SARS-CoV-2 infection and reinfection in a seroepidemiological workplace cohort in the United States

Emilie Finch[1‡], Rachel Lowe[1], Stephanie Fischinger[2,3], Michael de St Aubin[4], Sameed M. Siddiqui[5,6], Diana Dayal[7], Michael A Loesche[7,8], Justin Rhee[7], Samuel Berger[7], Yiyuan Hu[7], Matthew J. Gluck[7], Benjamin Mormann[7], Mohammad A. Hasdianda[8], Elon R. Musk[7], Galit Alter[2,6], Anil S. Menon[7], Eric J. Nilles[4,8*], Adam J. Kucharski[1*] on behalf of the CMMID COVID-19 working group and the SpaceX COVID-19 Cohort Collaborative*

[1]Centre for Mathematical Modelling of Infectious Diseases, London School of Hygiene & Tropical Medicine
[2]Ragon Institute of MGH, MIT and Harvard, Cambridge, MA, USA
[3]Institut für HIV Forschung, Universität Duisburg-Essen, Duisburg, Germany
[4]Harvard Humanitarian Initiative, Cambridge, MA, USA
[5]Computational and Systems Biology Program, Massachusetts Institute of Technology, Cambridge, MA, USA
[6]Broad Institute of MIT and Harvard, Cambridge, MA, USA
[7]Space Exploration Technologies Corp
[8]Brigham and Women's Hospital, Boston

[‡] Author for correspondence
** Equal contribution*

Identifying the extent of SARS-CoV-2 reinfection is crucial for understanding possible long-term epidemic dynamics (1). Existing studies examining the relative risk of reinfection in antibody positive individuals have typically involved specific cohorts (such as healthcare workers) who may not be representative of the wider community (2–5).To evaluate the risk of SARS-CoV-2 infection and reinfection over time, we analysed PCR and serological testing data from a prospective community cohort of 4411 employees at SpaceX in four states in the USA, with ages ranging from 18 to 71, between April 2020 and February 2021 (6). 309 individuals tested seropositive during the study period. We defined a possible reinfection as a new positive PCR test more than 30 days after initial seropositive result. This identified 14 possible reinfections with a median time of 66·5 days between initial seropositive result and subsequent PCR positive result.

Our analysis addressed two key sources of bias and uncertainty in estimating reinfection risk. First, confounders may inflate estimates; if a specific subset of the cohort is at higher risk of infection (e.g. due to underlying health conditions or increased risk of exposure), these participants will be more likely to be both initially seropositive and to have a subsequent reinfection. Second, the time period considered could increase uncertainty; defining the baseline seroprevalence at an early time point means few will be seropositive, whereas defining it at a later point means there is less time to observe possible reinfections. We accounted for these two factors by using logistic regression to identify predictors of baseline seropositivity (i.e. infection risk) in order to calculate an adjusted odds ratio for reinfection, and then performing a sensitivity analysis to identify the optimum cut-off date to define baseline seroprevalence (see Appendix).

We estimated an adjusted odds ratio of 0·09 (95% CI: 0·005–0·48) for reinfection, with the week of 26th July 2020 as the optimal baseline time point in a study period between April 2020 and February 2021. This suggests that the presence of SARS-CoV-2 antibodies at baseline is associated with around 91% reduced odds of a subsequent PCR positive test. Our findings are consistent with estimates of 0·17 (95% CI 0·13-0·24) odds ratio [2] and 0·11 (0·03-0·44) incidence rate ratio [3] for healthcare workers and 0·18 (0·11-0·28) incidence rate ratio for military recruits [4]. As well as quantifying reinfection risk over a six month period among a prospectively followed workplace population, our study highlights the importance of accounting for both individual-level heterogeneity in infection risk and population-level variation in epidemic dynamics when assessing the potential for reinfections.

