were met. A total of 18 adverse events deemed probably or possibly related to virus infection were largely due to transient and non-clinically significant leukopenia and neutropenia and mild muco-cutaneous abnormalities during the quarantine period (Table 1 and Supplementary Table 2).

Smell disturbance after SARS-CoV-2 challenge infection. To assess the degree and kinetics of smell disturbance, University of Pennsylvania smell identification tests (UPSITs) were conducted.
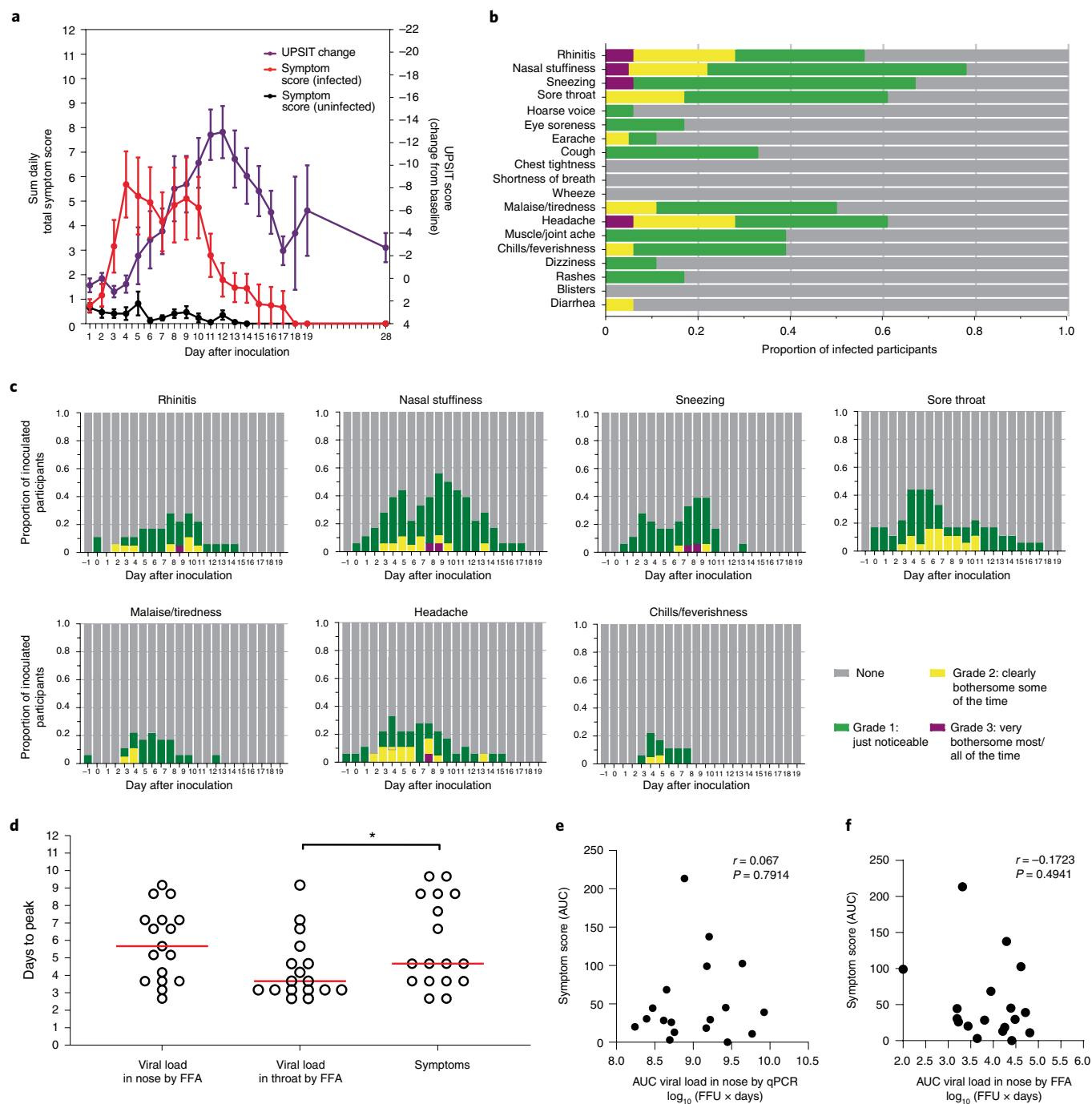
Fig. 4 | Human SARS-CoV-2 challenge induces mild-to-moderate symptoms that correlate with timing, but not magnitude, of VL. Symptom scores were collected using self-reported symptom diaries three times daily from infected participants ($n=18$ biologically independent participants). **a**, Total symptom (red) and fall in UPSIT (purple) scores are shown for all infected participants. Total symptom scores for uninfected participants are shown in black. Mean ± s.e.m. symptom scores are shown. **b**, The frequency and peak severity of symptoms of each type reported by infected participants over the course of the quarantine period are shown. **c**, The frequency and severity of the seven most commonly reported symptoms are shown on each day during the quarantine period scored as follows: none (gray); grade 1, just noticeable (green); grade 2, clearly bothersome some of the time (yellow); and grade 3, very bothersome most or all of the time (purple). **d**, The day after inoculation of peak VL in nose and throat are compared with the day of highest reported symptom score. Medians are shown, and two-sided Wilcoxon matched-pairs signed-rank test is used (*$P < 0.05$). **e,f**, AUC of VL in the nose by qPCR (**e**) and FFA (**f**) is correlated with AUC of total symptom scores over the quarantine period. Spearman's r and P values are shown.

No smell disturbance was observed during quarantine in uninfected participants (Extended Data Fig. 1d). However, 15 infected participants (83%) reported some degree of smell disturbance (14 of which were detected by UPSIT). Although other symptoms peaked with nasal VLs, the nadir of UPSIT scores was 6–7 days later (Fig. 4a

and Extended Data Figs. 4 and 5). Complete smell loss (anosmia) occurred in nine participants (50%), but most improved noticeably before day 28. At day 28, partial smell disturbance was still reported by 11 participants (61%), but, by day 180, this number had fallen to five. Of these, only one participant still had any measurable

smell impairment, although this was steadily improving both subjectively and objectively (UPSIT at baseline = 31, at day 11 = 9, at day 28 = 11, at day 90 = 17 and at day 180 = 23; UPSIT maximum score = 40; significant drop > 4). Two of the remaining participants reported mild parosmia, and two had mild subjective reduction in smell (although UPSIT scores had normalized). Six participants (including all five who had prolonged smell disturbance) received smell training advice, including two who also received treatment with short courses of oral and intranasal steroids.

Anosmia is, therefore, a common feature of human SARS-CoV-2 challenge that generally emerges several days later than viral shedding and resolves within 28 days in most individuals. Together, these findings suggest that human challenge infection with wild-type SARS-CoV-2 at this inoculum dose has low risk of causing severe symptoms in healthy young adults but leads to large amounts of nasopharyngeal virus even in the absence of respiratory or systemic disease.

Accuracy of rapid antigen testing by lateral flow assay. Lateral flow assay (LFA) rapid antigen tests are commonly used to identify potentially infectious people in the community, but their usefulness in early infection is unknown. To test the performance of LFA over the entire course of infection, antigen testing was performed using the same morning nose and throat swab samples assessed for VL. None of the uninfected participants had a positive LFA test at any time, whereas all infected participants had positive LFA for ≥2 days (Supplementary Fig. 1). Despite earlier viral detection in the throat by other methods, median time to first detection by daily LFA tests was the same in nose and throat at 4 days (range, 2–8 days) after inoculation (Fig. 5a). This was, on average, 24–48 hours after first qPCR positivity (Fig. 5b) and within 24 hours of FFA (Fig. 5c). Of note, in nine of 18 infected participants, viable virus became detectable by FFA one or more days before the first positive LFA. Toward the end of infection, the last LFA detection mainly occurred 24–72 hours after viable virus detection had ceased.

To assess the relationship between VL and probability of a positive LFA, logistic regression models were fitted using generalized estimating equations to control for repeated within-participant assessments. log_{10} VL was a significant predictor ($P < 2 \times 10^{-5}$) of LFA positivity with an odds ratio of 5.01 (95% CI (2.93, 8.57)) when predicting LFA from FFA in nose (Fig. 5e). Area under the receiver operating characteristic curves (AUROCs) were high at 0.96 for nasal qPCR and 0.89 for throat qPCR (Extended Data Fig. 9a) but lower for FFA, particularly in the throat (AUC 0.69). To test longitudinal performance as infection progressed, the sensitivity and specificity of LFA when compared with qPCR and FFA were calculated for each day after exposure (Fig. 5f). With both tests and anatomical sites, sensitivity of LFA was limited at the beginning and end of acute illness. However, from ~4 days after inoculation, LFA demonstrated high sensitivity as a surrogate for qPCR or FFA positivity. Overall, LFA was highly specific, although some 'false positives' were observed in relation to FFA (but not qPCR).

Where asymptomatic or pre-symptomatic LFA testing programs exist, testing is usually recommended twice weekly. To model the differential effect of LFA testing frequency that incorporate viral dynamics throughout infection, the mean proportion of VL AUC that had yet to occur (and might be responsible for transmission if undiagnosed) by the time of a first positive LFA test with testing cadences of 1–7 days was modeled. For both FFA (Fig. 5g) and qPCR (Extended Data Fig. 9b), infection would be recognized at or before more than 90% of the VL AUC had occurred if testing were daily. As the period between tests increased, the proportion of VL AUC declined, with twice-weekly testing capturing 70–80% of virus and weekly testing still exceeding 50% if nose and throat swabs were combined. Thus, LFA positivity is strongly associated with

culturable virus and, therefore, contagiousness and may be effective as a trigger for interventions to interrupt transmission.

Discussion

Here we report the virological and clinical results from the first SARS-CoV-2 human challenge study. With a low inoculum dose of 10 TCID$_{50}$, robust viral replication was observed in 53% of seronegative participants. After an incubation period of less than 2 days, VLs rose rapidly, peaking at high levels with infectious virus production for over 1 week. Symptoms were present in 89% of infected individuals but, despite high VLs, were consistently mild to moderate, transient and predominantly confined to the upper respiratory tract. Anosmia/dysosmia was common, occurred later than other symptoms and resolved without treatment in most participants within 90 days. In those with residual smell disturbance, their sense of smell steadily improved during the follow-up period, consistent with the good long-term prognosis seen in community cases[21]. There was no evidence of pulmonary disease in infected participants based on clinical and radiological assessments.

The natural infectious dose of SARS-CoV-2 is unknown, but, based on in vitro and preclinical models, the virus is understood to be highly infectious[22–24] and well-adapted to rapid and high-titer replication in human respiratory mucosa[25]. Early in the pandemic, a World Health Organization Advisory Group published expert consensus guidelines recommending a starting dose of 10^2 TCID$_{50}$ (ref. [17]). Here, based on in vitro data of high viral replication in primary human airway epithelial cells and Syrian hamster data[26], we started with a ten-fold-lower dose of 10 TCID$_{50}$ (equivalent to 55 FFU) and found it sufficient to meet the 50–70% target infection rate. With prospective household contact studies having similarly shown high secondary attack rates of ~38%[19], this suggests that the model can recapitulate higher exposure than naturally acquired infection events. In contrast, experimental infections of non-human primates have used 1,000–10,000 times more virus, with intra-tracheal or combined upper/lower airway administration, which results in markedly different kinetics to those observed during human infection[27,28]. In human challenge studies with other respiratory viruses, such as influenza viruses and respiratory syncytial virus, inoculum doses are typically also much higher at 10^4–10^6 TCID$_{50}$ because all volunteers have been exposed multiple times throughout life to those viruses, with pre-existing immunity reducing susceptibility and resulting in substantially lower peak viral loads at 10^3–10^4 copies per milliliter by PCR[8,9]. Thus, animal models and human data from other viral infections were of limited helpfulness in estimating the optimal SARS-CoV-2 inoculum dose.

