Report 44: Recent trends in SARS-CoV-2 variants of concern in England

Swapnil Mishra*[,1], Sören Mindermann*[,2], Mrinank Sharma*[,3,4,5], Charles Whittaker*[,1], Thomas A Mellan[1], Thomas Wilton[6], Dimitra Klapsa[6], Ryan Mate[6], Martin Fritzsche[6], Maria Zambon[7], Janvi Ahuja[5,8], Adam Howes[9], Xenia Miscouridou[9], Guy P Nason[9], Oliver Ratmann[9], Gavin Leech[10], Julia Fabienne Sandkühler[11], Charlie Rogers-Smith[12], Michaela Vollmer[1], H Juliette T Unwin[1], The COVID-19 Genomics UK (COG-UK) consortium[+], Yarin Gal[2], Meera Chand[7], Axel Gandy[9], Javier Martin[6], Erik Volz[1], Neil M Ferguson*[,1], Samir Bhatt*[,1,13], Jan M Brauner*[,2,5], Seth Flaxman*[,9]

1. Medical Research Council (MRC) Centre for Global Infectious Disease Analysis, Jameel Institute, School of Public Health, Imperial College London, UK
2. Oxford Applied and Theoretical Machine Learning (OATML) Group, Department of Computer Science, University of Oxford, UK
3. Department of Statistics, University of Oxford, UK
4. Department of Engineering Science, University of Oxford, UK
5. Future of Humanity Institute, University of Oxford, UK
6. National Institute for Biological Standards and Control (NIBSC), UK
7. Public Health England, London, UK
8. Medical Sciences Division, University of Oxford, UK
9. Department of Mathematics, Imperial College London, UK
10. Department of Computer Science, University of Bristol, UK
11. Department of Psychology, University of Bonn, Germany
12. OATML Group (work done while at OATML as an external collaborator), Department of Computer Science, University of Oxford, UK
13. Section of Epidemiology, Department of Public Health, University of Copenhagen, Denmark

*: equal contribution
Correspondence: s.mishra@imperial.ac.uk, s.bhatt@imperial.ac.uk
[+]: Full list of consortium names and affiliations are in the appendix

SUGGESTED CITATION

S Mishra, S Mindermann, M Sharma, *et al*. Recent trends in SARS-CoV-2 variants of concern in England. Imperial College London (20-05-2021), doi: https://doi.org/10.25561/88876

Summary

Since its emergence in Autumn 2020, the SARS-CoV-2 Variant of Concern (VOC) B.1.1.7 rapidly became the dominant lineage across much of Europe. Simultaneously, several other VOCs were identified globally. Unlike B.1.1.7, some of these VOCs possess mutations thought to confer partial immune escape. Understanding when, whether, and how these additional VOCs pose a threat in settings where B.1.1.7 is currently dominant is vital. This is particularly true for England, which has high coverage from vaccines that are likely more protective against B.1.1.7 than some other VOCs. We examine trends in B.1.1.7's prevalence in London and other English regions using passive-case detection PCR data, cross-sectional community infection surveys, genomic surveillance, and wastewater monitoring. Our results suggest shifts in the composition of SARS-CoV-2 lineages driving transmission in England between March and April 2021. Local transmission of non-B.1.1.7 VOCs may be increasing; this warrants urgent further investigation.

DOI: https://doi.org/10.25561/88876

1. Introduction

Since its emergence in Autumn 2020 in South East England, the SARS-CoV-2 variant of concern (VOC) B.1.1.7 has become the dominant lineage across much of Europe.[1] Characterised by several mutations in the spike protein receptor-binding domain (RBD), epidemiological studies suggest B.1.1.7 is 50-80% more transmissible[2,3] and causes more severe disease[4] than previously circulating lineages. B.1.1.7 rose rapidly, from near 0% to over 50% in under two months, and soon made up >98% of sequenced samples in England. Its rapid spread necessitated a third English national lockdown in January 2021. Subsequent spread in Europe[5] and North America[6] has similarly highlighted the threat this variant poses to continued control of community transmission.

The **Δ**69-70 deletion in B.1.1.7's Spike gene causes PCR tests to return negative results for that gene target,[3] allowing S-gene target failure (SGTF) to act as a proxy for genomic surveillance. Both community-based testing of symptomatic individuals ("Pillar 2"[7]) and a weekly survey of more than 100,000 randomly sampled UK residents conducted by the Office for National Statistics (ONS)[8] have shown trends in SGTF frequency which mirrored the pattern seen in sequenced samples. The frequency of SGTF increased from near 0% in October 2020 to 98.8% in March 2021.

Concurrent to B.1.1.7's emergence, additional VOCs have been identified globally, including B.1.351 (first identified in South Africa[9]) and P.1 (first identified in Brazil[10]). Both have been associated with extensive transmission following emergence, leading to substantial infection and mortality rates even in settings where seroprevalence was high (for example in Manaus, Brazil[11,12]). Epidemiological analysis suggests that, like B.1.1.7, these VOCs are more transmissible than ancestral SARS-CoV-2 lineages.[10,13] Neither have the **Δ**69-70 deletion and so test positive in Spike target PCR tests, but both share the E484K mutation thought to contribute to partial immune escape.[14,15,16]

The UK now has a high level of population immunity to SARS-CoV-2: at the beginning of April 2021, it was estimated that 55% (95% CI: 49%-60%) of the English population were seropositive, either due to prior infection or vaccination.[17] However, such high levels of immunity also represent an evolutionary selection pressure on the virus and may give VOCs with even a partial degree of immune escape (relative to B.1.1.7) a transmission fitness advantage –especially at a time where control measures are being progressively relaxed in the UK. Understanding when, how and if these VOCs pose a threat in settings where B.1.1.7 is currently dominant is vital. It is especially relevant for the UK, where vaccination rollout has relied heavily on the AstraZeneca vaccine; a vaccine that has proven highly protective against B.1.1.7 and prior variants,[18] but may possess reduced efficacy against other VOCs.[16]

Here, we use a combination of data from passive-case detection PCR data, cross-sectional community infection surveys, genomic sequencing surveillance, and wastewater monitoring to examine spatial and temporal trends in B.1.1.7's prevalence in England. We focus on London, which has the clearest trends and most available data, but we observe similar patterns in other regions. Our results suggest dynamic shifts in the composition of SARS-CoV-2 lineages driving transmission across England in March and April 2021, and an expansion of non-B.1.1.7 VOCs that warrants urgent further investigation.

DOI: https://doi.org/10.25561/88876

2. Methods

2.1 Data and Statistical Analysis

2.1.1 Pillar 2 symptomatic community testing

Public Health England's surveillance system assembles data from dozens of PCR testing laboratories, the largest of which are the three large "Lighthouse" laboratories developed specifically in response to the pandemic. Approximately 30% of the samples processed by the Lighthouse laboratories use the ThermoFisher TaqPath PCR assay, which includes Spike as a target. For tests that give a PCR cycle threshold (Ct) value for non-spike targets substantially below the positivity threshold of 40, SGTF is a highly accurate proxy for B.1.1.7. Thus we are able to categorise a substantial proportion of all lab-confirmed community SARS-CoV cases as B.1.1.7 or non-B.1.1.7.[2] SGTF becomes less reliable when Ct values for all targets are high since the Spike target is more likely to test negative by chance when sample viral load is low. Hence we estimate the frequency of SGTF only from cases with Ct values in non-Spike targets of 30 or less.

We consider the period from 31st January 2021 to 15th May 2021. We only consider test results in self-reported symptomatic cases and exclude tests conducted following a lateral flow test (used, for instance, for asymptomatic screening for infection in schools and workplaces). Unlike the COG-UK data detailed below, we do not have metadata to exclude individuals with recent travel history. Over that period and with these exclusions applied, there was a total of 72,881 S-gene positive (S+), and 586,854 S-gene negative (S-) cases in England processed by the Lighthouse laboratories and 4,246 S+ and 79,207 S- cases in London. Given that SGTF results are only available for a subset of samples, we estimate total Spike-positive (S+) case incidence by multiplying the frequency of S+ among all cases with SGTF results by the total Pillar 2 case incidence. Uncertainty estimates are detailed in Supplementary Text.

2.1.2 ONS Infection Survey

ONS conducts a fortnightly survey of randomly selected private households in the UK. In the two weeks prior to 16th April 2021, 139,948 participants from 73,328 households were tested using nose and throat self-swabs, analyzed with a PCR test. A Bayesian model was used to estimate the positivity rate for SARS-CoV-2 in the community, stratified by regions of England.[19] We use the ONS estimates of the percentage of PCR-positive samples that are "not compatible with UK variant" (gene pattern S + ORF1ab + N; indicated as S+ in Figure 1) and the estimates of samples that are "UK variant compatible" (gene pattern ORF1ab + N indicating likely infection with B.1.1.7). Uncertainty estimates are detailed in Supplementary Text.

Each ONS release provides estimates for a 6 week period. We combine all the ONS releases from 26th February 2021 to 14th May 2021. For duplicated dates, we take the most recent estimate available in the combined data. To estimate total infection prevalence for each region (Figure 1A and Supp Figure A), we multiply the estimated S+ infection prevalence for that region by its population size as reported by ONS.[20]

DOI: https://doi.org/10.25561/88876

2.1.3 Sewage water monitoring

Sequencing of viral RNA from sewage water has been a valuable tool for tracking the distribution of SARS-CoV-2 variants in the UK, both during the first wave[21] and the rise of B.1.1.7.[22] In particular, a key advantage of this method is low sampling bias as it captures all people in the catchment area and not only those that receive COVID-19 tests. Here, we analysed fortnightly samples from the Beckton Sewage Treatment Works plant, which has a catchment area containing approximately 4 million people in North London. The catchment area does not include Heathrow Airport and adjacent quarantine hotels, which drain into the Mogden Sewage Treatment Works plant (as confirmed by Thames Water). Sample collection, processing, and analysis are described in detail in previous work;[21,22] a short summary is given in Supplementary Text.