Funding: AJK was supported by Wellcome Trust (grant: 206250/Z/17/Z) and NIHR (NIHR200908). EF was funded by the Medical Research Council (grant number MR/N013638/1), RL is funded by a Royal Society Dorothy Hodgkin Fellowship. The work was also supported by the Translational Research Institute for Space Health through NASA Cooperative Agreement (NNX16AO69A), the Massachusetts Consortium on Pathogen Readiness (MassCPR), the NIH (3R37AI080289-11S1, R01AI146785, U19AI42790-01, U19AI135995-02, U19AI42790-01, 1U01CA260476-01), and the Musk Foundation. GA is a founder of Seromyx Systems Inc., a company developing platform technology that describes the antibody immune response. GA's interests were reviewed and are managed by Massachusetts General Hospital in accordance with their conflict-of-interest policies. All other authors declare no competing interests. We thank SpaceX employees Lindsay Chapman, Jordan Steinhart, Suzanne Siebert, and Kyle Meade for their valuable support, in addition to the dedication and commitment of the many SpaceX employees that volunteered to participate in this study.

Competing interests: GA is a founder of Seromyx Systems Inc. MJG, SB, DD, YH, JR, EP, BM, ASM, and ERM are employees of Space Exploration Technologies Corp. All other authors have declared that no conflict of interest exists.

References:

1. Sabino EC, Buss LF, Carvalho MPS, Prete CA, Crispim MAE, Fraiji NA, et al. Resurgence of COVID-19 in Manaus, Brazil, despite high seroprevalence. The Lancet. 2021 Feb 6;397(10273):452–5.
2. Hall V, Foulkes S, Charlett A, Atti A, Monk EJM, Simmons R, et al. Do antibody positive healthcare workers have lower SARS-CoV-2 infection rates than antibody negative healthcare workers? Large multi-centre prospective cohort study (the SIREN study), England: June to November 2020. medRxiv. 2021 Jan 15;2021.01.13.21249642.
3. Lumley SF, O'Donnell D, Stoesser NE, Matthews PC, Howarth A, Hatch SB, et al. Antibody Status and Incidence of SARS-CoV-2 Infection in Health Care Workers. N Engl J Med. 2021 Feb 11;384(6):533–40.
4. Letizia AG, Ge Y, Vangeti S, Goforth C, Weir DL, Kuzmina NA, et al. SARS-CoV-2 seropositivity and subsequent infection risk in healthy young adults: a prospective cohort study. medRxiv. 2021 Jan 29;2021.01.26.21250535.
5. Addetia A, Crawford KHD, Dingens A, Zhu H, Roychoudhury P, Huang M-L, et al. Neutralizing Antibodies Correlate with Protection from SARS-CoV-2 in Humans during a

Fishery Vessel Outbreak with a High Attack Rate. J Clin Microbiol [Internet]. 2020 Oct 21 [cited 2021 Mar 3];58(11). Available from: https://jcm.asm.org/content/58/11/e02107-20
6. Bartsch YC, Fischinger S, Siddiqui SM, Chen Z, Yu J, Gebre M, et al. Discrete SARS-CoV-2 antibody titers track with functional humoral stability. Nat Commun. 2021 Dec;12(1):1018.