Although some studies have measured the response to SARS-CoV-2 infection longitudinally in humans[29–31], none can capture host features at the time of virus exposure, the early events before symptom onset or the detailed course of infection that can be shown by experimental challenge. Although the incubation period from the estimated time of natural exposure to perceived symptom onset has previously been estimated as ~5 days[32,33], this best aligns with peak symptoms and is longer than the true incubation period. With close questioning, symptoms were found to be associated with viral shedding within 2–4 days of inoculation but did not peak until days 4–5. Thus, virus was first detected (first in the throat, then the nose) ~2 days before peak symptoms and increased steeply to achieve a sustained peak, in many cases before peak symptoms were reached, consistent with modeling data indicating that up to 44% of transmissions occur before symptoms are noted[6]. Anosmia was a later symptom, potentially explained by the proposed mechanism whereby only ACE2-expressing and TMPRSS2-expressing supporting cells, rather than neurons themselves, are directly infected, leading to delayed secondary olfactory dysfunction[34].

Pre-emptive remdesivir was administered to the first six infected participants as risk mitigation during early model development, as
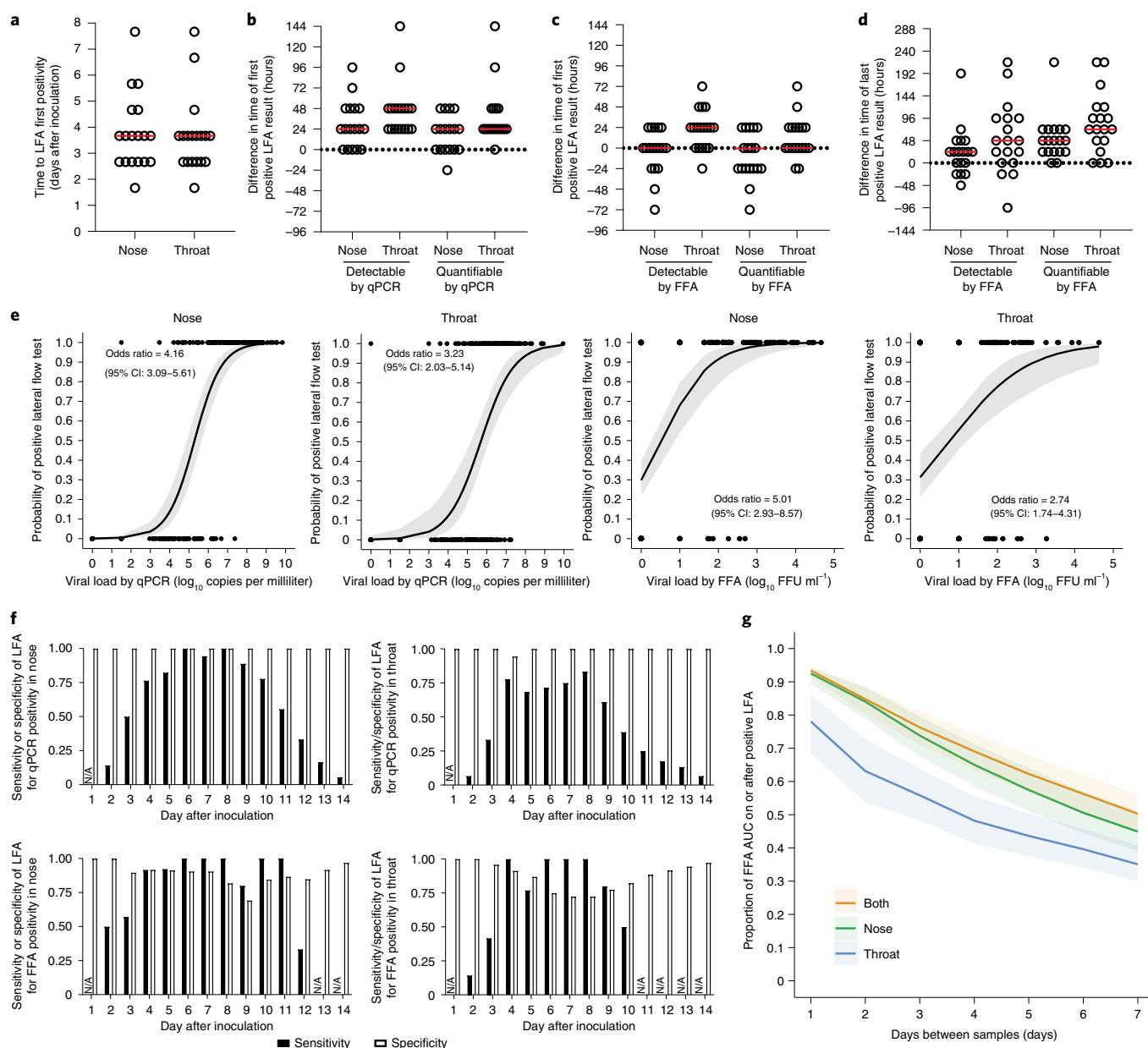
Fig. 5 | Rapid antigen testing by lateral flow accurately predicts infectious virus shedding. Nose and throat swab samples in viral transport medium from infected participants ($n = 18$ biologically independent participants) were tested by LFA. **a**, Time to first LFA positivity is shown. Median (red) difference in timing between LFA and nose or throat are shown for first qPCR (**b**) and FFA (**c**) detection and quantification. **d**, Median (red) and individual number of days between the last detectable or quantifiable FFA result compared with LFA are shown. **e**, Generalized estimating equations logistic regression showing the mean, 95% CIs and odds ratios for lateral flow test positivity at VLs by qPCR and FFA in the nose and throat. **f**, Sensitivity (black) and specificity (white) of LFA in determining qPCR and FFA viral detection in the nose and throat over the course of challenge infection. N/A indicates where there were no true-positive results. **g**, Effect of frequency of LFA testing on the proportion of viable virus shedding after LFA diagnosis from the nose (green), throat (blue) or combined nose and throat (orange). Mean and 95% CIs are shown.

trial data had suggested efficacy in shortening time to recovery in hospitalized patients[35]. However, no statistically significant effect on VL or symptoms was detectable in this small cohort. Field data have shown mixed results for the effectiveness of remdesivir in the hospitalized patient setting[36] but reduction in progression to severe disease in those with risk factors when given early in the course of infection[37]. This study was not designed nor powered to assess the efficacy of early treatment with remdesivir and especially its effect on severe disease, but such prospective human challenge studies are well-placed to answer the question of antiviral efficacy, with treatment commenced at different times relative to virus exposure.

A key unresolved question for public health has been whether transmission is less likely to occur during asymptomatic/mild infection compared to more severe disease. Some studies have shown a correlation between disease severity and extent of viral shedding[38,39], but others have not[40]. Overall, peak VLs reported in natural infection ($\sim10^5$–10^8 copies per milliliter) are lower than those observed in this study[6,41–44]. However, these are invariably sampled at the time of case ascertainment, and, where longitudinal samples have been taken, these indicate that patients are already in the downward phase of the VL curve[31]. It is, therefore, likely that most samples miss the peak of viral shedding. With virus present at significantly higher titers

in the nose than the throat, these data provide clear evidence that emphasizes the critical importance of wearing face coverings over the nose as well as the mouth. Furthermore, our data clearly show that SARS-CoV-2 viral shedding occurs at high levels irrespective of symptom severity, thus explaining the high transmissibility of this infection and emphasizing that symptom severity cannot be considered a surrogate for transmission risk in this disease. This remains relevant with the widespread transmission of the Delta and Omicron variants, where antigenic divergence along with waning vaccine-induced immunity lead to VL during breakthrough infection at similarly high levels to those in seronegative individuals[19,45].

An important limitation of this study, in keeping with first-in-human studies generally, was the small sample size that limits our ability to detect rare events or more subtle differences associated with SARS-CoV-2 infection. Further studies may be needed with larger numbers of participants for sufficient power to identify biomarkers that have less marked differential expression and/or are more inconsistent than VL and symptoms. However, although globally there are groups that remain naive to SARS-CoV-2 vaccination and infection, it is unlikely that a larger study in seronegative volunteers will be achievable going forward, and this study may, therefore, remain the only one of its kind. This likely will necessitate the further development of the model using variants of concern that have been shown to cause breakthrough infection despite vaccine-induced immunity (such as Delta and Omicron) to permit efficacy testing of novel interventions in that context and further understanding of correlates of protection.

Despite the relatively small sample size, limited variation in the kinetics and magnitude of VLs between infected study participants and longitudinal analysis permits several conclusions of public health importance. Detailed viral kinetics show that some individuals still shed culturable virus at 12 days after inoculation (that is, up to 10 days after symptom onset), and, on average, viable virus was still detectable 10 days after inoculation (up to 8 days after symptom onset). These data, therefore, support the isolation periods of 10 days after symptom onset advocated in many guidelines to minimize onward transmission[46]. High levels of asymptomatic/pauci-symptomatic VL also highlight the potential positive effect of routine asymptomatic testing programs that attempt to diagnose infection in the community so that infection control measures, such as self-isolation, can be implemented to interrupt transmission. In several jurisdictions, these rely on rapid antigen tests, with recent re-analysis of cross-sectional LFA validation data having suggested that sensitivity for infectious virus may be higher than previously estimated at ~80%[47]. In this study, longitudinal LFA data after SARS-CoV-2 challenge also strongly predicted culturable virus aside from the very earliest time points where sensitivity was lower. In addition, LFA was highly reliable in predicting the disappearance of viable virus and, therefore, also can underpin 'test to release' strategies, which are increasingly being used to shorten the period of self-isolation. Although positive LFA results were occasionally seen with negative FFA results (causing a reduction in specificity in relation to the viable virus assay), there were no false positives when comparing LFA to qPCR, which may imply the relatively lower sensitivity of viral culture rather than false positivity of LFA. Although some uncertainty remains in directly extrapolating these data to the community where self-swabbing and more concentrated samples may alter sensitivity, these results support their continued use for identifying those most likely to be infectious. Our modeling also suggests that this strategy remains effective even if imperfectly implemented, with routine testing as few as every 7 days able to interrupt more than half the virus still to be shed by an individual, if acted upon.

Although these first-in-human data do not preclude rare adverse events that can be detected only in larger-scale studies, our results indicate that symptoms of human challenge with SARS-CoV-2 in healthy young adults are consistent with those in natural infection.

In this cohort, we observed no severe consequences, which may support further development and expansion of this approach. This first report focuses on safety, tolerability and virological responses, but the uniquely controlled nature of the model will also enable robust identification of host factors present at the time of inoculation and associated with protection in those participants who resisted infection. Analysis of local and systemic immune markers (including potentially cross-reactive antibodies, T cells and soluble mediators) from this SARS-CoV-2 human challenge study that may explain these differences in susceptibility is, therefore, ongoing. In addition, with the feasibility of this approach having been demonstrated using a prototypic wild-type strain, further challenge studies are now underway in which previously infected and vaccinated volunteers will be challenged with escalating inoculum doses (ClinicalTrials.gov identifier NCT04864548) and/or viral variants to investigate the interplay between virus and host factors that influence clinical outcome. Together, these studies will, thus, optimize the platform for potential use in the rapid evaluation of vaccines, antivirals and diagnostics by generating efficacy data early during clinical development and avoiding the uncertainties of studies that require ongoing community transmission.