2.1.4 COG-UK Genomic Sequencing

We studied 5,277 sequences collected from Pillar 2 testing in the greater London area after March 1, 2021 and provided by the COG-UK consortium.[23] Sequence quality control, alignment, and lineage classification was carried out as described in previous work[24] and computed with the MRC-CLIMB computational infrastructure.[25] Lineage classification for novel variants under investigation B.1.617.1 and B.1.617.2 were checked manually using the pangolin tool.[26] Among the 5,277 sequences, 461 were found to be from a lineage other than B.1.1.7 with 336 sequences in the set of VOCs and variants under investigation (VUIs) P.1 (n=21), B.1.1.318 (n=27), B.1.525 (n=69), B.1.617.2 (n=27), B.1.617.1 (n=52) and B.1.351 (n=140).

We estimated the frequency over time for each lineage with more than 20 samples using a Gaussian process generalized additive model with a multinomial response for each lineage. A large majority of the non-B.1.1.7 sequences (n=344) were found to be collected from managed quarantine facilities and individuals with recent travel history or surge testing. We repeated the analysis excluding this set.

DOI: https://doi.org/10.25561/88876

3. Results

Since the beginning of March 2021, S+ infection prevalence (ONS) and S+ case incidence (Pillar 2) have both started to increase against a background of falling overall case numbers. Figure 1 displays the data for London, where this trend is clearest, but there are signs of similar patterns in nearly every other region in England (Supplementary Figure 1). However, Pillar 2 is based on non-random testing and the ONS survey may suffer from sampling artifacts due to the low overall incidence in London in recent weeks.

Examination of the Pillar 2 Ct values supports a qualitative change in S+ transmission patterns. Ct values in community testing are both inversely related to viral load and associated with transmission levels[27]–declining epidemics are correlated with lower mean Ct values, and vice-versa. Figure 2 shows that until March 2021, S- samples (primarily B.1.1.7) had statistically significantly lower Ct values than S+ samples, especially for the N gene. This is as expected; reports suggest B.1.1.7 has higher viral loads, and thus lower Ct values, than prior lineages.[28] Since the end of March 2021, however, mean Ct values for S+ samples have significantly decreased and are now comparable to values for S- samples. This suggests either a change in the genetic composition of S+ cases, with variants causing higher viral loads becoming more dominant, and/or an increase in transmission of S+ lineages.

Figure 3 shows the frequency of mutations in SARS-CoV-2 viral RNA found in sewage water[21,22] from North London. This data source includes all people living in the sewage plant's catchment area, not just those that are tested. Figure 3 confirms that the increase in the proportion of S+ observed in other data sources is due to a decrease in the proportion of B.1.1.7. Mutations HV69-70del, Y144del, and A570D are relatively unique to B.1.1.7 (Supplementary Table 1).[22] All three mutations were detected at a stable frequency >95% from early January[22] to mid-March 2021 and then decreased to mean frequencies of 67% - 75% by April 13th (Figure 3). Conversely, the frequency of the E484K mutation – absent in B.1.1.7 but present in many other variants of concern/interest – has increased to over 30% by April 13th. Analysis of independent subsamples further reveals that E484K is indeed only present in non-B.1.1.7 viruses (Supplementary Text). These data suggest that variants with E484K are replacing B.1.1.7 in North London. This non-B.1.1.7 population can be further differentiated by analysing additional mutations (Supplementary Text), albeit with considerable uncertainty due to low viral loads. The non-B.1.1.7 population likely contains B.1.351 and B.1.525, while P.1 and B.1.617.1 were not found.

Figure 4 shows results from COG-UK sequencing of SARS-CoV-2 samples from London since March 1st 2021, also indicating recent growth in non-B.1.1.7 lineages. This trend is largely driven by increases in non-B.1.1.7 infections from travel-linked cases and surge testing (Figure 4 A-B). Smaller increases are observed in a subset which excludes such cases (Figure 4 C-D). The rise of non-B.1.1.7 lineages after April 1 is largely driven by imports of the B.1.617 lineages associated with the current epidemic wave in India.[29]

DOI: https://doi.org/10.25561/88876

4. Discussion

Experiences across the globe to date have highlighted the significant public health threat that new SARS-CoV-2 VOCs can pose, even in settings where transmission is currently under control or where population-level immunity should preclude resurgence. They have also highlighted the importance of early detection and identification of emerging viral threats, which provides the opportunity for prompt implementation of measures to control spread. Here, using four independent data sources, we present evidence supporting recent increases in the proportion of COVID-19 infections that are S+; an increase possibly driven by B.1.351, B.1.525, and B.1.617.

A key question is whether these trends reflect local transmission of those variants, or imported infections detected on the background of very low overall incidence (Pillar 2 incidence was below 0.5 cases/1000/week in London at the end of April 2020). In a context of high and sometimes rising incidence in many origin countries for international travellers and low and declining incidence in the UK, importations would be expected to represent an increasing proportion of detected cases, and this alone might explain the observed increase over time in S+ lineage frequency.

While frequency of non-B.1.1.7 lineages has trended upwards since mid-March, genomic sequencing data suggest a majority of these cases may be linked directly or indirectly to overseas travel. While >20% of sequenced cases were from non-B.1.1.7 lineages as of mid-April (Figure 4B), the fraction is smaller in cases not known to be associated with travel or surge testing (Figure 4D). An upward trend in non-B.1.1.7 lineages could suggest that local transmission of such lineages is occurring, consistent with detected clusters in London and elsewhere[9,30,31], though we do not yet know the extent to which this transmission is self-sustaining or is associated with typically short chains of transmission initiated by individual importation events. In addition, VOCs such as B.1.351 are subject to enhanced public health interventions, and thus the patterns we observe may deviate substantially from what would be observed otherwise.

However, further lines of evidence suggest that local transmission of non-B.1.1.7 VOCs may be increasing. The recent uptick in E484K[14,15,16] frequency in wastewater sequencing in North London is a particular concern, given the large catchment of this data stream and that it is not subject to the same surveillance biases as symptomatic case testing. Less directly, the observation of recent decreases in the average Ct values for S+ cases also provides support for ongoing community transmission. Recent work has shown that population-level average Ct values can provide an indication about the epidemic's dynamics, with average Ct values declining when epidemics are growing and increasing when epidemics are declining.[27] Trends in mean Ct values could be consistent with a change in the transmission patterns of S+ lineages. However, multiple VOC/non-VOCs are nested within the S+ classification; it is therefore not possible to disentangle the comparative contributions of each lineage with confidence. As shown in Supplementary Figures 1 and 2, trends similar to those we have described in London may be occurring in other regions of England, though overall S+ cases are so far at lower levels. Last, the detection of several clusters of VOC B.1.351[29–31] in London and the rest of England also suggest community transmission.

We note that it is not inevitable that E484K/Q-carrying variants will outcompete B.1.1.7. Variants under investigation such as B.1.525 and A.23.1 have undergone periods of rapid expansion in January-

March 2021 associated with travel-related importation and limited local spread, only to subside in the most recent period. The outcome of competition between two variants depends on their relative transmission fitness, which is determined by the intrinsic transmissibility of each strain, the extent each can evade prior immunity and any targeted non-pharmaceutical interventions in place.

Several studies suggest that B.1.1.7,[2,3] P.1,[10] and B.1.351[13] are more transmissible than previously circulating lineages, but precise estimates of their relative transmissibility are not yet available. However, even if B.1.351 and P.1 are less intrinsically transmissible than B.1.1.7, any substantive ability to evade prior immunity may give those VOCs an overall transmission advantage over B.1.1.7 in the context of a highly immunised population such as the UK's. Mounting evidence from *in vitro,*[14,32] epidemiological,[10,13] and vaccine studies[15,16,33,34] suggests that variants with E484K or E484Q mutations may partially evade prior immunity – indeed, rapid resurgences followed variant emergence in both Manaus, Brazil (where P.1 was first identified) despite potential evidence of high levels of immunity in the population.[11,35] The extent of evasion against vaccine-based and natural immunity remains uncertain, though trials and observational studies suggest reduced efficacy of a number of vaccines against B.1.351[15,16,36]. There have been suggestions however that residual protection against severe disease may be higher[37].

Events following the emergence of novel SARS-CoV-2 variants have emphasised the value of identifying and responding to changes in lineage frequency early. Overall, our analysis provides a still ambiguous but potentially concerning early signal of current transmission of non-B.1.1.7 VOCs in England which suggest a need for intensified monitoring. Rapid increases in such VOCs may threaten the success to date of the UK vaccination programme. More generally, our results underscore the value of utilising a diverse array of data sources in community surveillance and underscore the value of timely genomic surveillance to provide real-time information on the highly dynamic composition and trajectory of different SARS-CoV-2 lineages in the country. Such information is critical to the epidemic's immediate control and to future vaccine development and deployment - both in the UK and other countries where the potential emergence of other novel SARS-CoV-2 variants remains a serious public health threat.

5. Data Availability

Data underlying the figures, source code, and links to publicly available data sources can be found at https://github.com/ImperialCollegeLondon/SARS_CoV_2_variants_uk.