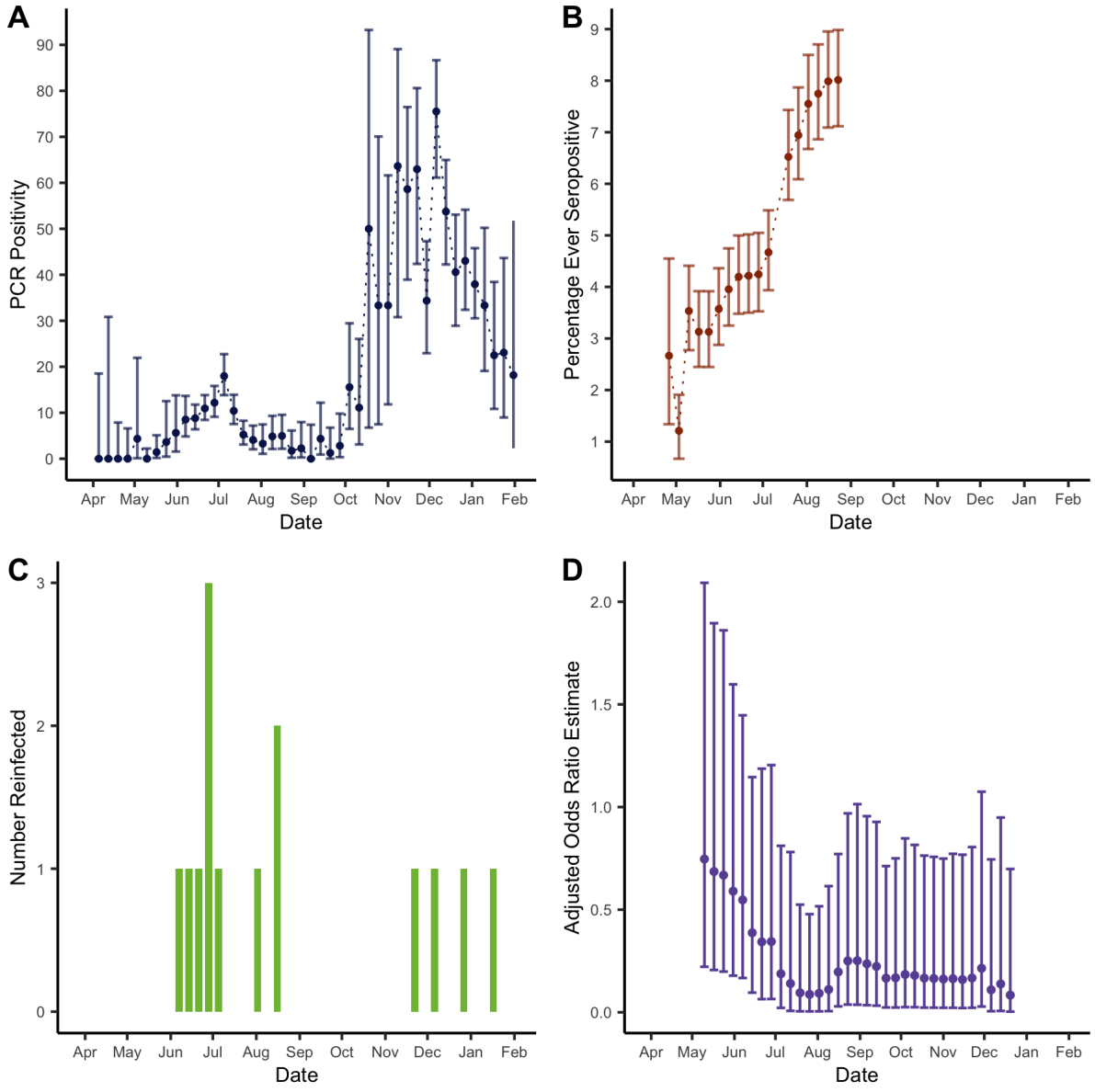


Figure: A) PCR positivity in the cohort between 5th April 2020 and 31st January 2021. B) Percentage ever seropositive in the cohort (number ever seropositive/ cumulative number enrolled) between 29th March 2020 and 23rd August 2020. Note that the percentage ever positive decreases initially as participants continue to be enrolled in the study. C) Number of possible reinfections in cohort over time (defined as a new positive PCR test more than 30 days after initial seropositive result). D) Odds ratio estimates comparing odds of reinfection in the seropositive group with odds of primary infection in the seronegative group, estimated using logistic regression and adjusted for race, ethnicity, state, job category and BMI. The