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/s41591-022-01780-9.

References

1. Drake, T. M. et al. Characterisation of in-hospital complications associated with COVID-19 using the ISARIC WHO Clinical Characterisation Protocol UK: a prospective, multicentre cohort study. *Lancet* **398**, 223–237 (2021).
2. Voysey, M. et al. Safety and efficacy of the ChAdOx1 nCoV-19 vaccine (AZD1222) against SARS-CoV-2: an interim analysis of four randomised controlled trials in Brazil, South Africa, and the UK. *Lancet* **397**, 99–111 (2021).
3. Thomas, S. J. et al. Safety and efficacy of the BNT162b2 mRNA Covid-19 vaccine through 6 Months. *N. Engl. J. Med.* **385**, 1761–1773 (2021).
4. Fischerii, W. A. et al. A phase 2a clinical trial of molnupiravir in patients with COVID-19 shows accelerated SARS-CoV-2 RNA clearance and elimination of infectious virus. *Sci. Transl. Med.* **14**, eabl7430 (2022).
5. He, J., Guo, Y., Mao, R. & Zhang, J. Proportion of asymptomatic coronavirus disease 2019: a systematic review and meta-analysis. *J. Med. Virol.* **93**, 820–830 (2021).
6. He, X. et al. Temporal dynamics in viral shedding and transmissibility of COVID-19. *Nat. Med.* **26**, 672–675 (2020).
7. Syangtan, G. et al. Asymptomatic SARS-CoV-2 carriers: a systematic review and meta-analysis. *Front. Public Health* **8**, 587374 (2020).
8. Jozwik, A. et al. RSV-specific airway resident memory CD8+ T cells and differential disease severity after experimental human infection. *Nat. Commun.* **6**, 10224 (2015).
9. Paterson, S. et al. Innate-like gene expression of lung-resident memory CD8. *Am. J. Respir. Crit. Care Med.* **204**, 826–841 (2021).
10. Reed, S. E. The behaviour of recent isolates of human respiratory coronavirus in vitro and in volunteers: evidence of heterogeneity among 229E-related strains. *J. Med. Virol.* **13**, 179–192 (1984).
11. Habibi, M. S. et al. Neutrophilic inflammation in the respiratory mucosa predisposes to RSV infection. *Science* **370**, eaba9301 (2020).
12. DeVincenzo, J. P. et al. Viral load drives disease in humans experimentally infected with respiratory syncytial virus. *Am. J. Respir. Crit. Care Med.* **182**, 1305–1314 (2010).
13. Bagga, B. et al. Comparing influenza and RSV viral and disease dynamics in experimentally infected adults predicts clinical effectiveness of RSV antivirals. *Antivir. Ther.* **18**, 785–791 (2013).
14. Nguyen-Van-Tam, J. S. et al. Minimal transmission in an influenza A (H3N2) human challenge-transmission model within a controlled exposure environment. *PLoS Pathog.* **16**, e1008704 (2020).

15. Rosenau, M. Experiments to determine mode of spread of influenza. *JAMA* **73**, 311–313 (1919).

16. Rapeport, G. et al. SARS-CoV-2 human challenge studies—establishing the model during an evolving pandemic. *N. Engl. J. Med.* **385**, 961–964 (2021).

17. Levine, M. M. et al. Viewpoint of a WHO advisory group tasked to consider establishing a closely-monitored challenge model of coronavirus disease 2019 (COVID-19) in healthy volunteers. *Clin. Infect. Dis.* **72**, 2035–2041 (2021).

18. Deming, M. E., Michael, N. L., Robb, M., Cohen, M. S. & Neuzil, K. M. Accelerating development of SARS-CoV-2 vaccines—the role for controlled human infection models. *N. Engl. J. Med.* **383**, e63 (2020).

19. Singanayagam, A. et al. Community transmission and viral load kinetics of the SARS-CoV-2 delta (B.1.617.2) variant in vaccinated and unvaccinated individuals in the UK: a prospective, longitudinal, cohort study. *Lancet Infect. Dis.* **22**, 183–195 (2021).

20. Baay, M. & Neels, P. SARS-CoV-2 controlled human infection models: ethics, challenge agent production and regulatory issues. *Biologicals* **67**, 69–74 (2020).

21. Renaud, M. et al. Clinical outcomes for patients with anosmia 1 year after COVID-19 diagnosis. *JAMA Netw. Open* **4**, e2115352 (2021).

22. Imai, M. et al. Syrian hamsters as a small animal model for SARS-CoV-2 infection and countermeasure development. *Proc. Natl Acad. Sci. USA* **117**, 16587–16595 (2020).

23. Sia, S. F. et al. Pathogenesis and transmission of SARS-CoV-2 in golden hamsters. *Nature* **583**, 834–838 (2020).

24. Abdelnabi, R. et al. Comparing infectivity and virulence of emerging SARS-CoV-2 variants in Syrian hamsters. *EBioMedicine* **68**, 103403 (2021).

25. Hou, Y. J. et al. SARS-CoV-2 reverse genetics reveals a variable infection gradient in the respiratory tract. *Cell* **182**, 429–446 (2020).

26. Rosenke, K. et al. Defining the Syrian hamster as a highly susceptible preclinical model for SARS-CoV-2 infection. *Emerg. Microbes Infect.* **9**, 2673–2684 (2020).

27. Corbett, K. S. et al. Evaluation of the mRNA-1273 vaccine against SARS-CoV-2 in nonhuman primates. *N. Engl. J. Med.* **383**, 1544–1555 (2020).

28. Williamson, B. N. et al. Clinical benefit of remdesivir in rhesus macaques infected with SARS-CoV-2. *Nature* **585**, 273–276 (2020).

29. Ke, R., Zitzmann, C., Ho, D. D., Ribeiro, R. M. & Perelson, A. S. In vivo kinetics of SARS-CoV-2 infection and its relationship with a person's infectiousness. *Proc. Natl Acad. Sci. USA* **118**, e2111477118 (2021).

30. Kim, M. C. et al. Duration of culturable SARS-CoV-2 in hospitalized patients with Covid-19. *N. Engl. J. Med.* **384**, 671–673 (2021).

31. Zou, L. et al. SARS-CoV-2 viral load in upper respiratory specimens of infected patients. *N. Engl. J. Med.* **382**, 1177–1179 (2020).

32. Yu, P., Zhu, J., Zhang, Z. & Han, Y. A familial cluster of infection associated with the 2019 novel coronavirus indicating possible person-to-person transmission during the incubation period. *J. Infect. Dis.* **221**, 1757–1761 (2020).

33. Lauer, S. A. et al. The incubation period of coronavirus disease 2019 (COVID-19) from publicly reported confirmed cases: estimation and application. *Ann. Intern. Med.* **172**, 577–582 (2020).

34. Butowt, R. & von Bartheld, C. S. Anosmia in COVID-19: underlying mechanisms and assessment of an olfactory route to brain infection. *Neuroscientist* **27**, 582–603 (2020).

35. Beigel, J. H. et al. Remdesivir for the treatment of Covid-19—final report. *N. Engl. J. Med.* **383**, 1813–1826 (2020).

36. Barratt-Due, A. et al. Evaluation of the effects of remdesivir and hydroxychloroquine on viral clearance in COVID-19: a randomized trial. *Ann. Intern. Med.* **174**, 1261–1269 (2021).

37. Gottlieb, R. L. et al. Early remdesivir to prevent progression to severe Covid-19 in outpatients. *N. Engl. J. Med.* **386**, 305–315 (2022).

38. Eberhardt, K. A. et al. RNAemia corresponds to disease severity and antibody response in hospitalized COVID-19 patients. *Viruses* **12**, 1045 (2020).

39. Fajnzylber, J. et al. SARS-CoV-2 viral load is associated with increased disease severity and mortality. *Nat. Commun.* **11**, 5493 (2020).

40. Marks, M. et al. Transmission of COVID-19 in 282 clusters in Catalonia, Spain: a cohort study. *Lancet Infect. Dis.* **21**, 629–636 (2021).

41. Munker, D. et al. Dynamics of SARS-CoV-2 shedding in the respiratory tract depends on the severity of disease in COVID-19 patients. *Eur. Respir. J.* **58**, 2002724 (2021).

42. Walsh, K. A. et al. SARS-CoV-2 detection, viral load and infectivity over the course of an infection. *J. Infect.* **81**, 357–371 (2020).

43. Baggio, S. et al. Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) viral load in the upper respiratory tract of children and adults with early acute coronavirus disease 2019 (COVID-19). *Clin. Infect. Dis.* **73**, 148–150 (2021).

44. Wölfel, R. et al. Virological assessment of hospitalized patients with COVID-2019. *Nature* **581**, 465–469 (2020).

45. Levine-Tiefenbrun, M. et al. Viral loads of delta-variant SARS-CoV-2 breakthrough infections after vaccination and booster with BNT162b2. *Nat. Med.* **27**, 2108–2110 (2021).

46. Quilty, B. J. et al. Quarantine and testing strategies in contact tracing for SARS-CoV-2: a modelling study. *Lancet Public Health* **6**, e175–e183 (2021).

47. Petersen, I., Crozier, A., Buchan, I., Mina, M. J. & Bartlett, J. W. Recalibrating SARS-CoV-2 antigen rapid lateral flow test relative sensitivity from validation studies to absolute sensitivity for indicating individuals shedding transmissible virus. *Clin. Epidemiol.* **13**, 935–940 (2021).

Methods

Ethics statement. This study was conducted in accordance with the protocol; the consensus ethical principles derived from international guidelines, including the Declaration of Helsinki and Council for International Organizations of Medical Sciences International Ethical Guidelines; applicable ICH Good Clinical Practice guidelines; and applicable laws and regulations. The screening protocol and main study were approved by the UK Health Research Authority's Ad Hoc Specialist Ethics Committee (references 20/UK/2001 and 20/UK/0002). Written informed consent was obtained from all volunteers before screening and study enrollment. Participants were given a donation of up to £4,565 to compensate for the time and inconvenience of taking part in the study (including at least a 17-day quarantine). This was calculated using the National Institute for Health Research (NIHR) formula and the UK national living wage. The study was overseen by a medical oversight committee (trial steering committee) with advice from an independent data and safety monitoring board, which assessed the study data. Discussion was had with the MHRA to determine whether the study was to be deemed a clinical trial of an investigational medicinal product. Because no medicinal product was being investigated, the study was deemed not a clinical trial; as such, a EudraCT number was not assigned, and the clinical study was registered with ClinicalTrials. gov (identifier NCT04865237). The study was delayed by these discussions and the ClinicalTrials.gov registration and, therefore, went live after the first participants were enrolled.

Public consultation and involvement. Building on earlier work in the UK[48], broad consultation was undertaken to explore public understanding of the concept of a human challenge study in general and the acceptability of a human challenge study with SARS-CoV-2 taking place in the UK[49]. This involved a cross-sectional survey (2,441 participants) and a series of focus groups (57 participants in nine groups). A group of public advisors provided input into study design and materials. The opinions and concerns gathered were used to inform study design and document preparation.