DOI: https://doi.org/10.25561/88876

6. Figures

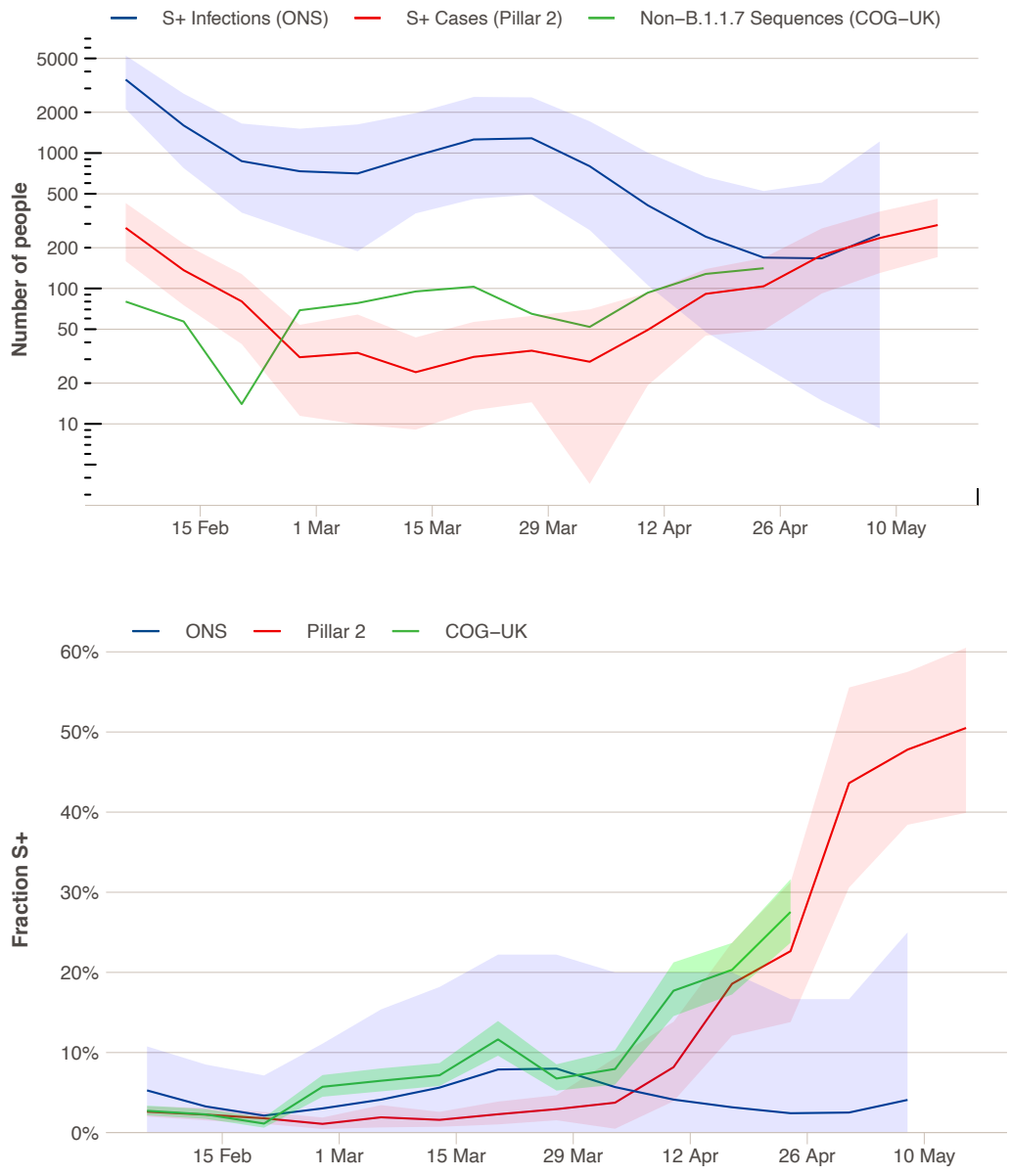


Figure 1: Trends in S+ infections in London, February-May 2021. A) Estimated aggregated weekly incidence (log scale) of symptomatic S+ cases diagnosed via community testing (Pillar 2), S+ infections estimated from the ONS infection survey[38], and non-B.1.1.7 SARS-CoV-2 sequences (COG-UK public data, which may include travelers and surge testing. **B)** Temporal trends in the proportion of cases and infections that are S+, estimated from symptomatic community testing (Pillar 2), the ONS infection survey, and from SARS-CoV-2 sequence data (non-B.1.17 fraction is shown). Results for other regions of England can be found in Supplementary Figures 1 and 2. Details on uncertainty intervals can be found in Supplementary Text.

DOI: https://doi.org/10.25561/88876

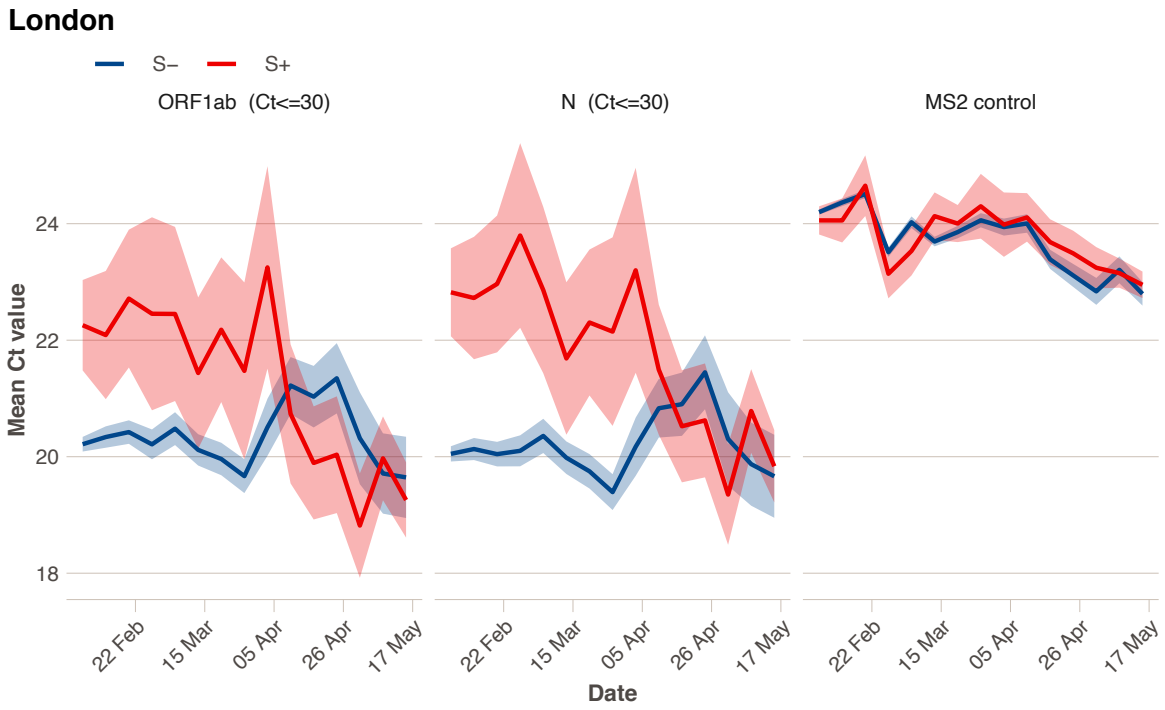


Figure 2: Mean Cycle threshold (Ct) values by week for Pillar 2 symptomatic community testing in London. Shaded ribbons show 95% confidence intervals for the mean. Ct values for ORF1ab gene and N gene are shown, with S+ in blue and S- in red. MS2 control indicates the mean Ct value of Bacteriophage MS2, which is added to samples for calibration purposes. Results for other regions of England can be found in Supplementary Figure 3.

DOI: https://doi.org/10.25561/88876

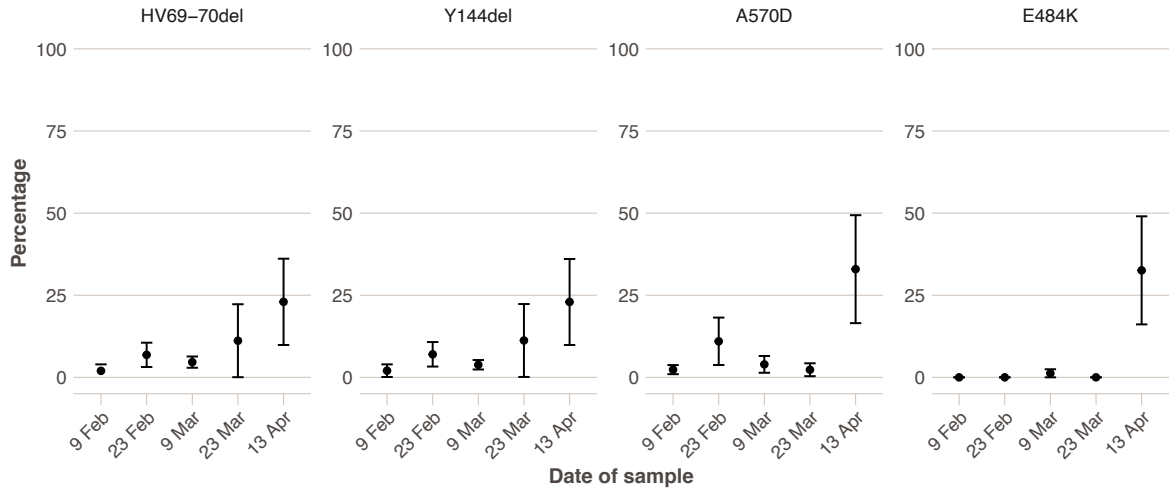


Figure 3: Fraction of viral RNA showing mutations at key spike protein amino acid positions, identified in sewage samples from North London. Mean values from replicate sequences (n=8-12) for each sampling date are shown. Error bars indicate standard error of the mean. HV69-70del, Y144del, and A570D are relatively uniquely found in B.1.1.7 (Supplementary Table 1). E484K is absent in B.1.1.7. but present in several other variants of interest/concern.