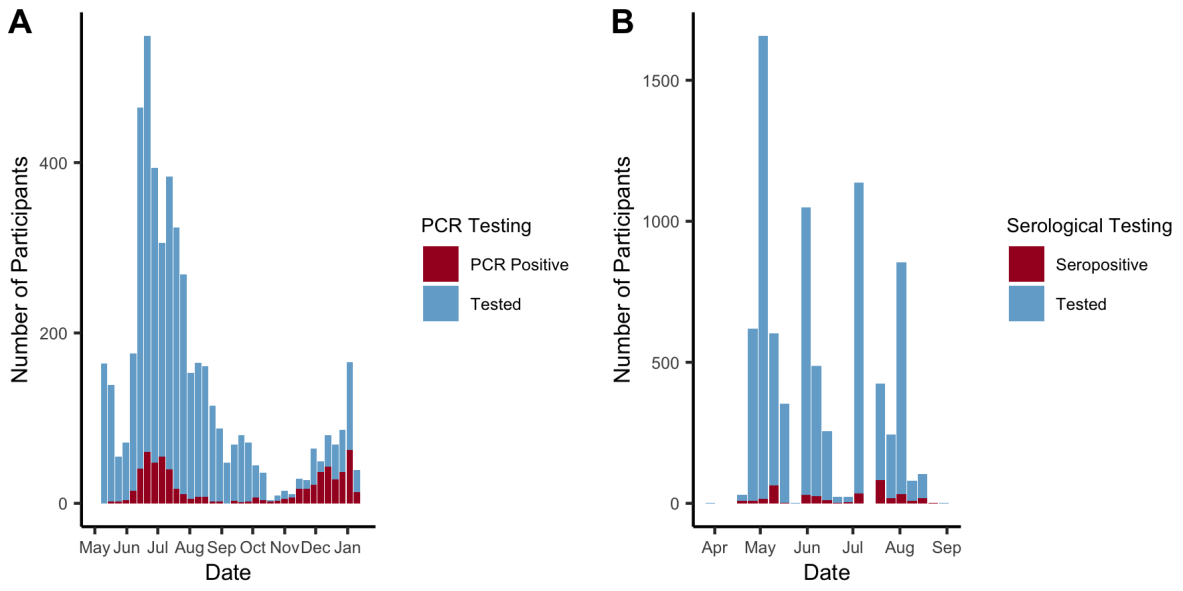
estimates are presented with their associated 95% confidence intervals and with the cut-off week used to define baseline seroprevalence on the x-axis.

Appendix

Cohort description

We used data from a seroepidemiological cohort study of US employees at SpaceX, also described elsewhere (1). In brief, this study involved 4411 employees from four states in the United States who were followed from April 2020. All employees were invited to participate by email and there were no exclusion criteria. Participants occupied a range of job positions within SpaceX including office-based and factory-based jobs. Participants completed a questionnaire including demographic, symptom and exposure information at enrolment, and with each round of serological testing. Participants were offered SARS-CoV-2 IgG receptor-binding domain (RBD) antibody testing with an in-house ELISA assay with 82·4% sensitivity and 99·6% specificity (2). Serological samples were taken during four rounds of testing between April - September 2020. Participants continued to be enrolled throughout the study period at each round of testing and around half of the total participants (48%) were tested at more than one time point.

Additionally, symptomatic and asymptomatic PCR testing were widely available for participants, with PCR testing data available from April 2020 - January 2021. Both serology and PCR testing data were available for 1800 participants over the entire study period.