Study design. Healthy adults aged 18–30 years with no evidence of previous SARS-CoV-2 infection or vaccination were recruited to this single-center, phase 1, open-label, first-in-human study. The first date of participant enrollment was 6 March 2021; the last date was 8 July 2021. Volunteers were excluded if positive for anti-SARS-CoV-2 S protein antibodies using the MosaiQ COVID-19 antibody microarray (Quotient) and on the basis of risk factors assessed by clinical history, physical examination and screening assessments. The QCOVID tool[50] was used to provide a personalized estimated absolute risk of hospitalization and death, with those above a pre-defined risk threshold (equivalent to that for a 30-year-old individual with no risk factors, calculated as a 1:250,000 risk of death or 1:4,902 risk of hospitalization) excluded. Echocardiography and chest X-ray were performed before inoculation. See protocol for full inclusion and exclusion criteria. Additionally, per-protocol analysis was performed after exclusion of participants who fulfilled enrollment criteria at screening but were later found to have neutralizing and spike-binding antibodies on admission to the quarantine unit.

The primary objective was to identify a dose of wild-type SARS-CoV-2 in healthy volunteers with an acceptable safety profile that induced laboratory-confirmed infection in ≥50% of participants, suitable for future human challenge studies. Laboratory-confirmed infection was defined as quantifiable RT–PCR detection greater than the lower limit of quantification (LLOQ) from mid-turbinate and/or throat swabs on two or more consecutive 12-hourly time points, starting from 24 hours after inoculation and up to discharge from quarantine. Secondary objectives and exploratory endpoints are listed in the protocol.

The study was conducted in a high-containment clinical trials unit at the Royal Free London NHS Foundation Trust. Participants were housed in single-occupancy, negative pressure side rooms. Participants were inoculated intranasally by pipette with 10 $TCID_{50}$ of wild-type SARS-CoV-2 (100 μl per naris) between both nostrils, with an initial sentinel group ($n = 3$) followed by the remaining individuals in the cohort. Participants remained supine (face and torso facing up) for 10 minutes, followed by 20 minutes in a sitting position wearing a nose clip after inoculation to ensure maximum contact time with the nasal and pharyngeal mucosa. Mid-turbinate nose and throat samples were collected twice daily using flocked swabs, placed in 3 ml of viral transport medium (BSV-VTM-001, Bio-Serv) that was aliquoted and stored at −80 °C. For the first ten challenged participants, pre-emptive intravenous remdesivir (100 mg intravenously for 5 days; Gilead) was initiated after two consecutive quantifiable viral detections and administered twice daily for 5 days to infected participants only. Participants remained in quarantine for a minimum of 14 days after inoculation until the following discharge criteria were met: two consecutive daily nose and or throat swabs with no viral detection or a qPCR Ct value >33.5 and no viable virus by overnight incubation viral culture with detection by immunofluorescence. Participants will be followed for 1 year after inoculation and for data collected after the day 28 data lock. The analyses are exploratory.

Clinical assessments. Safety was assessed with daily blood tests, spirometry, electrocardiograms, clinical assessments (vital signs, symptom diaries and clinical examination) and CT scan of chest on day 5 (in all participants) and day 10

(in infected participants only). Self-completed symptom diaries were completed three times daily from the day before inoculation to day 14 after inoculation. A total of 19 symptoms covering upper respiratory, lower respiratory and systemic symptoms were scored on a scale of 0–3 (that is, absence of symptoms, mild, moderate and severe, respectively). Blood and respiratory samples were obtained before infection and at time points indicated in the text. Smell was monitored using the UPSIT, a well-validated and reliable (test–retest $r = 0.94$) test that employs microencapsulated 'scratch and sniff' odorants provided as booklets containing a series of cards that the participants scratch[51]. The total number of odorant stimuli out of 40 that is correctly identified serves as the test measure. A fall in score of more than 4 was considered significant.

Challenge virus. The SARS-CoV-2 challenge virus (full formal name: SARS-CoV-2/human/GBR/484861/2020) was obtained with consent from a nose/throat swab taken from a patient in the UK with COVID-19, facilitated by the International Severe Acute Respiratory and Emerging Infection Consortium (ISARIC) Coronavirus Clinical Characterisation Consortium (ISARIC4C)[52] using their study protocol (study registry ISRCTN66726260) approved by the South Central–Oxford C Research Ethics Committee in England (reference 13/SC/0149) and the Scotland A Research Ethics Committee (reference 20/SS/0028).

The virus was isolated by inoculation of a qualified cGMP Vero cell line with the clinical sample. Sequence analysis showed this to be from the 20A clade of the B.1 lineage and possessed the D614G mutation. A seed virus stock was then generated by a further passage on the same cGMP Vero cell line. The seed virus stock was then used to manufacture the challenge virus in accordance with cGMP at the Zayed Centre for Research GMP manufacturing facility of Great Ormond Street Hospital, and a challenge virus master virus bank (MVB) was produced. Individual dose inoculum vials were then produced in accordance with cGMP by dilution of the cGMP MVB with sucrose diluent. The challenge virus underwent quality testing performed as part of the GMP manufacturing release processes according to pre-determined specifications (including identity, infectivity and contaminant/adventitious agent tests). This included whole-genome sequencing for confirmation that the GMP virus was unaltered compared to the original clinical isolate. The sequence of the challenge virus has been deposited in GenBank (accession number OM294022). In the UK, because they are not medicinal products, challenge viruses are not regulated by the MHRA. However, the challenge virus was manufactured according to GMP, and the supporting paperwork was reviewed by the MHRA, which confirmed that the manufacture was suitable for the challenge agent to be used in future efficacy studies of investigational medicinal products. Therefore, in future clinical trials of an investigational medicinal product, the challenge virus will be reviewed as part of the clinical trial application to the MHRA. The challenge virus was stored in a secure −80 °C freezer (normal temperature range, −60 °C to −90 °C) until use. The SARS-CoV-2 inoculum dose (10^1 $TCID_{50}$) was selected as the lowest infectious dose that could be reliably quantified by viral culture.

Virology. Longitudinal measures of pharyngeal and nasal viral kinetics were measured using two independent assays: (1) qRT–PCR with N gene primers/probes adapted from the Centers for Disease Control and Prevention (CDC) protocol[53] (updated 29 May 2020) and (2) quantitative culture by FFA.

- RT–PCR

Aliquots of the clinical samples were processed for PCR analysis by the addition of Qiagen's ATL lysis buffer to the sample to lyse the virus before subsequent RNA extraction, PCR and reaction setup using Qiagen's QIAsymphony RNA extraction (SP) and assay setup (AS) modules. Each sample was spiked with an internal control RNA before PCR amplification in triplicate in a multiplex PCR assay run on Applied Biosystems' ViiA7 PCR machines.

The SARS-CoV-2 primer and probe sequences were derived from the CDC protocol (see below).

- Forward primer 5′–3′: GACCCCAAAATCAGCGAAAT
- Reverse primer 5′–3′: TCTGGTTACTGCCAGTTGAATCTG
- Probe 5′–3′: ACCCCGCAT/ZEN/TACGTTTGGTGGACC (labeled with FAM, ZAN and 31ABkFQ)

Ct values were converted to copies per milliliter by comparison with a standard curve run concurrently within the assay. The standard curve was generated from in vitro transcribed RNA from linearized plasmid of known concentrations containing sequences of the SARS-CoV-2 virus nucleocapsid and spike.

- Quantitative culture

Quantitative virus infectivity was assessed using FFAs. Samples were assayed in triplicate inoculating Vero cells seeded the day prior at 3×10^4 cells per well in a 96-well plate format. Serial dilutions of the test samples were inoculated onto cells and incubated at 37 °C and 5% CO_2 for 1 hour before the addition of methylcellulose overlay and further incubation for 1 day. Cells were fixed and stained using standard neutral buffered formalin (10%) and Triton X-100 (0.1%). The presence of SARS-CoV-2-infected cells was determined after the addition of primary antibody of anti-SARS-CoV-2 (Invitrogen), followed by secondary antibody of goat anti-mouse IgG conjugated with HPR antibody (Abcam) and use of TrueBlue peroxidase substrate. Foci were read using Autoimmun Diagnostika

V-Spot image analyzer, with virus titer determined by calculating the average spot number and subtraction of background spot count from the negative control wells.

Serum antibody assays. Serum samples were analyzed at Nexelis to determine SARS-CoV-2 anti-spike IgG concentrations by ELISA (reported as ELU ml^{-1}). Neutralizing antibody titers for live SARS-CoV-2 virus (lineage Victoria/01/2020) were determined by microneutralization assay at the UK Health Security Agency and reported as the 50% neutralizing antibody titer (NT$_{50}$). For the microneutralization assay, lower limit of detection (LLOD) was 58, and undetectable samples were assigned a value of 29. For the spike protein IgG ELISA, LLOD was 50.2 ELU ml^{-1}, and undetectable samples were assigned a value of 25 ELU ml^{-1}.

Lateral flow rapid antigen assays. LFAs were performed using the Innova SARS-CoV-2 antigen rapid quantitative test kit (BT1309) as per the manufacturer's recommendations with adaptations as follows. This commercially available kit is designed to detect the presence of the SARS-CoV-2 nucleocapsid protein through in vitro immunochromatographic assays. Viral transport medium from daily throat and mid-turbinate swab samples at days 1–15 after inoculation were tested. Samples (60 μl) were directly added to the sample window of the LFA test device at room temperature. Bands in both control (C) and test (T) windows were indicative of a SARS-CoV-2-positive status. A single control band indicated SARS-CoV-2 negativity.

Statistical analysis. Statistical analysis was performed using GraphPad Prism version 9.2 and R version 4.05. The study was a first-in-human experimental medicine study, and only analyses of VL were pre-specified in the protocol; all others were post hoc. Two-group comparisons were tested using two-sided Mann–Whitney test for unpaired and two-sided Wilcoxon matched-pairs signed-rank test for paired groups. Spearman's rank-order correlation was used for correlation analysis. Exact (Clopper–Pearson) CIs were used for proportions; 95% CI of the median was calculated using the binomial distribution[54]. For all tests, a value of $P < 0.05$ was considered significant. P values were two-sided and unadjusted for multiplicity, as these investigations were exploratory. In general, missing data rarely occurred and were not imputed, and summary statistics were reported on observed data. Two nose qPCR VL data points (one in an infected participant and one in an uninfected participant on morning day 3 after inoculation) were invalid; there were no other missing data.

No formal sample size calculation was performed for this early-stage dose-finding study. However, a sample size of up to an expected 30 participants for a dose-level and treatment regimen was thought to be sufficient to meet the primary objective of escalating/expanding the dose in a safe manner while providing information on the attack rate.

The LLOQ for qPCR was 3 log$_{10}$ copies per milliliter, with positive detections less than the LLOQ assigned a value of 1.5 log$_{10}$ copies per milliliter and undetectable samples assigned a value of 0 log$_{10}$ copies copies per milliliter. For FFA, LLOQ was 1.27 FFU ml^{-1}; viral detection less than the LLOQ was assigned 1 log$_{10}$ FFU ml^{-1}; and undetectable samples were assigned 0 log$_{10}$ FFU ml^{-1}. AUCs for VL (Figs. 2e and 4g and Supplementary Figs. 2c and 7b) and total symptom score (Fig. 3e,f) were calculated using the trapezoid rule on all collected data between 24 hours after inoculation and discharge from quarantine. AUCs for VL were calculated for both qPCR and FFA measurements, using VL on a linear scale (rather than summing log$_{10}$ VL) with the derived AUCs presented on a log$_{10}$ scale.