DOI: https://doi.org/10.25561/88876

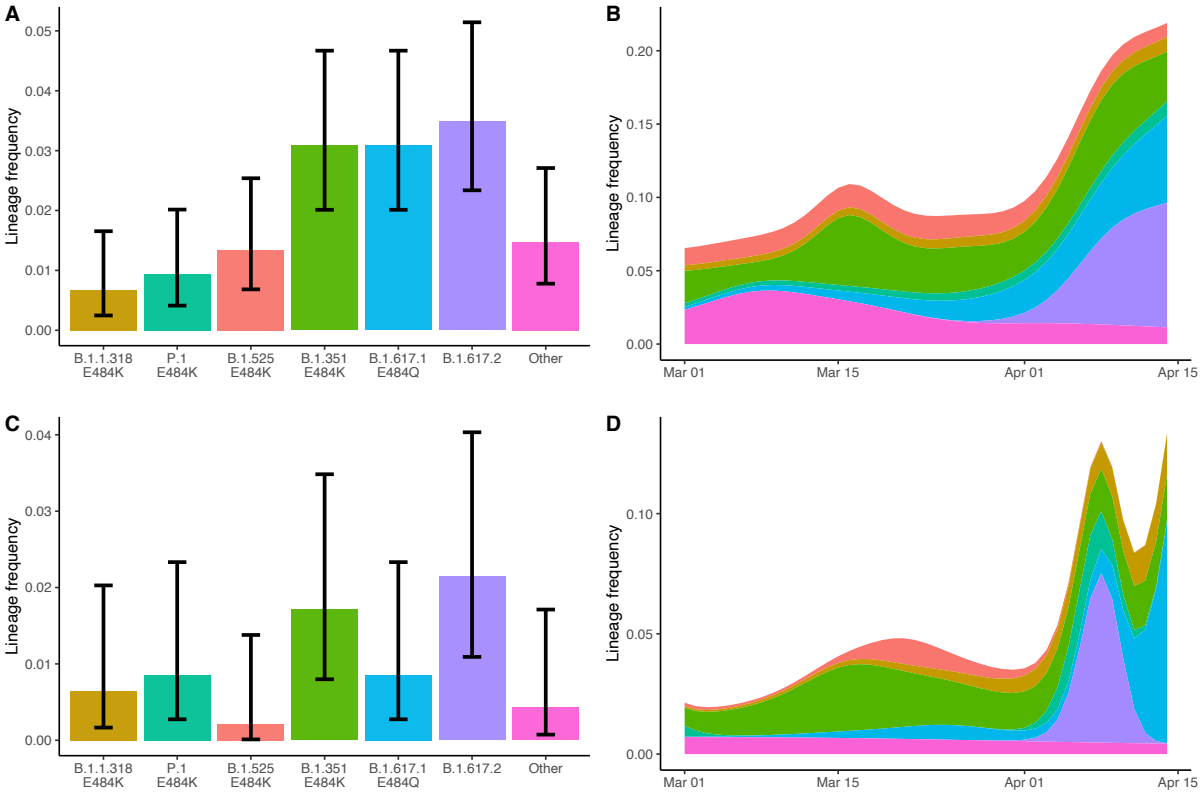


Figure 4. The sample frequency of non-B.1.1.7 lineages in Greater London in community testing. Bar charts show the sample proportion of lineages with at least 20 samples after 31 March 2021. Error bars show 95% confidence intervals based on binomial sampling. Stacked area charts show estimates over time of the frequency of lineages in the period 1 March to 17 April. Panels A-B show results for all Pillar 2 tests (n=461). Panels C-D show results excluding samples from managed quarantine facilities (travel associated) and surge testing (n=117).

DOI: https://doi.org/10.25561/88876

7. Acknowledgements

We thank Mihaela Cirdei, Ola Miloszewska, and Julian Hand from Public Health England for arranging collection and transport of raw sewage samples. We thank A. Sarah Walker (Oxford) for clarifying discussions about the analysis of ONS data. We thank Elizaveta Semenova for comments on a draft of the manuscript.

8. Funding and relations

Sewage water analysis is based on independent research commissioned and funded by the National Institute for Health Research (NIHR) Policy Research Programme (NIBSC Regulatory Science Research Unit) and Medicines and Healthcare products Regulatory Agency (MHRA). S. Mindermann's funding for graduate studies was from Oxford University and DeepMind. M. Sharma was supported by the EPSRC Centre for Doctoral Training in Autonomous Intelligent Machines and Systems (EP/S024050/1) and a grant from the EA Funds programme. J Ahuja was supported by Open Philanthropy. C. Rogers-Smith was supported by Open Philanthropy. G. Leech was funded by UKRI grant EP/S022937/1. G. Nason is a member of the Royal Statistical Society's COVID-19 Taskforce. Y. Gal has received a research grant (studentship) from GlaxoSmithKline, outside of the submitted work. S. Bhatt acknowledges the Academy of Medical Sciences Springboard Award (SBF004/1080), The BMGF (OPP1197730), Imperial College Healthcare NHS Trust- BRC Funding (RDA02) and The Novo Nordisk Young Investigator Award (NNF20OC0059309). S. Mishra, S. Bhatt, C. Whittaker, TA Mellan, H.J.T. Unwin, E. Volz, and N.M. Ferguson acknowledge funding from the MRC Centre for Global Infectious Disease Analysis (reference MR/R015600/1) and Community Jameel. S. Bhatt and N.M. Ferguson acknowledge funding from the National Institute for Health Research (NIHR) Health Protection Research Unit in Modelling and Health Economics (grant code NIHR200908) and UK Research and Innovation (MR/V038109/1). Outside of the submitted work, N.M. Ferguson has received grants from NIH NIGMS, Janssen Pharmaceuticals, Bill and Melinda Foundation, Gavi, the vaccine alliance. O. Ratmann acknowledges The BMGF (OPP1175094). J.M. Brauner was supported by the EPSRC Centre for Doctoral Training in Autonomous Intelligent Machines and Systems (EP/S024050/1) and by Cancer Research UK. A. Howes was supported by the EPSRC Centre for Doctoral Training in Modern Statistics and Statistical Machine Learning (EP/S023151/1). S. Flaxman acknowledges the EPSRC (EP/V002910/1) and the Imperial College COVID-19 Research Fund. COG-UK is supported by funding from the Medical Research Council (MRC) part of UK Research & Innovation (UKRI), the National Institute of Health Research (NIHR) and Genome Research Limited, operating as the Wellcome Sanger Institute.

The views expressed in the publication are those of the authors and not necessarily those of the funders, NHS, the NIHR, the Department of Health, arm's length bodies, or other government departments.

9. Competing interests

All authors declare no conflicts of interests or competing interests.

10. Contributions

All authors fulfilled these criteria:

•Substantial contributions to the conception or design of the work; or the acquisition, analysis, or interpretation of data for the work; AND

•Drafting the work or revising it critically for important intellectual content; AND

•Final approval of the version to be published; AND

•Agreement to be accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.

Additionally, authors contributing to the formal analysis and design were S. Mishra, S. Flaxman, S. Bhatt, S. Mindermann, M. Sharma, J.M. Brauner, C. Whittaker, T.A. Mellan, E. Volz, J. Martin, N.M. Ferguson.

S. Mishra, S. Flaxman, S. Bhatt, S. Mindermann, M. Sharma, J.M. Brauner, C. Whittaker, N.M. Ferguson led the investigation and conceptualisation of the idea.

T. Wilton, D. Klapsa, R. Mate, M. Fritzsche, M. Zambon and J. Martin ran experiments and analysis for the sewage samples.

E. Volz and S. Flaxman analysed sequence data from COG-UK.

11. References

1 European Centre for Disease Prevention and Control. SARS-CoV-2-increased circulation of variants of concern and vaccine rollout in the EU/EEA, 14th update. 2021. https://www.brief.com.cy/sites/default/files/2021-03/RRA-covid-19-14th-update-15-feb-2021.pdf.

2 Volz E, Mishra S, Chand M, et al. Assessing transmissibility of SARS-CoV-2 lineage B.1.1.7 in England. Nature 2021; published online March 25. DOI:10.1038/s41586-021-03470-x.

3 Davies NG, Abbott S, Barnard RC, et al. Estimated transmissibility and impact of SARS-CoV-2 lineage B.1.1.7 in England. Science 2021; 372. DOI:10.1126/science.abg3055.

4 Davies NG, Jarvis CI, CMMID COVID-19 Working Group, et al. Increased mortality in community-tested cases of SARS-CoV-2 lineage B.1.1.7. Nature 2021; published online March 15. DOI:10.1038/s41586-021-03426-1.

5 Gaymard A, Bosetti P, Feri A, et al. Early assessment of diffusion and possible expansion of SARS-CoV-2 Lineage 20I/501Y.V1 (B.1.1.7, variant of concern 202012/01) in France, January to March 2021. Euro Surveill 2021; 26. DOI:10.2807/1560-7917.ES.2021.26.9.2100133.

6 Washington NL, Gangavarapu K, Zeller M, et al. Emergence and rapid transmission of SARS-CoV-2 B.1.1.7 in the United States. Cell 2021; published online March 30. DOI:10.1016/j.cell.2021.03.052.

7 Department of Health and Social Care. COVID-19 testing data: methodology note. https://www.gov.uk/government/publications/coronavirus-covid-19-testing-data-methodology/covid-19-testing-data-methodology-note (accessed April 22, 2021).

8 Office for National Statistics. Coronavirus (COVID-19) infection survey, UK - office for national statistics. 2021; published online April 22. https://www.ons.gov.uk/peoplepopulationandcommunity/healthandsocialcare/conditionsanddiseases/bulletins/coronaviruscovid19infectionsurveypilot/23april2021 (accessed April 27, 2021).

9 Tegally H, Wilkinson E, Giovanetti M, et al. Emergence and rapid spread of a new severe acute respiratory syndrome-related coronavirus 2 (SARS-CoV-2) lineage with multiple spike mutations in South Africa. bioRxiv. 2020; published online Dec 22. DOI:10.1101/2020.12.21.20248640.

10 Faria NR, Mellan TA, Whittaker C, et al. Genomics and epidemiology of the P.1 SARS-CoV-2 lineage in Manaus, Brazil. Science 2021; published online April 14. DOI:10.1126/science.abh2644.

11 Buss LF, Prete CA Jr, Abrahim CMM, et al. Three-quarters attack rate of SARS-CoV-2 in the Brazilian Amazon during a largely unmitigated epidemic. Science 2021; 371: 288–92.

12 Sabino EC, Buss LF, Carvalho MPS, et al. Resurgence of COVID-19 in Manaus, Brazil, despite high seroprevalence. Lancet 2021; 397: 452–5.

13 Carl A. B. Pearson, Timothy W Russell, Nicholas Davies, Adam J Kucharski, CMMID COVID-19 working group, W John Edmunds & Rosalind M Eggo. Estimates of severity and transmissibility of novel South Africa SARS-CoV-2 variant 501Y.V2. LSHTM CMMID COVID-19 Repository 2021; published online Jan 11. https://cmmid.github.io/topics/covid19/sa-novel-variant.html (accessed Jan 19, 2021).