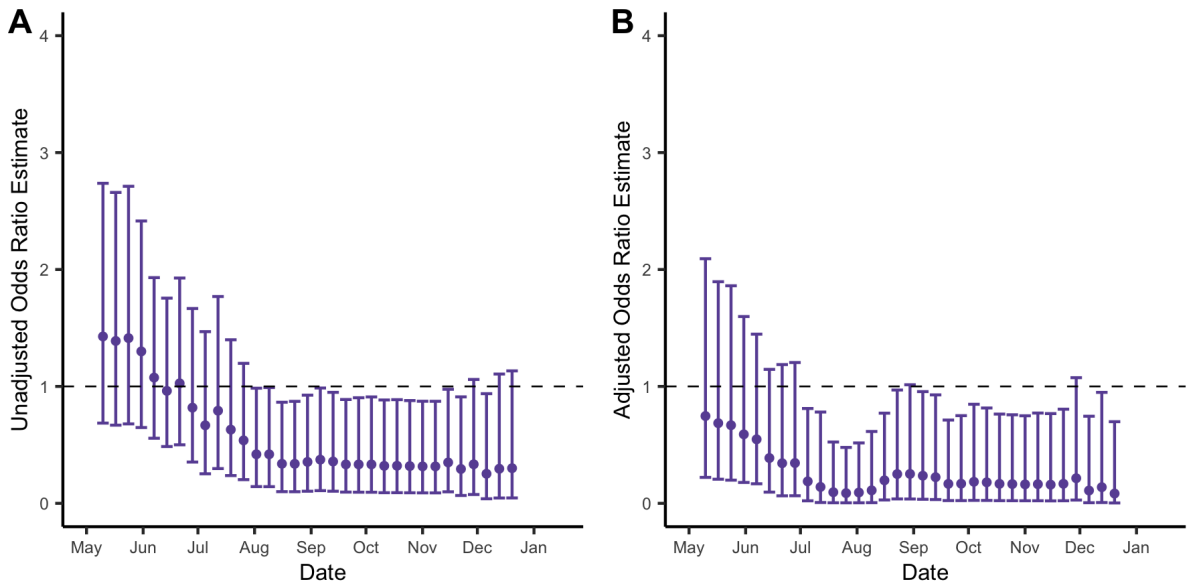


Appendix Figure 1: A) Number of PCR tests and PCR positive tests in the cohort between 5th April 2020 and 31st January 2021. B) Number of serological tests and seropositive tests between 29th March 2020 and 23rd August 2020.

Statistical Analysis

To estimate the risk of SARS-CoV-2 infection in seropositive and seronegative individuals we conducted multivariable logistic regression analysis to assess the association between serological status (the exposure) and subsequent PCR test result, given a test was sought (the outcome). To assess how the odds ratio estimate changed depending on the cut-off week used to define baseline seroprevalence and the subsequent observation period for PCR testing, we conducted a sensitivity analysis using each week in the study period as a cut-off week. For each cut-off week, the analysis was restricted to only include individuals who had a serological test result prior to or including the threshold week, and a PCR test result in any week after the threshold week. We examined whether the propensity to get a PCR test differed by baseline serostatus and found that the percentage PCR tested was broadly similar between seronegative and seropositive individuals for each cut-off week.

We then compared the odds of SARS-CoV-2 reinfection in the seropositive cohort with the odds of infection in the seronegative cohort. We defined a possible reinfection as a PCR positive test > 30 days after their initial seropositive test result (any PCR positive test results ≤ 30 days after the initial seropositive test result were excluded from the analysis). Potential confounding variables were selected a priori and included; age, sex, race, ethnicity, BMI, state, work location, job category, household size, history of chronic disease or history of smoking. We then used logistic regression to identify predictors of baseline seropositivity (i.e. infection risk) using a forwards modelling strategy comparing the Akaike Information Criterion (AIC). Finally, we used goodness of fit statistics, including the AIC, to select the optimally sized variable set to adjust for in the entire analysis (considering all cut-off weeks), by comparing the AIC between models for the same cut-off week for different sized variable sets. Odds ratio estimates were adjusted for race, ethnicity, state, job category and BMI. Age was not strongly predictive in this cohort, likely due to the limited age range of the cohort, and was therefore not included in the final model.



Appendix Figure 2: A) Unadjusted odds ratio estimates comparing odds of reinfection in the seropositive group with odds of primary infection in the seronegative group. The estimates are presented with their associated 95% confidence intervals and with the cut-off week used to define baseline seroprevalence on the x-axis. B) Odds ratio estimates comparing odds of reinfection in the seropositive group with odds of primary infection in the seronegative group, estimated using logistic regression and adjusted for race, ethnicity, state, job category and BMI. The estimates are presented with their associated 95% confidence intervals and with the cut-off week used to define baseline seroprevalence on the x-axis.