Logistic regression models to predict LFA positivity from log$_{10}$ VL measurements (qPCR or FFA) were used. Models were fitted using general estimating equations[55] to control for repeated within-participant assessments, using R with the geepack, ggeffects and pROC packages. Quasi-likelihood information criterion-based model selection favored a constant (exchangeable) within-cluster correlation structure. ROC curves were computed for each model, for the fit to all data, for fits to a random 60% (training) sample of the data and for predictions for the remaining 40% (validation) sample of the data.

The mean proportion of the AUC for VL, which would occur on or after the day of the first positive LFA test, was estimated assuming regular asymptomatic LFA testing occurring on a 1–7-day cadence. In doing so, we averaged this proportion over all participants and all possible days of the first test relative to the day of infection. Estimates were computed for nasal and throat measurements separately and for both combined. For the combined analysis, VL measurements for nose and throat were added, and the combined LFA was assumed to be positive if either or both of the nose and throat LFAs were positive. Mean proportions and the 95% CI around the mean were calculated.

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

Data availability
Individual participant data that underlie the results reported in this article after de-identification will be made available for individual participant data meta-analysis beginning 12 months and ending 5 years after article publication upon written request. Data will be shared with investigators whose proposed use

of the data has been approved by the UK Vaccine Taskforce Human Challenge Steering Committee to achieve the aims in the approved proposal. Additional shareable documents include the Statistical Analysis Plan. Proposals should be directed to c.chiu@imperial.ac.uk. To gain access, data requestors will need to complete a data request form and sign a data access agreement.

References
48. Gbesemete, D. et al. Exploring the acceptability of controlled human infection with SARSCoV2—a public consultation. *BMC Med.* **18**, 209 (2020).
49. Barker, C. et al. Public attitudes to a human challenge study with SARS-CoV-2: a mixed-methods study. Preprint at https://doi.org/10.12688/wellcomeopenres.17516.1 (2022).
50. Clift, A. K. et al. Living risk prediction algorithm (QCOVID) for risk of hospital admission and mortality from coronavirus 19 in adults: national derivation and validation cohort study. *BMJ* **371**, m3731 (2020).
51. Doty, R. L., Frye, R. E. & Agrawal, U. Internal consistency reliability of the fractionated and whole University of Pennsylvania Smell Identification Test. *Percept. Psychophys.* **45**, 381–384 (1989).
52. Coronavirus Clinical Characterisation Consortium. Site set-up. *ISARIC 4C* https://isaric4c.net/protocols
53. Center for Disease Control and Prevention. CDC's Diagnostic Test for COVID-19 *Only* and Supplies. *COVID-19* https://www.cdc.gov/coronavirus/2019-ncov/lab/virus-requests.html (2021).
54. Zar, J. H. *Biostatistical Analysis* 5th edn (Pearson, 2010).
55. Liang, K.-Y. & Zeger, S. L. Longitudinal data analysis using generalized linear models. *Biometrika* **73**, 13–22 (1986).

Author contributions
C. Chiu., B.K., A.M., A.P.C., J.S.N.-V.-T., G.R., P.O., H.M., T.W., E.S., M.J., C.H., S.H., A.G., M.G.S., N.M.F., C. Charman, P.M., K.L., R.C.R. and W.S.B. conceived and/or designed the work. C. Chiu, B.K., M.K., A.M., A.P.C., J.Z., A.B., N.G., S.S.H., J.B., R.F., N.N., B.L., K.L., C. Charman and P.M. conducted the study and/or collected data.