14 Garcia-Beltran WF, Lam EC, St Denis K, et al. Multiple SARS-CoV-2 variants escape neutralization by vaccine-induced humoral immunity. Cell 2021; published online March 12. DOI:10.1016/j.cell.2021.03.013.

15 Madhi SA, Baillie V, Cutland CL, et al. Efficacy of the ChAdOx1 nCoV-19 Covid-19 Vaccine against the B.1.351 Variant. N Engl J Med 2021; published online March 16. DOI:10.1056/NEJMoa2102214.

16 Shinde V, Bhikha S, Hoosain Z, et al. Preliminary efficacy of the NVX-CoV2373 Covid-19 vaccine against the B.1.351 variant. 2021; published online March 3. DOI:10.1101/2021.02.25.21252477.

17 Office for National Statistics. Coronavirus (COVID-19) Infection Survey, antibody and vaccination data for the UK - Office for National Statistics. 2021; published online April 13. https://www.ons.gov.uk/peoplepopulationandcommunity/healthandsocialcare/conditionsa

nddiseases/articles/coronaviruscovid19infectionsurveyantibodydatafortheuk/14april2021 (accessed April 27, 2021).

18 Vasileiou E, Simpson CR, Shi T, et al. Interim findings from first-dose mass COVID-19 vaccination roll-out and COVID-19 hospital admissions in Scotland: a national prospective cohort study. Lancet 2021; published online April 23. DOI:10.1016/S0140-6736(21)00677-2.

19 Pouwels KB, House T, Pritchard E, et al. Community prevalence of SARS-CoV-2 in England from April to November, 2020: results from the ONS Coronavirus Infection Survey. Lancet Public Health 2021; 6: e30–8.

20 Office for national statistics. Estimates of the population for the UK, England and wales, Scotland and northern Ireland - office for national statistics. https://www.ons.gov.uk/peoplepopulationandcommunity/populationandmigration/populationestimates/datasets/populationestimatesforukenglandandwalesscotlandandnorthernireland (accessed April 29, 2021).

21 Martin J, Klapsa D, Wilton T, et al. Tracking SARS-CoV-2 in Sewage: Evidence of Changes in Virus Variant Predominance during COVID-19 Pandemic. Viruses 2020; 12. DOI:10.3390/v12101144.

22 Wilton T, Bujaki E, Klapsa D, Fritzsche M, Mate R, Martin J. Rapid increase of SARS-CoV-2 variant B.1.1.7 detected in sewage samples from England between October 2020 and January 2021. medRxiv. 2021; published online March 7. DOI:10.1101/2021.03.03.21252867.

23 COVID-19 Genomics UK (COG-UK) consortiumcontact@cogconsortium.uk. An integrated national scale SARS-CoV-2 genomic surveillance network. Lancet Microbe 2020; 1: e99–100.

24 Volz E, Hill V, McCrone JT, et al. Evaluating the Effects of SARS-CoV-2 Spike Mutation D614G on Transmissibility and Pathogenicity. Cell 2021; 184: 64–75.e11.

25 Connor TR, Loman NJ, Thompson S, et al. CLIMB (the Cloud Infrastructure for Microbial Bioinformatics): an online resource for the medical microbiology community. Microb Genom 2016; 2: e000086.

26 Rambaut A, Holmes EC, O'Toole Á, et al. A dynamic nomenclature proposal for SARS-CoV-2 lineages to assist genomic epidemiology. Nat Microbiol 2020; 5: 1403–7.

27 Hay JA, Kennedy-Shaffer L, Kanjilal S, Lipsitch M, Mina MJ. Estimating epidemiologic dynamics from single cross-sectional viral load distributions. bioRxiv. 2020; published online Oct 13. DOI:10.1101/2020.10.08.20204222.

28 Kidd M, Richter A, Best A, et al. S-variant SARS-CoV-2 lineage B1.1.7 is associated with significantly higher viral loads in samples tested by ThermoFisher TaqPath RT-qPCR. J Infect Dis 2021; published online Feb 13. DOI:10.1093/infdis/jiab082.

29 Public Health England. SARS-CoV-2 variants of concern and variants under investigation in England: Technical briefing 9. 2021; published online April. https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/979818/Variants_of_Concern_VOC_Technical_Briefing_9_England.pdf.

30 Public Health England. SARS-CoV-2 variants of concern and variants under investigation in England. https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/975742/Variants_of_Concern_VOC_Technical_Briefing_8_England.pdf.

31 Department of Health and Social Care. Surge testing to be deployed in Wandsworth and Lambeth. GOV. UK. 2021; published online April. https://www.gov.uk/government/news/surge-testing-to-be-deployed-in-wandsworth-and-lambeth.

32 Zhou D, Dejnirattisai W, Supasa P, et al. Evidence of escape of SARS-CoV-2 variant B.1.351 from natural and vaccine-induced sera. Cell 2021; published online Feb 23. DOI:10.1016/j.cell.2021.02.037.

33 Emary KRW, Golubchik T, Aley PK, et al. Efficacy of ChAdOx1 nCoV-19 (AZD1222) vaccine against SARS-CoV-2 variant of concern 202012/01 (B.1.1.7): an exploratory analysis of a randomised controlled trial. Lancet 2021; 397: 1351–62.

34 de Souza WM, Amorim MR, Sesti-Costa R, et al. Levels of SARS-CoV-2 Lineage P.1 Neutralization by Antibodies Elicited after Natural Infection and Vaccination. 2021; published online March 1. DOI:10.2139/ssrn.3793486.

35 Lalwani P, Salgado BB, Pereira Filho IV, et al. SARS-CoV-2 Seroprevalence and Associated Factors in Manaus, Brazil: Baseline Results from the DETECTCoV-19 Cohort Study. 2021; published online Feb 26. DOI:10.2139/ssrn.3795816.

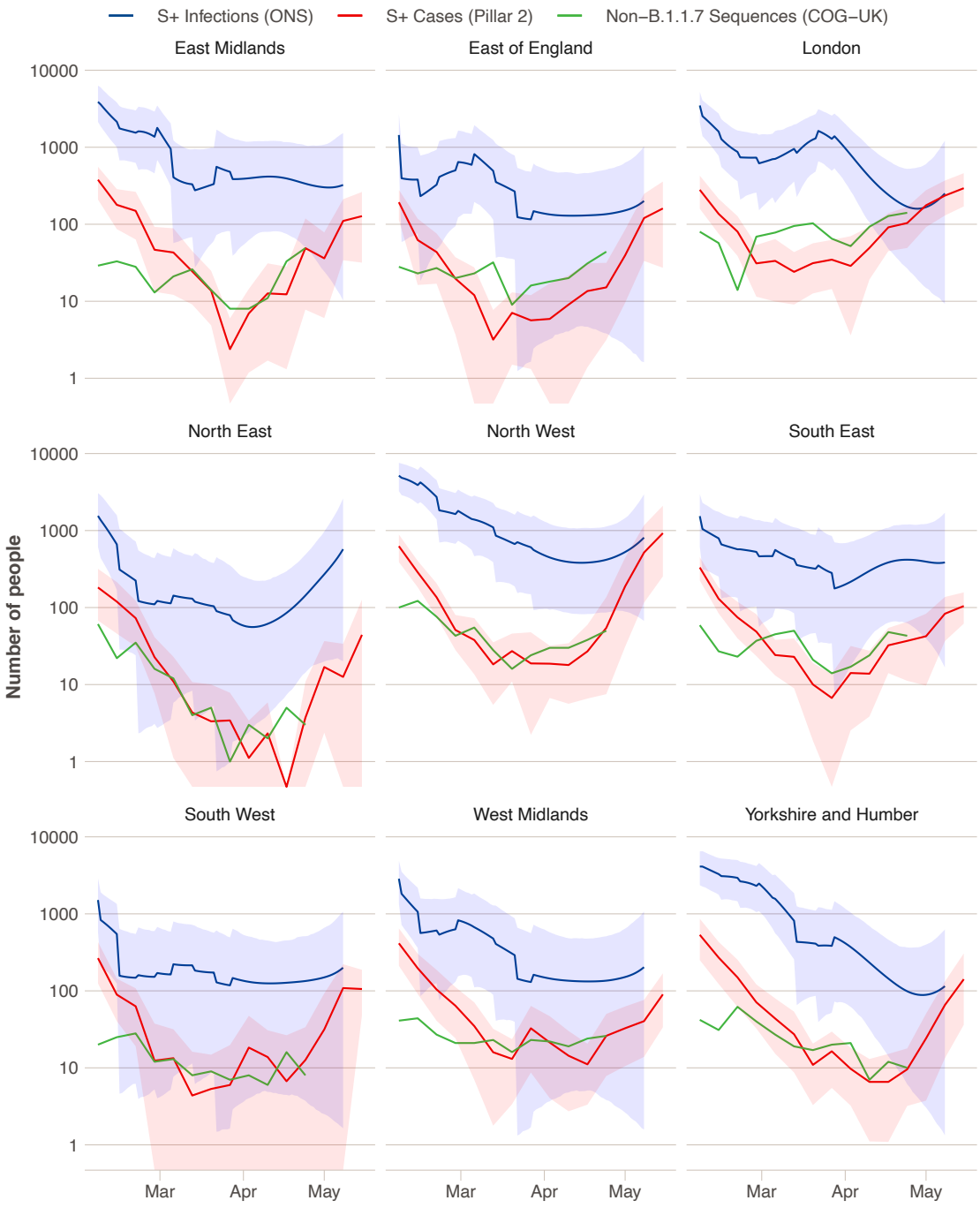
36 Kustin T, Harel N, Finkel U, et al. Evidence for increased breakthrough rates of SARS-CoV-2 variants of concern in BNT162b2 mRNA vaccinated individuals. bioRxiv. 2021; published online April 9. DOI:10.1101/2021.04.06.21254882.