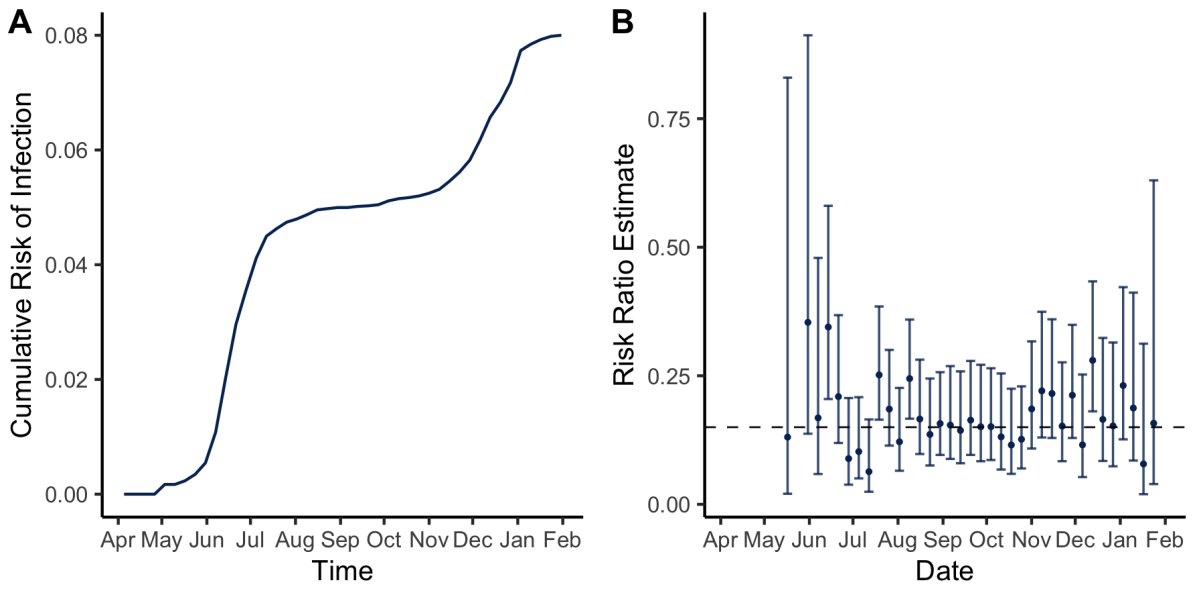
Unadjusted odds ratio estimates tended to overestimate the odds ratio of reinfection compared with primary infection, particularly when using cut-off weeks early in the study period. Notably, in these early threshold weeks the unadjusted analysis estimated a higher odds of reinfection compared to primary infection, albeit with wide confidence intervals (for instance on the week of 24th May, OR = 1·41, 95% CI: 0·7 - 2·71). We hypothesise that individuals who are at higher risk of seroconversion (and thus would be included in analyses at earlier time thresholds) would also then be at higher risk of later reinfection, giving a biased estimate of the effect of antibodies on subsequent infection. While adjusting for confounding removed some of this effect (and reduced all point estimates to below 1), higher odds ratios at earlier time points could be the result of further residual confounding. We estimated an adjusted odds ratio of 0·09 (95% CI: 0·005–0·48) for reinfection, with the week of 26th July 2020 as the optimal baseline time point in a study period between April 2020 and February 2021. The optimal cut-off week was defined in our main analysis as the week associated with the narrowest confidence interval around the odds ratio estimate for reinfection.

Simulation Analysis

In order to validate our choice of optimal cut-off week, we conducted a simulation analysis with a known underlying distribution of the risk of infection and reinfection. To simulate the epidemic dynamics observed in our cohort, we used the cumulative risk of infection derived

from PCR testing data from the study cohort as our underlying distribution of the risk of infection. This showed a bimodal risk of infection with a peak in late June and a second smaller peak in December. This was then scaled so that the overall cumulative risk was 8%, to reflect the level of seropositivity in the cohort by the end of the study period. Considering a sample size of 2000 individuals over a period of 44 weeks, with a pre-set probability of reinfection given seropositivity of 0·15, we obtained risk ratio estimates as shown below (see Appendix Figure 3).