a

Nose

b

Throat

c

d

**Extended Data Fig. 1 | Isolated viral detections in participants deemed uninfected following human SARS-CoV-2 challenge.** Healthy adult volunteers were challenged intranasally with SARS-CoV-2. Viral load in twice-daily (**a**) nose and (**b**) throat swab samples was measured by qPCR. Infection was defined by two consecutive quantifiable viral detections by qPCR from day 1 post-inoculation onwards. Individuals who did not meet these criteria are shown (n = 18 biologically independent subjects). (**c**) Change in peak self-reported symptoms of each type and (**d**) UPSIT score in uninfected individuals are shown over the 28 days post-inoculation. Data are presented as mean +/- S.E.M.
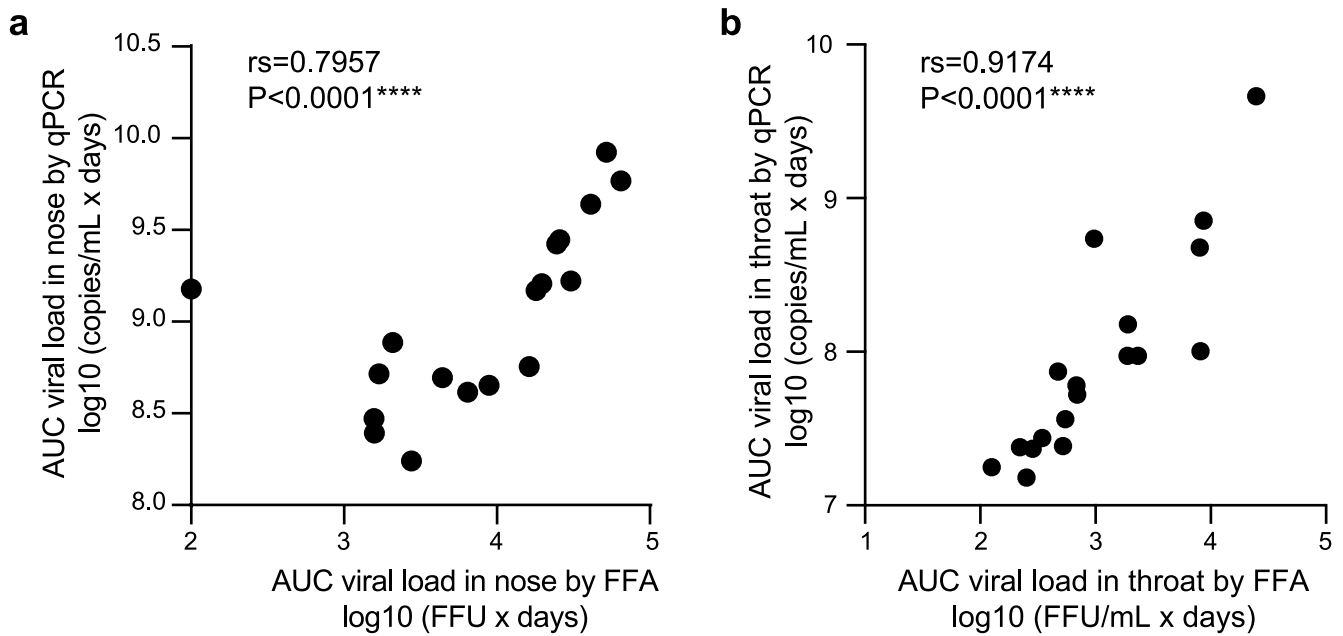
Extended Data Fig. 2 | VL by qPCR and FFA correlate strongly. Twice-daily nose and throat swabs were assayed by qPCR and FFA for SARS-CoV-2 in infected individuals (n = 18 biologically independent subjects). Trapezoidal area under the curve (AUC) was calculated and Spearman's correlation between qPCR and FFA was performed in nose and throat. Spearman's r (rs) and p-values are shown.
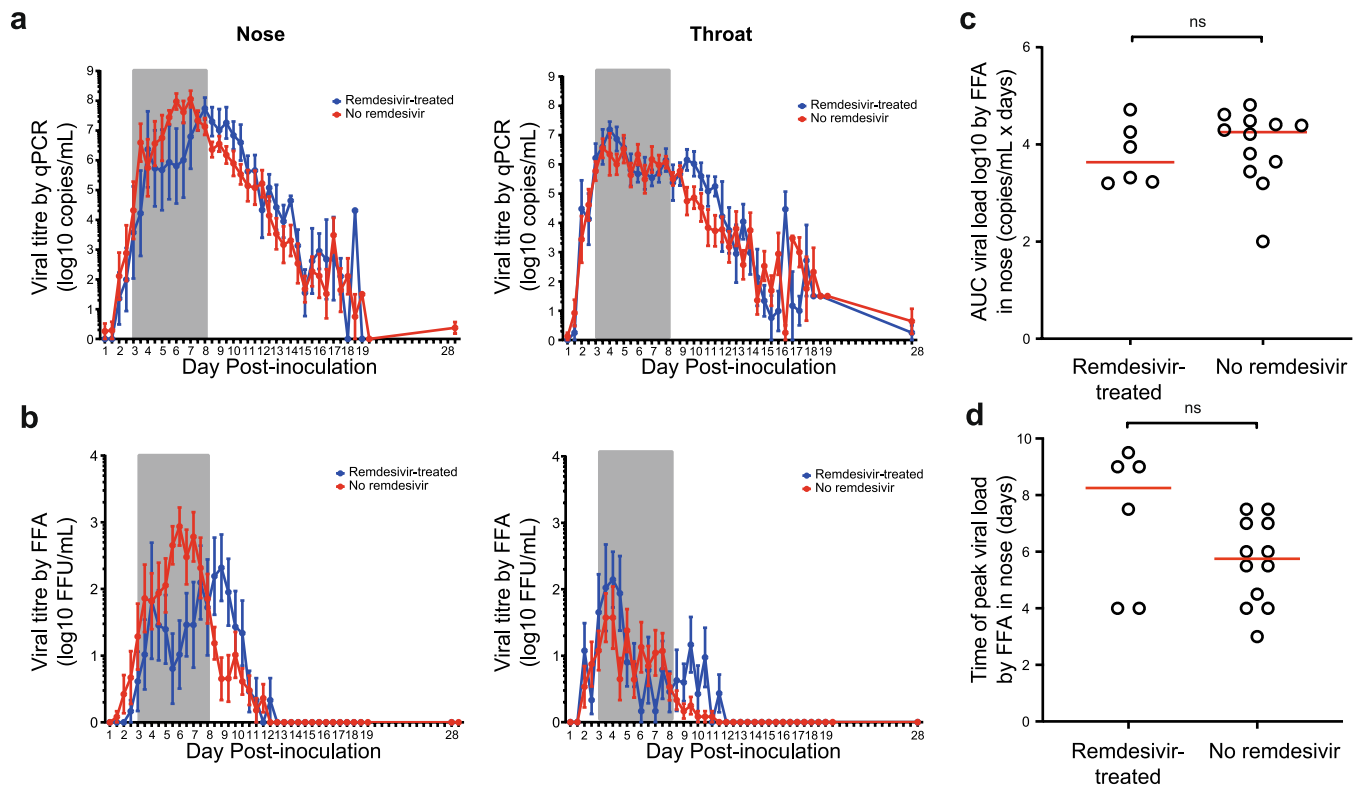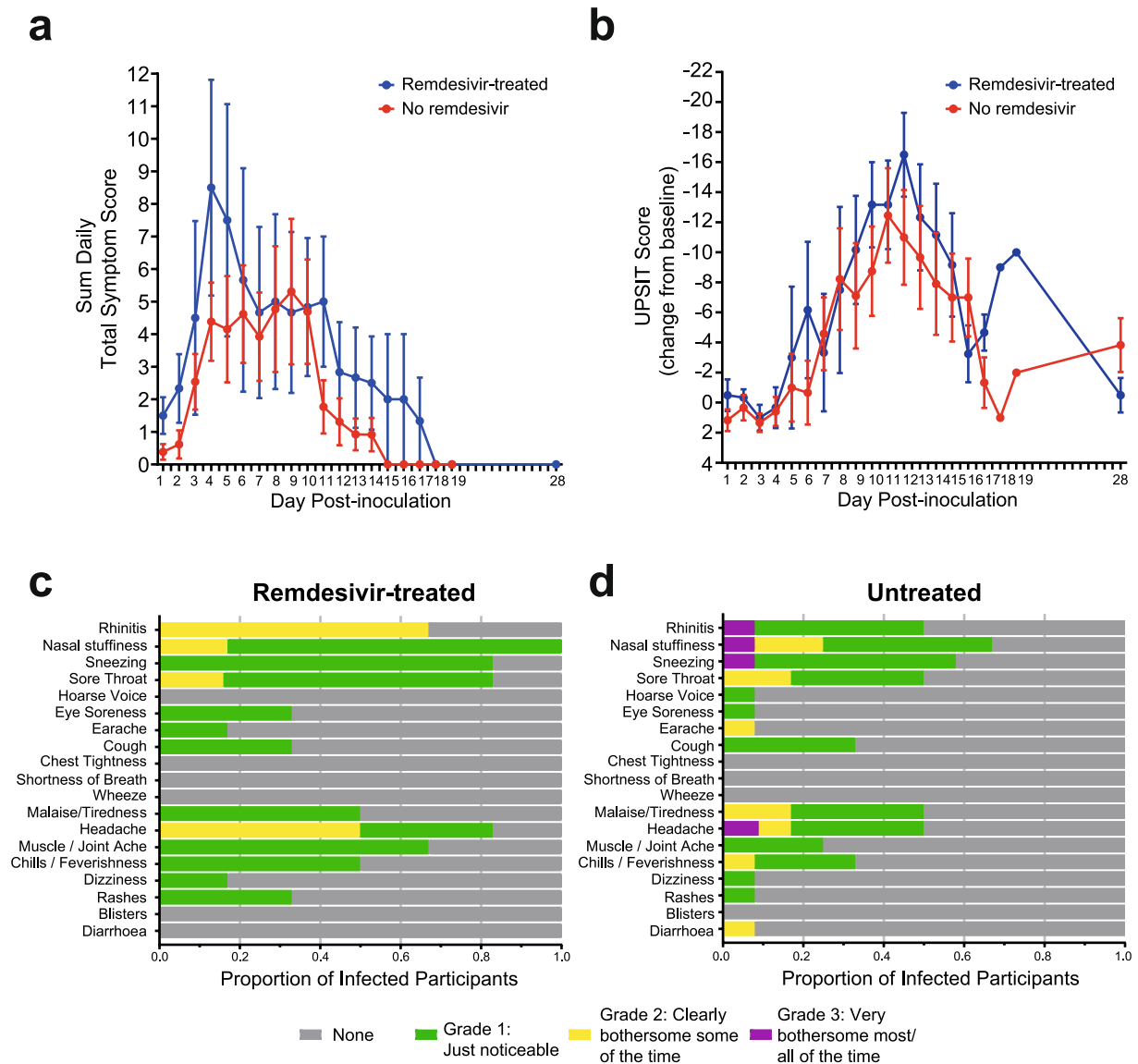
Extended Data Fig. 3 | Pre-emptive remdesivir treatment is associated with no statistically significant changes in VL following human SARS-CoV-2 challenge. Healthy adult volunteers were challenged intranasally with SARS-CoV-2. VL was measured twice-daily in nose and throat swab samples by (**a**) qPCR and (**b**) FFA from remdesivir-treated (blue) and untreated (red) participants (n=6 biologically independent subjects). Results are expressed as means+/- S.E.M. Grey shading indicates the average 5-day remdesivir treatment period. AUC of VL by FFA (**c**) and days to peak VL by FFA (**d**) were compared between remdesivir-treated and untreated groups using two-sided Mann-Whitney tests (ns=non-significant).
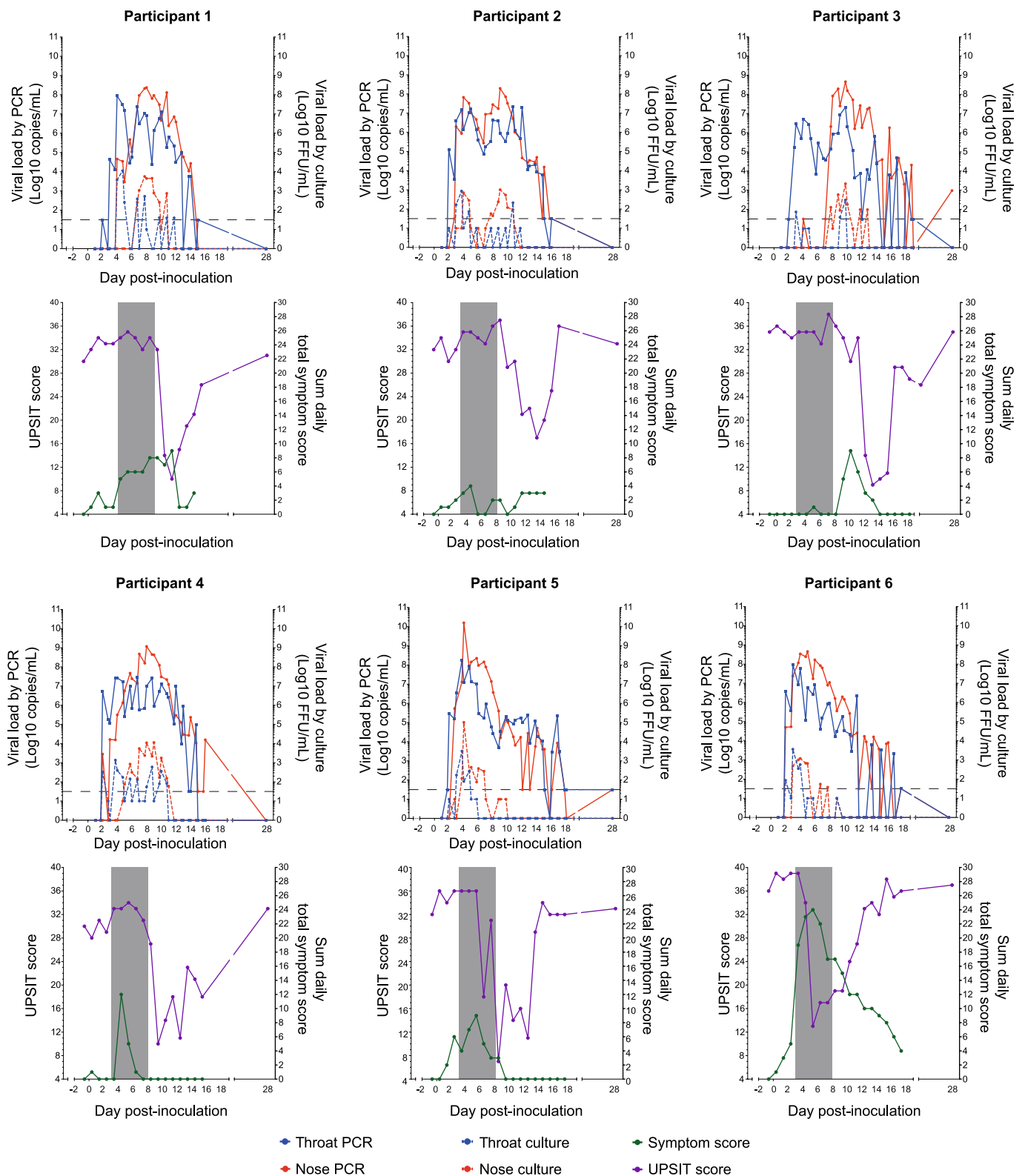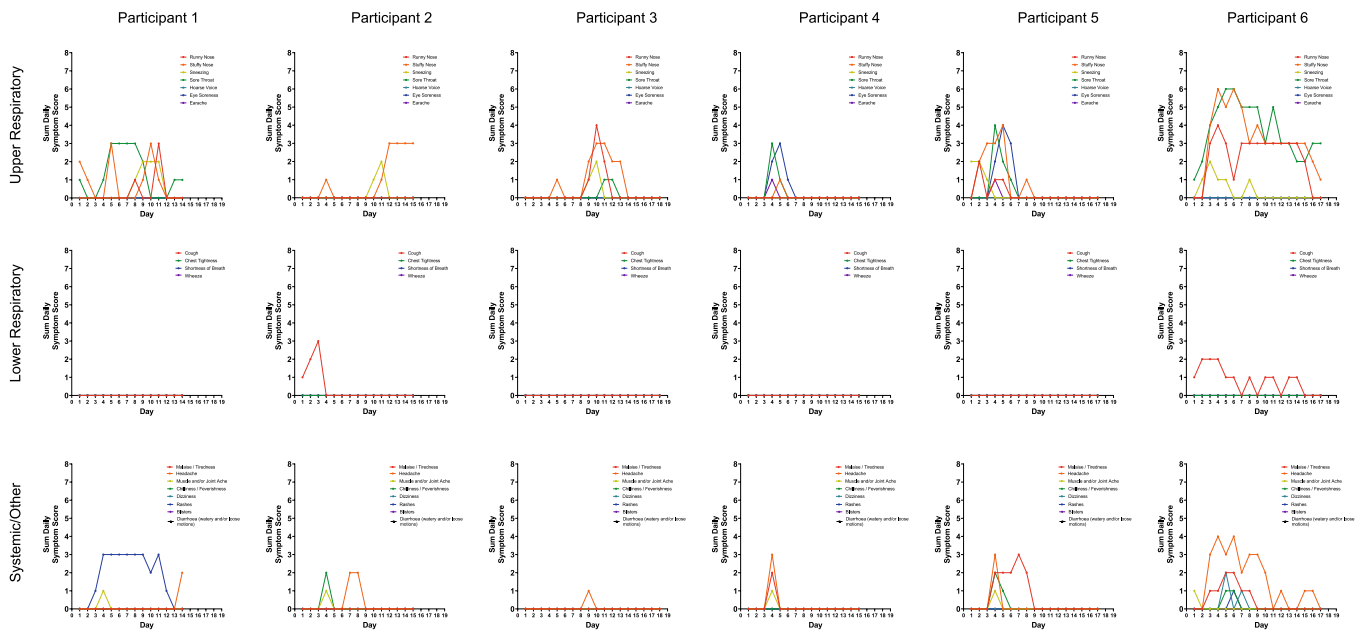
Extended Data Fig. 4 | Human SARS-CoV-2 challenge infection treated pre-emptively with remdesivir causes a range of mild-moderate symptoms. Six individuals developed PCR-confirmed infection following inoculation with SARS-CoV-2 and were treated with Remdesivir for 5 days after 2 consecutive viral detections (grey shading). Nose swabs were analysed by qPCR (red lines) and viral culture (red dotted lines). Throat swabs were analysed by qPCR (blue lines) and viral culture (blue dotted lines). Objective smell testing was achieved using the University of Pennsylvania Smell Identification Test (UPSIT, purple lines). Symptoms were assessed using self-reported symptom diaries 3 times daily and shown as the total for each day (blue lines). Subjective smell disturbance was assessed daily (anosmia in dark orange, partial smell reduction in light orange). Dashed horizontal lines represent the lower limit of quantification of the qPCR assay. Grey shaded area represents the period of remdesivir treatment.
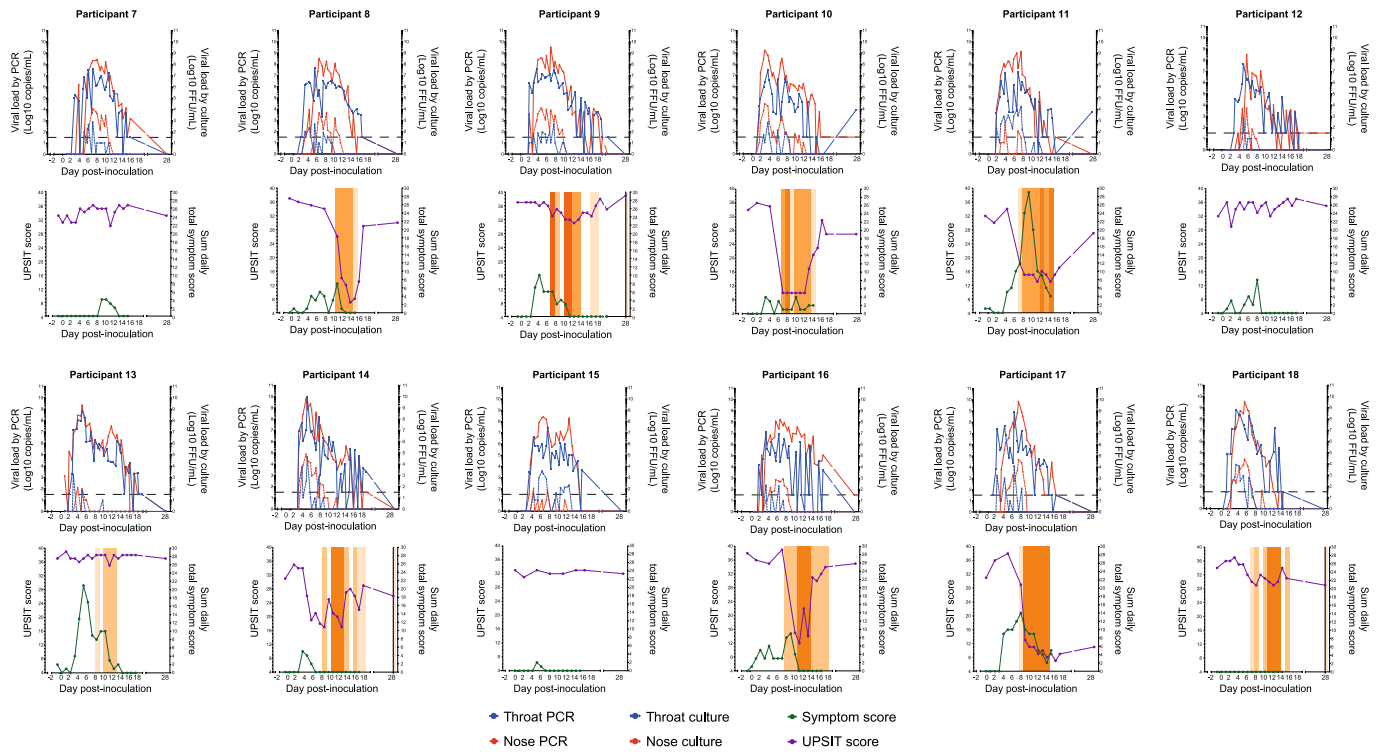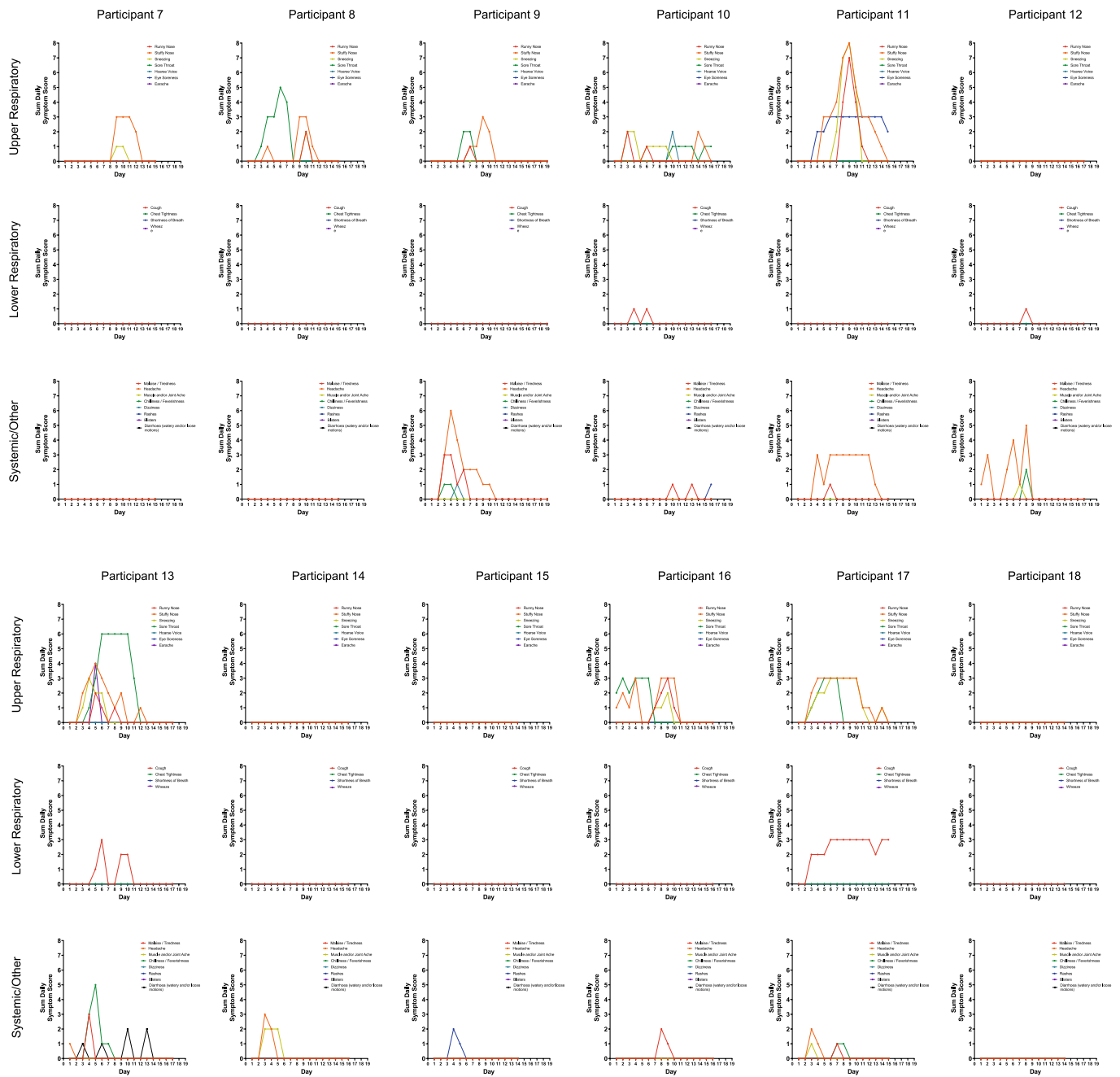
Extended Data Fig. 5 | Human SARS-CoV-2 challenge infection causes a range of mild-moderate symptoms without remdesivir treatment. Twelve participants were inoculated with SARS-CoV-2 and developed PCR-confirmed infection. Nose swabs were analysed by qPCR (red lines) and viral culture (red dotted lines). Throat swabs were analysed by qPCR (blue lines) and viral culture (blue dotted lines). Objective smell testing was achieved using the University of Pennsylvania Smell Identification Test (UPSIT, purple lines). Symptoms were assessed using self-reported symptom diaries 3 times daily and shown as the total for each day (blue lines). Subjective smell disturbance was assessed daily (anosmia in dark orange, partial smell reduction in light orange). Dashed horizontal lines represent the lower limit of quantification of the qPCR assay.