37 Sadoff J, Gray G, Vandebosch A, et al. Safety and Efficacy of Single-Dose Ad26.COV2.S Vaccine against Covid-19. New England Journal of Medicine. 2021. DOI:10.1056/nejmoa2101544.

38 Davies KSA. Coronavirus (COVID-19) Infection Survey, UK - Office for National Statistics. 2021; published online April 15. https://www.ons.gov.uk/peoplepopulationandcommunity/healthandsocialcare/conditionsanddiseases/bulletins/coronaviruscovid19infectionsurveypilot/16april2021 (accessed April 21, 2021).

12. Appendix

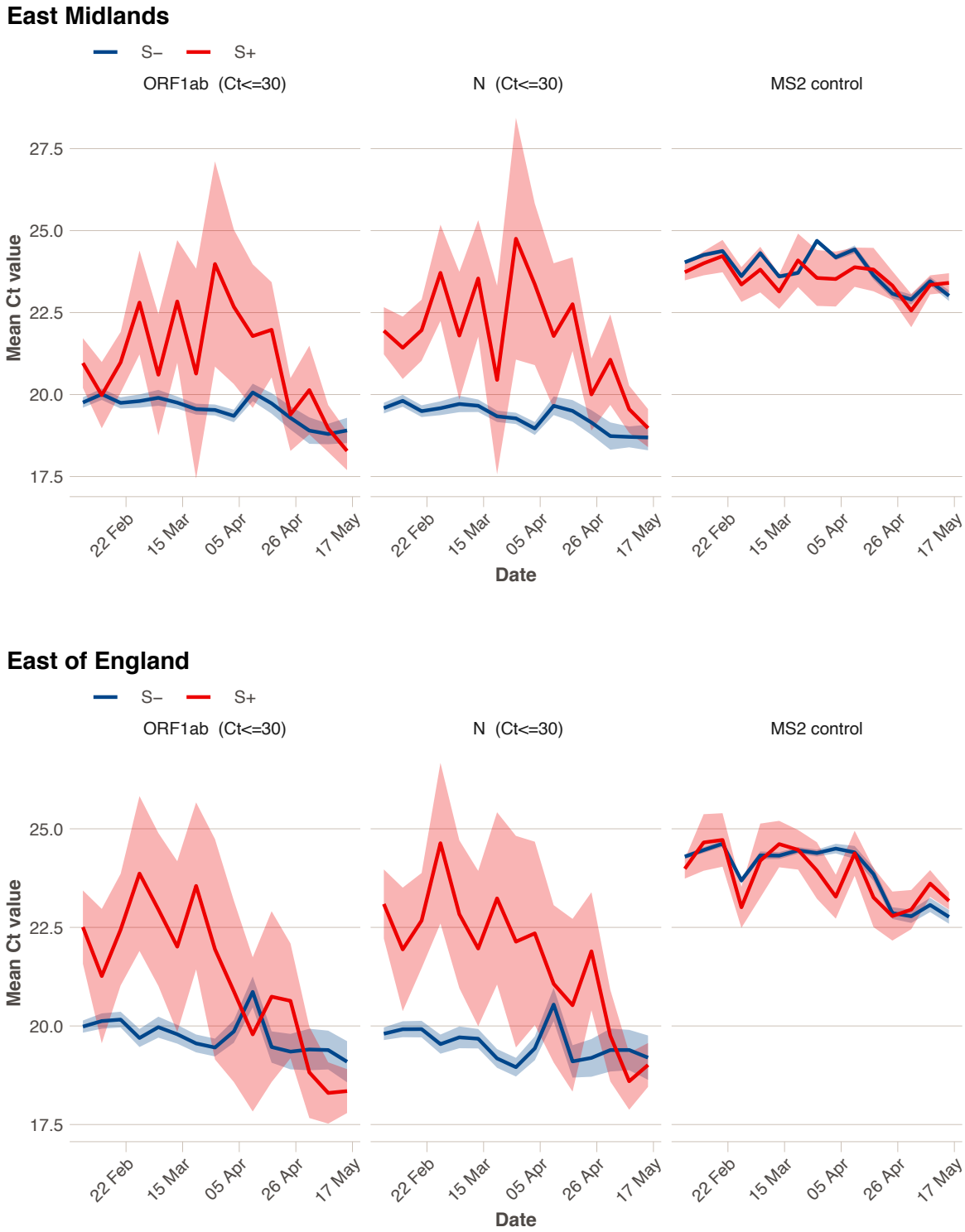
Supplementary Figure 1. Time trend of S-gene positive PCR results and count of non-B.1.1.7 sequences in regions of England in early 2021. Figure 1 has details on data sources.



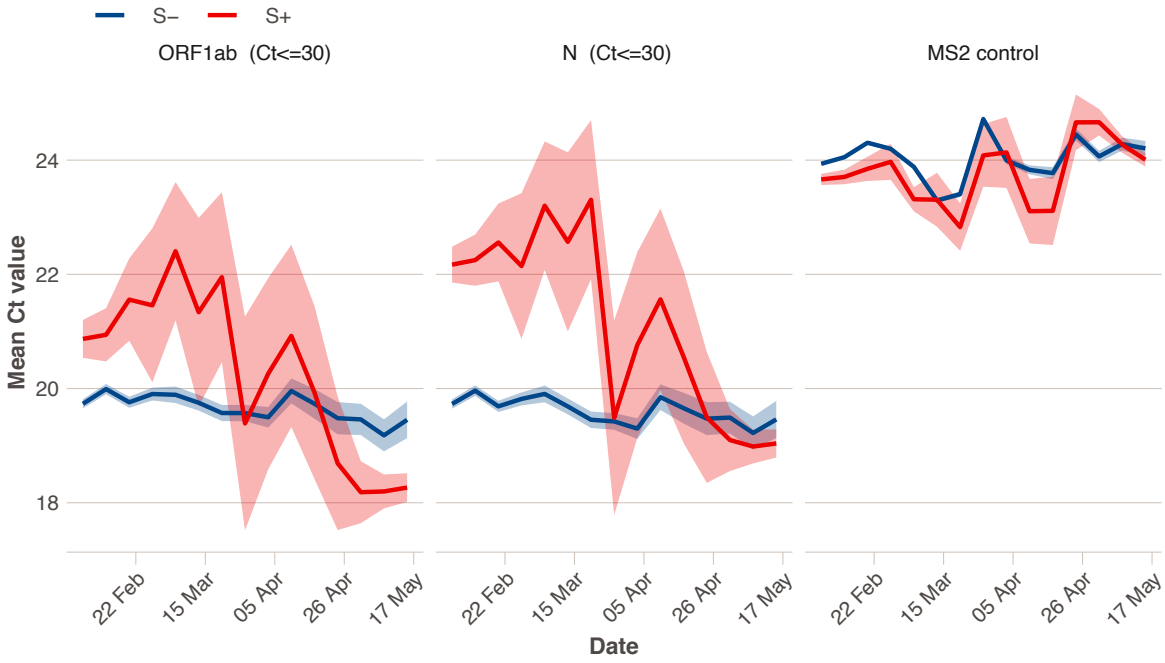
DOI: https://doi.org/10.25561/88876

Supplementary Figure 2. Time trend of fraction of S+ PCR results divided by PCR results which were identified as either S+ or S- in regions of England in early 2021. Figure 1 has details on data sources.

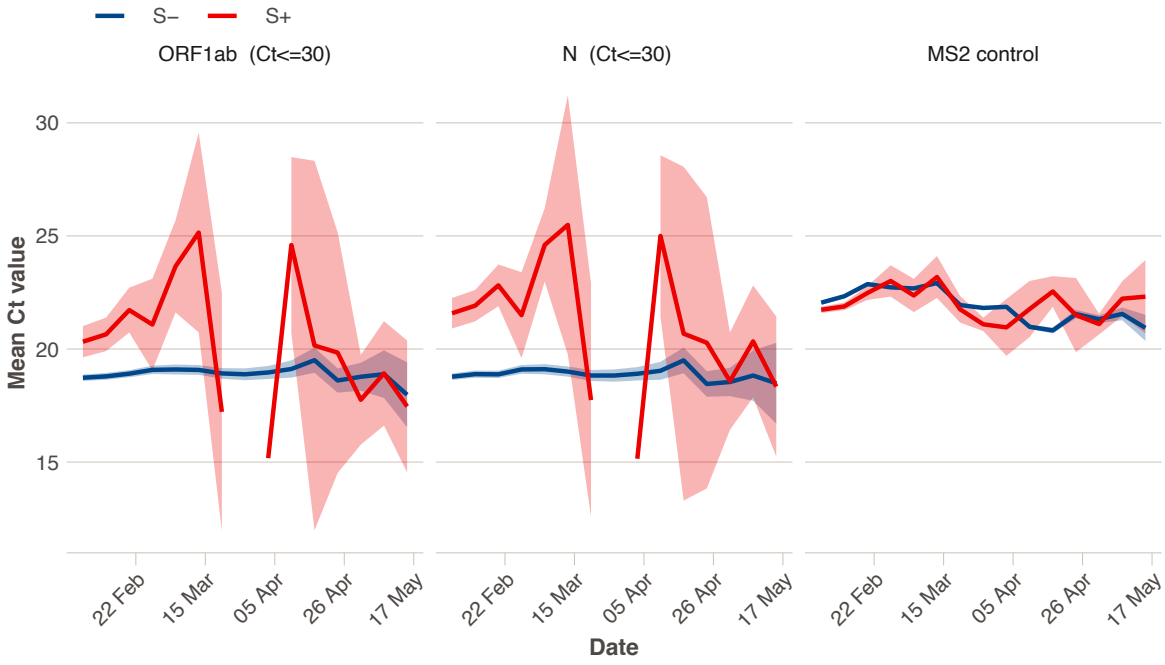
Supplementary Figure 3. Mean Cycle threshold (Ct) values for regions of England. See Figure 2 for data sources and processing.