When considering a two-wave epidemic scenario modelled using the cumulative risk of infection in our cohort, we found that the uncertainty in the estimated risk of reinfection was reduced in the middle of the simulation period (i.e. between the two 'waves' of infection risk). This supports our choice of optimal cut-off week for estimating an odds ratio for reinfection in our study cohort. In our main analysis we considered odds ratios given the estimates required an underlying regression analysis.



Appendix Figure 3: A) Cumulative risk of infection in the cohort used for simulation analysis. B) Risk ratio estimates comparing the risk of reinfection with the risk of primary infection. The estimates are presented with their associated 95% confidence intervals and with the cut-off week used to define baseline infection status on the x-axis.

Code to reproduce the figures and simulation analysis presented here and in the main text can be found at https://github.com/EmilieFinch/covid-reinfection.

References

1. Bartsch YC, Fischinger S, Siddiqui SM, Chen Z, Yu J, Gebre M, et al. Discrete SARS-CoV-2 antibody titers track with functional humoral stability. Nat Commun. 2021 Dec;12(1):1018.
2. Roy V, Fischinger S, Atyeo C, Slein M, Loos C, Balazs A, et al. SARS-CoV-2-specific ELISA development. J Immunol Methods. 2020 Sep 1;484–485:112832.

CMMID COVID-19 working group information

The following authors were part of the Centre for Mathematical Modelling of Infectious Disease COVID-19 Working Group. Each contributed in processing, cleaning and interpretation of data, interpreted findings, contributed to the manuscript, and approved the work for publication: Jiayao Lei, Sebastian Funk, Fiona Yueqian Sun, Amy Gimma, Emily S Nightingale, Graham Medley, Sam Abbott, Fabienne Krauer, Nicholas G. Davies, Mark Jit, Akira Endo, Oliver Brady, Anna M Foss, Yung-Wai Desmond Chan, Thibaut Jombart, Kevin van Zandvoort, Rosalind M Eggo, Yang Liu, Gwenan M Knight, Carl A B Pearson, Kaja Abbas, Katherine E. Atkins, Samuel Clifford, Mihaly Koltai, Yalda Jafari, Damien C Tully, Christopher I Jarvis, Kathleen O'Reilly, Nikos I Bosse, Kiesha Prem, Billy J Quilty, Simon R Procter, Rosanna C Barnard, William Waites, Ciara McCarthy, James D Munday, David Hodgson, W John Edmunds, Alicia Rosello, C Julian Villabona-Arenas, Hamish P Gibbs, Stefan Flasche, Timothy W Russell, Sophie R Meakin, Joel Hellewell, Naomi R Waterlow, Matthew Quaife, Frank G Sandmann.