Extended Data Fig. 6 | Pre-emptive remdesivir treatment is not associated with differences in symptom severity following human SARS-CoV-2 challenge. Symptom scores were collected using self-reported symptom diaries 3 times daily. (**a**) Mean +/- S.E.M. symptom scores are shown as the total scores for each day. Remdesivir-treated individuals (n = 6 biologically independent subjects) are shown in blue and untreated individuals (n = 12 biologically independent subjects) in red. (**b**) Smell disturbance was assessed using the University of Pennsylvania Smell Identification Test (UPSIT) in remdesivir-treated (blue) and untreated (red) infected participants. Data are presented as mean +/- S.E.M. The frequency and peak severity of symptoms of each type reported by infected participants (**c**) treated and(**d**) untreated with remdesivir are shown, scored as follows: none (grey), grade 1 just noticeable (green), grade 2 clearly bothersome some of the time (yellow and grade 3 very bothersome most/all of the time (purple).

Extended Data Fig. 7 | SARS-CoV-2 human challenge infection with remdesivir treatment causes mild-to-moderate upper, lower and systemic symptoms only. Six healthy adult volunteers inoculated with SARS-CoV-2 intranasally developed PCR-confirmed infection and were treated with remdesivir. Upper, lower and systemic symptoms were ascertained by self-reported symptom diary.