North West

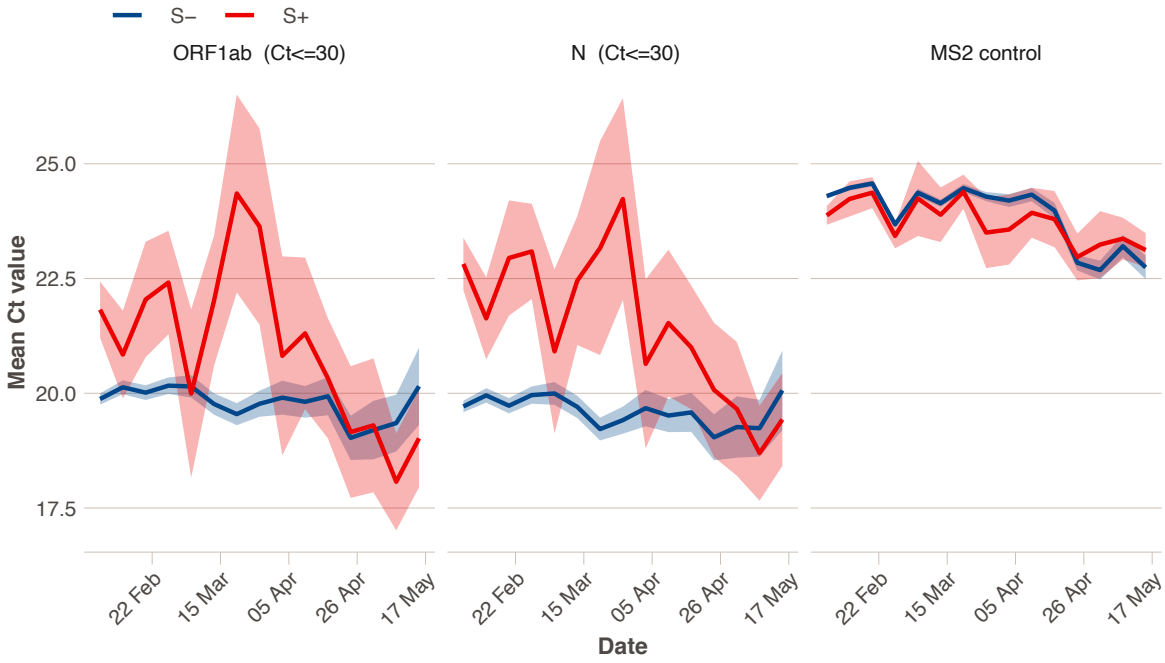


North East

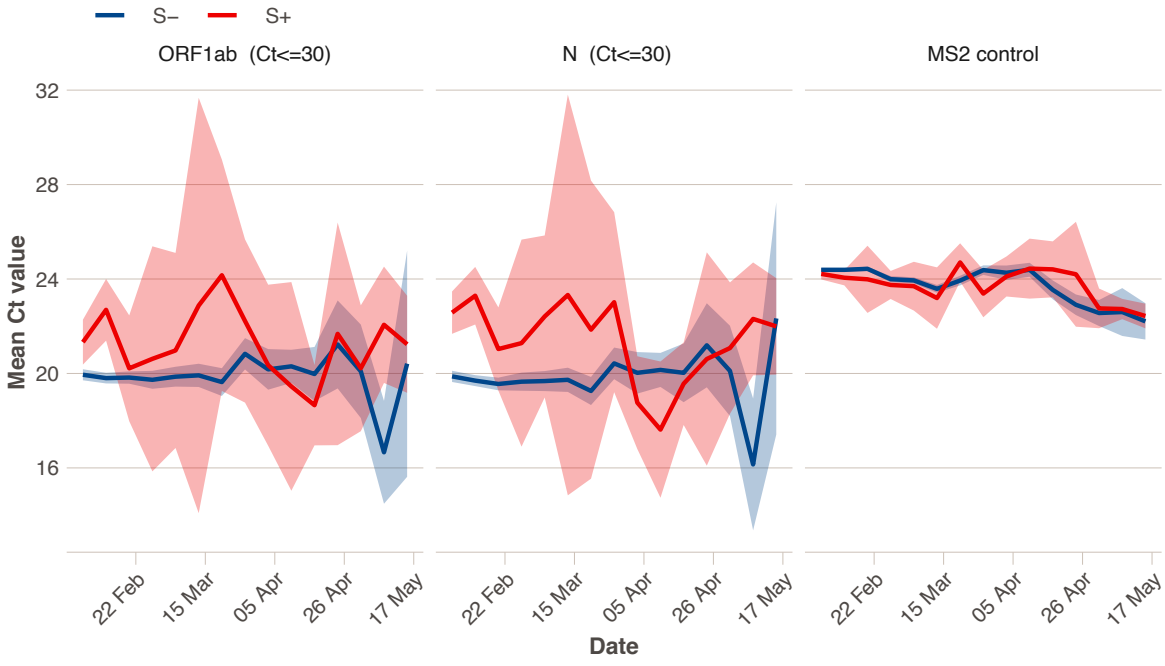


DOI: https://doi.org/10.25561/88876

South East

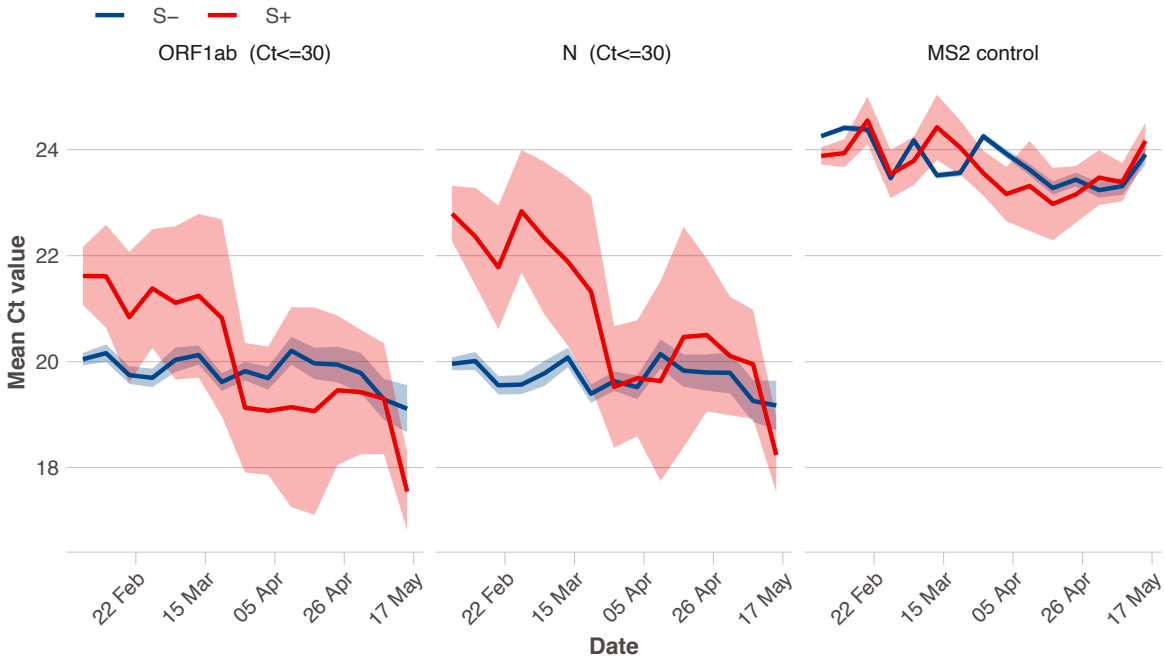


South West

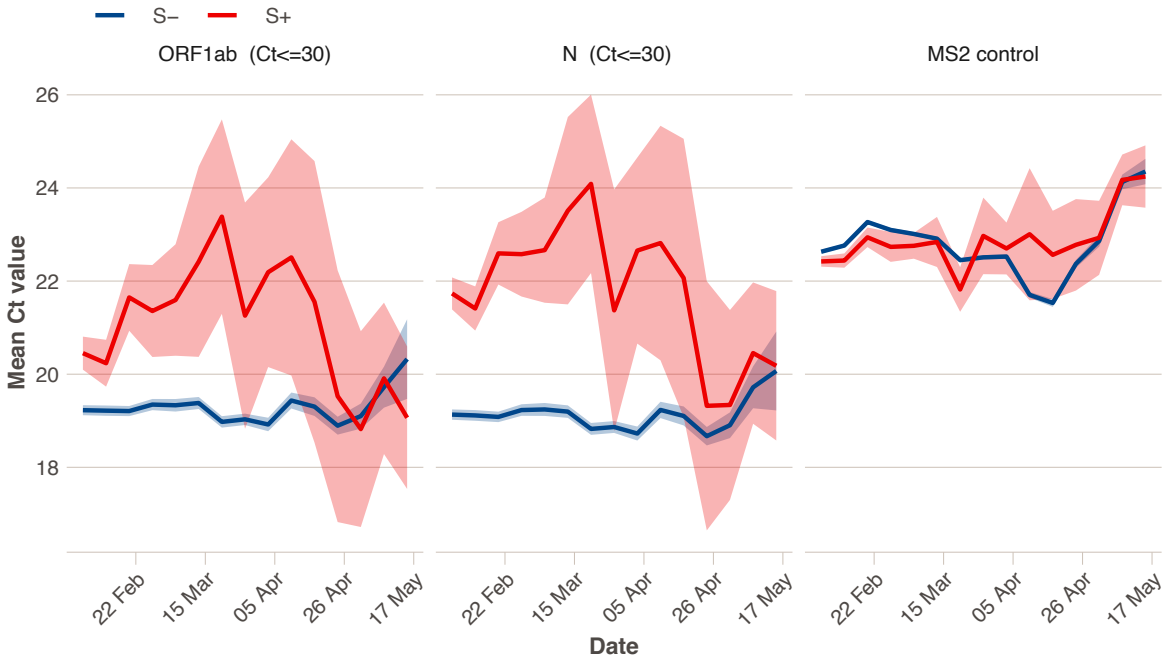


DOI: https://doi.org/10.25561/88876

West Midlands

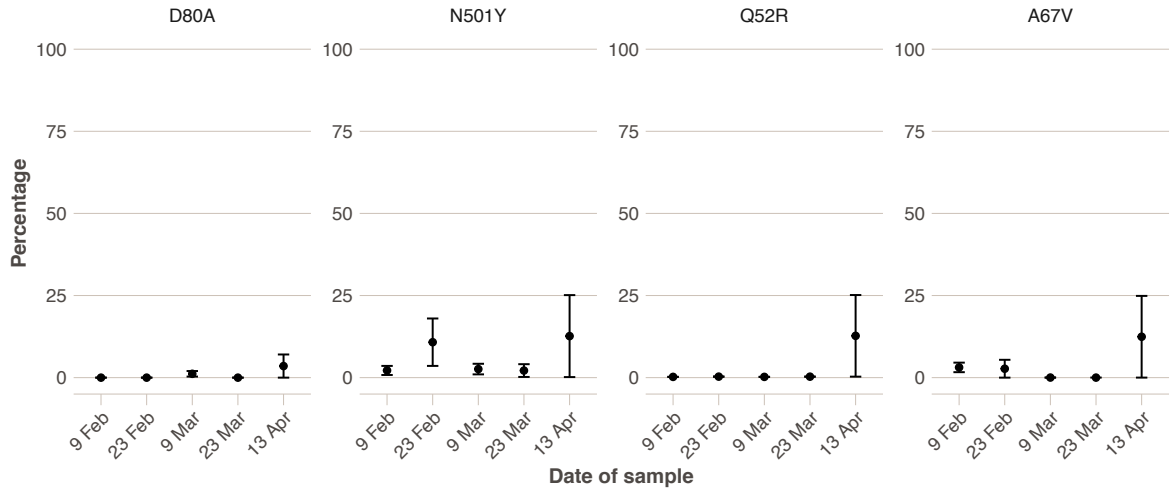


Yorkshire and Humber



DOI: https://doi.org/10.25561/88876

Supplementary Figure 4. Fraction of viral RNA showing mutations at key spike protein amino acid positions, identified in sewage samples from North London. Mean values from replicate sequences (n=8-12) for each sampling date are shown. Error bars indicate standard error of the mean. See Supplementary Text 2 and Supplementary Table 1 for conclusions. D138Y and E484Q, which are characteristic of variants P.1 and B.1.617.1, were not detected.



Supplementary Table 1. Main SARS-CoV-2 variants and their associated mutations. The last two columns describe the trend observed in the fraction of these mutations in North London sewage water (Fig. 3 and Suppl. Fig. 4) and the conclusions that can be drawn from these trends. Sources: Public Health England investigation of novel SARS-CoV-2 variants of concern, Technical Briefings 1, 6, 7, 9.

Mutation	Wuhan-Hu-1 strain (orig. wild type)	B.1.1.7	P.1	B.1.351	B.1.525	B.1.617.1	Observed trend in sewage water	Conclusion from observed trend
HV69-70del	no	yes	no	no	yes	no	Large decrease	Decrease in fraction of B.1.1.7.
Y144del	no	yes	no	no	yes	no	Large decrease	
A570D	no	yes	no	no	no	no	Large decrease	
E484K	no	no	yes	yes	yes	no	Large increase	The variants that replace B.1.1.7 carry E484K.
D80A	no	no	no	yes	no	no	Small increase	Increase in fraction of B.1.351.
N501Y	no	yes	yes	yes	no	no	Moderate decrease	At least one other variant, besides B.1.351, is growing.
Q52R	no	rare	no	no	yes	no	Moderate increase	Increase in fraction of B.1.525.
A67V	no	rate	no	no	yes	no	Moderate increase	
D138Y	no	no	yes	no	no	no	Not found	P.1 not found.
E484Q	no	no	no	no	no	yes	Not found	B.1.617.1 not found.

Supplementary Text: Methods

Sewage water monitoring

Sample collection, processing, and analysis are described in detail in previous work.[21,22] In short, one-litre inlet wastewater composite samples were collected during a 24-hours window on each sampling day. Samples were pre-processed by centrifugation, molecular-weight cut-off filtration, and concentration. Viral RNA was purified from sewage concentrates using the High Pure viral RNA kit (Roche).

RNA aliquots were amplified using a two-step nested reverse-transcription PCR (RT-PCR) process. Because of the length of the spike protein gene, two different primer sets were used, targeting different regions of the gene. The resulting PCR fragments contain the positions of the most relevant mutations (for example, on PCR fragment A: HV69-70del, D80A, D138Y, Y144del; and on PCR fragment B: E484K, N501Y, A570D). Good laboratory practices were ensured in all assays to reduce the possibility of cross-contamination. RNA extraction and negative template controls were included in every assay.

PCR products were analysed with next-generation sequencing to quantify single nucleotide polymorphisms (SNPs) at each nucleotide position. Sequencing was performed with 250 base pair paired-end reads on MiSeq v2 (500 cycles) kits (Illumina). The resulting sequence data were further processed and analysed using Geneious 10.2.3 software. After filtering the reads, paired-end reads were combined and sequence contigs were built by reference-guided assembly. SNPs were identified using Geneious default settings, with the original SARS-CoV-2 Wuhan-Hu-1 strain (GenBank accession number MN908947) as reference.

To reduce sampling effects, we PCR-amplified 12 independent aliquots of RNA concentrate per sampling date, and sequenced all samples with positive PCR results (n=8-12). The results of individual aliquots can reveal which mutations co-occur. For example, positions 484 and 570 map to the same PCR fragment (fragment B) and the detected fractions of E484K and A570D sum to approximately 1 in each of the 8 PCR-positive RNA aliquots. We can thus conclude that the mutation E484K is only present in non-B.1.1.7 viruses.

Further analysis of the non-B.1.1.7 population found in sewage water.