Funding statements for the CMMID COVID-19 working group are as follows. KvZ: KvZ is supported by the UK Foreign, Commonwealth and Development Office (FCDO)/Wellcome Trust Epidemic Preparedness Coronavirus research programme (ref. 221303/Z/20/Z), and Elrha's Research for Health in Humanitarian Crises (R2HC) Programme, which aims to improve health outcomes by strengthening the evidence base for public health interventions in humanitarian crises. The R2HC programme is funded by the UK Government (FCDO), the Wellcome Trust, and the UK National Institute for Health Research (NIHR). SC: Wellcome Trust (grant: 208812/Z/17/Z). FYS: NIHR EPIC grant (16/137/109). SFunk: Wellcome Trust (grant: 210758/Z/18/Z), NIHR (NIHR200908). GFM: NTD Modelling Consortium by the Bill and Melinda Gates Foundation (OPP1184344). YJ: LSHTM, DHSC/UKRI COVID-19 Rapid Response Initiative. SRM: Wellcome Trust (grant: 210758/Z/18/Z). WJE: European Commission (EpiPose 101003688), NIHR (NIHR200908). MQ: European Research Council Starting Grant (Action Number #757699); Bill and Melinda Gates Foundation (INV-001754). NRW: Medical Research Council (grant number MR/N013638/1). RME: HDR UK (grant: MR/S003975/1), MRC (grant: MC_PC 19065), NIHR (grant: NIHR200908). NGD: UKRI Research England; NIHR Health Protection Research Unit in Immunisation (NIHR200929); UK MRC (MC_PC_19065). JYL: Bill & Melinda Gates Foundation (INV-003174). MK: Foreign, Commonwealth and Development Office / Wellcome Trust. FK: Innovation Fund of the Joint Federal Committee (Grant number 01VSF18015), Wellcome Trust (UNS110424). DCT: No funding declared. JDM: Wellcome Trust (grant: 210758/Z/18/Z). AS: No funding declared. AMF: No funding declared. KP: Gates (INV-003174), European Commission (101003688). SFlasche: Wellcome Trust (grant: 208812/Z/17/Z). SA: Wellcome Trust (grant: 210758/Z/18/Z). BJQ: This research was partly funded by the National Institute for Health Research (NIHR) (16/137/109 & 16/136/46) using UK aid from the UK Government to support global health research. The views expressed in this publication are those of the author(s) and not necessarily those of the NIHR or the UK Department of Health and Social Care. BJQ is supported in part by a grant from the Bill and Melinda Gates Foundation (OPP1139859). TJ: RCUK/ESRC (grant: ES/P010873/1); UK PH RST; NIHR HPRU Modelling & Health Economics (NIHR200908). AR: NIHR (grant: PR-OD- 1017-20002). GMK: UK Medical Research Council (grant: MR/P014658/1). MJ: Gates (INV- 003174, INV-016832), NIHR (16/137/109, NIHR200929, NIHR200908), European Commission (EpiPose 101003688). YL: Gates (INV-003174), NIHR (16/137/109), European Commission (101003688). JW: NIHR Health Protection Research Unit and NIHR HTA. JH: Wellcome

Trust (grant: 210758/Z/18/Z). KO'R: Bill and Melinda Gates Foundation (OPP1191821). YWDC: No funding declared. TWR: Wellcome Trust (grant: 206250/Z/17/Z). CIJ: Global Challenges Research Fund (GCRF) project ‘RECAP’ managed through RCUK and ESRC (ES/P010873/1). SRP: Bill and Melinda Gates Foundation (INV-016832). AE: The Nakajima Foundation. ESN: Gates (OPP1183986). NIB: Health Protection Research Unit (grant code NIHR200908). CJVA: European Research Council Starting Grant (Action number 757688). FGS: NIHR Health Protection Research Unit in Modelling & Health Economics, and in Immunisation. AG: European Commission (EpiPose 101003688). KA: Bill & Melinda Gates Foundation (OPP1157270, INV-016832). WW: MRC (grant MR/V027956/1). KEA: European Research Council Starting Grant (Action number 757688). RCB: European Commission (EpiPose 101003688). PK: This research was partly funded by the Royal Society under award RP\EA\180004, European Commission (101003688), Bill & Melinda Gates Foundation (INV-003174). HPG: This research was produced by CSIGN which is part of the EDCTP2 programme supported by the European Union (grant number RIA2020EF-2983-CSIGN). The views and opinions of authors expressed herein do not necessarily state or reflect those of EDCTP. This research is funded by the Department of Health and Social Care using UK Aid funding and is managed by the NIHR. The views expressed in this publication are those of the author(s) and not necessarily those of the Department of Health and Social Care (PR-OD-1017-20001). CABP: CABP is supported by the Bill & Melinda Gates Foundation (OPP1184344) and the UK Foreign, Commonwealth and Development Office (FCDO)/Wellcome Trust Epidemic Preparedness Coronavirus research programme (ref. 221303/Z/20/Z). OJB: Wellcome Trust (grant: 206471/Z/17/Z). WJE: European Commission (EpiPose 101003688), NIHR (NIHR200908, MRC: MC_PC_19065). CIJ: Global Challenges Research Fund (GCRF) project ‘RECAP’ managed through RCUK and ESRC (ES/P010873/1). NGD: UKRI Research England; NIHR Health Protection Research Unit in Immunisation (NIHR200929); UK MRC (MC_PC_19065).