Extended Data Fig. 8 | SARS-CoV-2 human challenge infection causes mild-to-moderate upper, lower and systemic symptoms only. Twelve healthy adult volunteers inoculated with SARS-CoV-2 intranasally developed PCR-confirmed infection with no treatment administered. Upper, lower and systemic symptoms were ascertained by self-reported symptom diary.
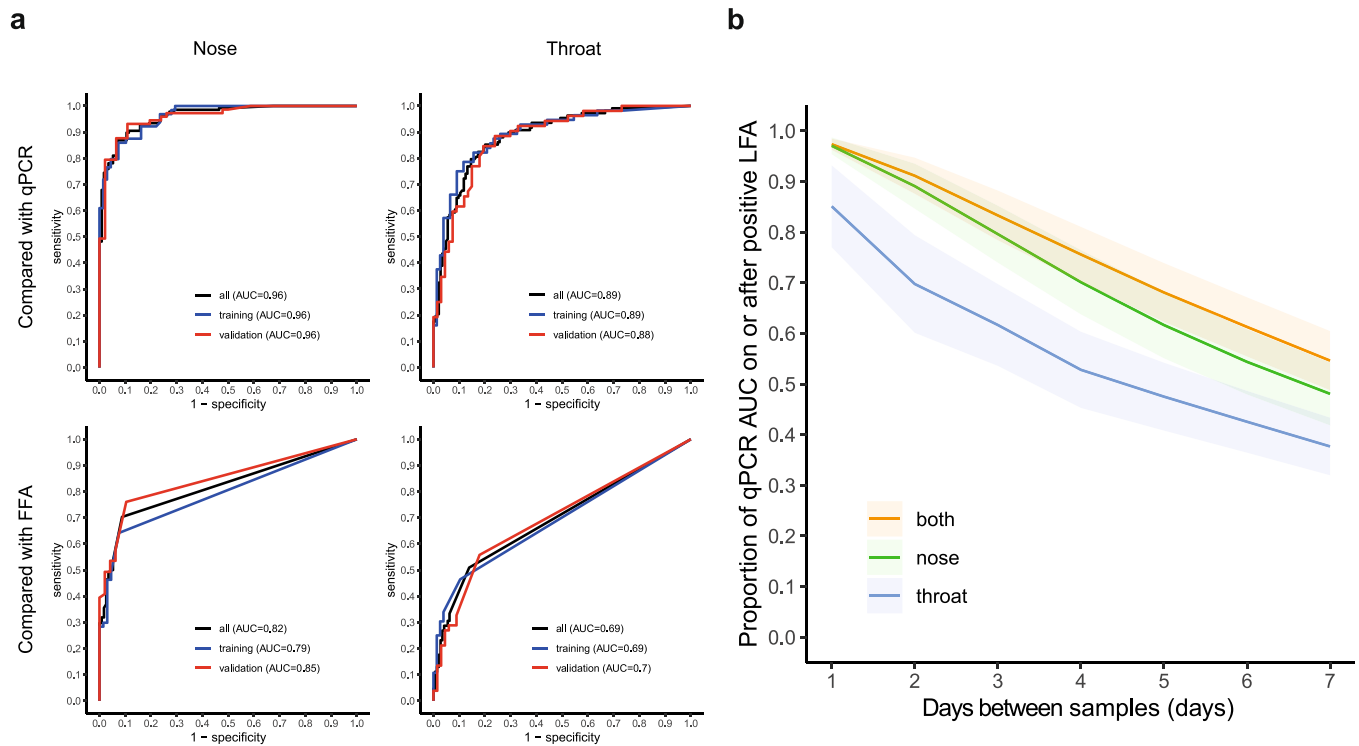
Extended Data Fig. 9 | Lateral flow assay is sensitive and specific for SARS-CoV-2 detection. Nose and throat swab samples in viral transport medium were tested by lateral flow assay (LFA) in n = 18 biologically independent subjects. (**a**) Receiver operating characteristic curves for LFA to predict qPCR and FFA positivity. (**b**) Impact of frequency of LFA testing on the proportion of qPCR quantified virus after first LFA diagnosis from the nose (green), throat (blue) or combined nose and throat (orange). Mean and 95% confidence intervals are shown.

nature research

Corresponding author(s): Prof Christopher Chiu

Last updated by author(s): Nov 30, 2021

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	No software was used to collect data
Data analysis	Graphpad Prism v9.2, R v4.05 with the geepack, ggeffects and pROC packages

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

Individual participant data that underlie the results reported in this article after deidentification will be made available for individual participant data meta-analysis beginning 12 months and ending 5 years following article publication upon written request. Data will be shared with investigators whose proposed use of the data has been approved by the UK Vaccine Taskforce Human Challenge Steering Committee to achieve the aims in the approved proposal. Additional shareable documents include the Statistical Analysis Plan. Proposals should be directed to c.chiu@imperial.ac.uk. To gain access, data requestors will need to complete a data request form and sign a data access agreement. The sequence of the challenge virus has been deposited in Genbank (Accession number OM294022).

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	The primary objective of the study was to identify a safe and infectious dose of wild type SARS-CoV-2 in healthy volunteers, suitable for future intervention studies. No formal sample size calculation was therefore performed for this early-stage dose finding study. However, a sample size of up to an expected 30 subjects for a dose level and treatment regimen was felt sufficient to meet the primary objective of escalating/expanding the dose in a safe manner whilst providing information on the attack rate. The cohort expansion criterion used was based on targeting a 70% target attack rate, but also reducing the risk of seeing an attack rate of less than 50%, after an expected 30 subjects. This was similar to previous human infection dose escalation studies. Although the observed attack rate was less than 70% after the first cohorts, expansion was undertaken at the "most promising" dose. Since a 70% attack rate was not reached after the first cohorts the "≥70% infected" criterion became a "guide". For the level of precision, once say 30 subjects were recruited, and an observed 21 out of 30 (70%) subjects were infected then this would have provided a 95% CI of (51%, 85%), and therefore this would have provided the necessary CI width to have the lower bound above the required 50%. In addition to this, by recruiting the later cohort (e.g. 30 subjects in total), this allowed for inspection of any variation in the attack rate between the cohorts at the particular dose level, and also increased the chance of potentially seeing any less frequent adverse events.
Data exclusions	Two virology datapoints were invalidated due to laboratory error. These are indicated in the manuscript. No other data were excluded from the analysis
Replication	The study took place over 5 distinct quarantine groups with similar findings across all studies
Randomization	None; participants were enrolled into the study chronologically as they completed screening and according to their availability for at least 2 weeks quarantine. All participants were inoculated with challenge virus in this open-label study.
Blinding	Study participants and participant-facing staff conducting the clinical study were blinded to virologically-confirmed infection status where possible until protocol-defined procedures (i.e. pre-emptive remdesivir treatment and day 10 thoracic CT scan) made this apparent. Study participants and clinical staff could not be blinded to the clinical data and symptoms scores there were generating and collecting. Laboratory staff processing blood and respiratory specimens were blinded to infection status. Laboratory staff conducting and analysing qPCR, FFA and LFA virology assays were not blinded to infection status as this was defined by the results of these assays. Laboratory staff undertaking antibody measurements were blinded.

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

Antibodies

Antibodies used	Anti-spike rabbit polyclonal antibody (Sino Biologicals, cat. no. 40592-T62) at 1 in 2000, Anti-rabbit IgG goat–horseradish peroxidase conjugate (Invitrogen, cat. no. G21234) at 1 in 4000.
Validation	Primary and secondary detection antibodies were validated for the application by the National Infection Service, UK Health Security Agency, Salisbury, UK and published in Bewley, K.R., Coombes, N.S., Gagnon, L. et al. Quantification of SARS-CoV-2 neutralizing antibody by wild-type plaque reduction neutralization, microneutralization and pseudotyped virus neutralization assays. Nat Protoc 16, 3114–3140 (2021). https://doi.org/10.1038/s41596-021-00536-y.

Eukaryotic cell lines

Cell line source(s)	Vero (WHO) cells purchased from the European collection of authenticated cell cultures (ECACC, catalogue no. 88020401).
Authentication	Vero cells were authenticated by STR profile analysis
Mycoplasma contamination	Vero cells were tested negative for Mycoplasma by culture isolation, Hoechst DNA staining and PCR
Commonly misidentified lines (See ICLAC register)	No commonly misidentified cell lines were used in the study

Human research participants

Population characteristics	Sero-suitable (no evidence of COVID 19 infection or previous vaccination) healthy male and female volunteers 18-30 years of age (inclusive) with no known risk factors for severe COVID-19
Recruitment	Screening of potential participants took place in two stages with an initial screening visit, followed by a study specific remote consultation to go through the full study participant information following adequate time for the informed consent form (ICF) and participation in the study to be considered. Screening visits took place between Day -90 to Day -2. Potential participants were screened under a separate study-specific screening protocol using a screening ICF and advertising material that was approved by the Research Ethics Committee (REC) and Health Research Authority (HRA). Screening activities under the separate screening protocol continued up until subjects sign the study specific consent. Recruitment was done through a number of channels: • Approved advertising, including social media • hVIVO volunteer database (Volunteers already registered with any other hVIVO database may be contacted to determine their interest in participating in SARS-CoV-2 research.) • Referral • Organic search (e.g. via Google or other search engines) The participant sample was biased by the age criteria (18-30 years) and requirement to be healthy with no co-morbidities or known risk factors for severe COVID-19 based on clinical history, blood tests and radiology. There was potential self-selection bias as participation was voluntary and instigated by the volunteers. Due to these factors, direct extrapolation of the results to young children, older adults, those with pre-existing conditions and minority groups may not be possible.
Ethics oversight	This study was conducted in accordance with the protocol, the Consensus ethical principles derived from international guidelines including the Declaration of Helsinki and Council for International Organizations of Medical Sciences (CIOMS) International Ethical Guidelines, applicable ICH Good Clinical Practice guidelines, applicable laws and regulations. The screening protocol and main study were approved by the UK Health Research Authority – Ad Hoc Specialist Ethics Committee (reference: 20/UK/2001 and 20/UK/0002).

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

Clinical data

All manuscripts should comply with the ICMJE guidelines for publication of clinical research and a completed CONSORT checklist must be included with all submissions.

Clinical trial registration	Clinicaltrials.gov NCT04865237
Study protocol	Following acceptance of the manuscript we will upload the protocol to the protocol exchange portal
Data collection	The study was conducted at the Queen Mary BioEnterprises (QMB) Innovation Centre, London, UK (outpatient screening and follow-up visits) and Royal Free London NHS Trust, London, UK (in-patient quarantine). The first date of participant enrollment was 6th March 2021 and the last was 8th July 2021. Data collection occurred at: Study specific screening Day -90 to Day -2 Quarantine Phase Day -2 to Day 14 (+ extended days) Follow up visits Day 28 (+/-3 days), Day 90 (+/- 7 days), Day 180 (+/- 14 days), Day 270 (+/- 14 days) and Day 360 (+/- 14 days)
Outcomes	Primary Objective /Endpoint • To identify a safe and infectious dose of wild type SARS-CoV-2 in healthy volunteers, suitable for future intervention studies, that: • has an acceptable safety profile as measured by: o Occurrence of Adverse Events (AEs) within 30 days post-viral challenge (Day 0) up to Day 28 follow up. o Occurrence of Serious Adverse Events (SAEs) from the viral challenge (Day 0) up to Day 28 follow up. • induces laboratory confirmed infection in ≥50% of participants

(ideally between 50% and 70%). Laboratory confirmed
infection is defined by:
o Two quantifiable greater than lower limit of
quantification (viral load ≥LLOQ) RT-PCR
measurements from mid turbinate and/or throat
samples, reported on 2 or more consecutive
timepoints, starting from 24 hours post-inoculation and
up to discharge from quarantine.

Secondary
Objectives/endpoint
• To further assess SARS-CoV-2 viral infection rates in upper
respiratory samples by qRT-PCR and cell culture
• To assess the incidence of symptomatic SARS-CoV-2 infection
• To assess the SARS-CoV-2 viral dynamics in upper respiratory
samples (AUC, peak, duration, incubation period)
• To assess the SARS-CoV-2 induced symptoms (Sum, AUC,
peak, peak daily, frequency)
• To assess the incidence of SARS-CoV-2 illness (Upper
Respiratory Tract illness [URTI], Lower Respiratory Tract illness
[LRTI], Systemic Illness (SI), Febrile Illness [FI], grade 1, 2 & 3
symptoms)

Exploratory/Tertiary
Objectives /
Endpoints
• To explore the safety of the wild type SARS-CoV-2 human
challenge model (smell, cognition, pulmonary changes [CT,
spirometry, FOT], safety laboratory tests, blood type, concomitant
medications)
• To explore SARS-CoV-2 viral infection rates in saliva, by qRTPCR
and cell culture
• To explore the SARS-CoV-2 viral dynamics in saliva, by qRT-PCR
and cell culture (AUC, peak, duration, incubation period)
• To explore the host-pathogen relationship in the SARS-CoV-2
human challenge model (including humoral and cellular immunity,
proteomics, transcriptomics, host and viral genomics, microbiome
and systems biology)
• To explore the Minimal Clinically Important Difference (MCID) in
instrument change (e.g. Symptom diary cards)
• To explore environmental contamination in the SARS-CoV-2
human challenge model (quantitation and detection of virus in air
sampling, exhaled breath and surface swabbing)