Mutations D138Y and E484Q were not found, indicating the absence of P.1 and B.1.617.1 variants at detectable levels. Mutation D80A, characteristic of B.1.351, however, has increased in frequency to 4% (Suppl. Figure 4). This suggests that a part, but not all, of the non-B.1.1.7 viruses belong to the B.1.351 lineage. Mutation N501Y is present in B.1.1.7 and B.1.351, but not some other variants (Supplementary Table 1). N501Y's frequency decreased to 87% in April, a decrease less pronounced than that of mutations unique to B.1.1.7 (Figure 3). This further implies that some, but not all, of the non-B.1.1.7 viruses belong to the B.1.351 lineage. Mutations Q52R and A67V are unique to B.1.525 and increased to 12% in April, suggesting that B.1.525 might be one of the other lineages contained in the non-B.1.1.7 population. This further matches the observation that the decrease of A570D (present in only B.1.1.7) was more pronounced than that of HV69-70del and Y144del (present in both B.1.1.7 and B.1.525) (Figure 3).

To summarise, sewage water samples suggest that >25% of the North London viral population on 13th April 2021 did not belong to the B.1.1.7 lineage. The non-B.1.1.7 population is likely composed of B.1.351, B.1.525, and possibly other lineages with E484K. P.1 and B.1.617 were not detected.

Statistical analysis of Pillar 2 data.

We use a bootstrap approach to quantify the uncertainty in estimates of Pillar 2 S+ and S- counts. For each week and region, we sample with replacement the counts of S+/S- of local areas as a pair and use these samples to calculate confidence intervals of the regional estimates. The purpose of the bootstrap approach is to account for clustering within local areas and randomness in the amount of PCR tests being sent for SGTF within a local area.

Statistical analysis of ONS Infection Survey data

The ONS Infection Survey reports posterior mean and 95% Bayesian credible intervals (bCI) for the daily positivity rate of infections consistent with B.1.1.7 (positive on ORF1ab and N genes) and separately posterior mean and 95% bCI for the daily positive rate of infections inconsistent with B.1.1.7 (positive on ORF1ab, N, and S genes).

Due to limited information released with the public ONS Infection Survey our ONS confidence intervals should only be seen as pseudo-intervals that give an approximate understanding of the uncertainty around a classical random sampled proportion estimate conditioned on the ONS estimate.

COG-UK full author list

Funding acquisition, Leadership and supervision, Metadata curation, Project administration, Samples and logistics, Sequencing and analysis, Software and analysis tools, and Visualisation:
Samuel C Robson [13].

Funding acquisition, Leadership and supervision, Metadata curation, Project administration, Samples and logistics, Sequencing and analysis, and Software and analysis tools:
Nicholas J Loman [41] and Thomas R Connor [10, 69].

Leadership and supervision, Metadata curation, Project administration, Samples and logistics, Sequencing and analysis, Software and analysis tools, and Visualisation:
Tanya Golubchik [5].

Funding acquisition, Metadata curation, Samples and logistics, Sequencing and analysis, Software and analysis tools, and Visualisation:
Rocio T Martinez Nunez [42].

Funding acquisition, Leadership and supervision, Metadata curation, Project administration, and Samples and logistics:
Catherine Ludden [88].

Funding acquisition, Leadership and supervision, Metadata curation, Samples and logistics, and Sequencing and analysis:
Sally Corden [69].

Funding acquisition, Leadership and supervision, Project administration, Samples and logistics, and Sequencing and analysis:
Ian Johnston [99] and David Bonsall [5].

Funding acquisition, Leadership and supervision, Sequencing and analysis, Software and analysis tools, and Visualisation:
Colin P Smith [87] and Ali R Awan [28].

Funding acquisition, Samples and logistics, Sequencing and analysis, Software and analysis tools, and Visualisation:
Giselda Bucca [87].

Leadership and supervision, Metadata curation, Project administration, Samples and logistics, and Sequencing and analysis:
M. Estee Torok [22, 101].

Leadership and supervision, Metadata curation, Project administration, Samples and logistics, and Visualisation:
Kordo Saeed [81, 110] and Jacqui A Prieto [83, 109].

Leadership and supervision, Metadata curation, Project administration, Sequencing and analysis, and Software and analysis tools:
David K Jackson [99].

Metadata curation, Project administration, Samples and logistics, Sequencing and analysis, and Software and analysis tools:
William L Hamilton [22].

Metadata curation, Project administration, Samples and logistics, Sequencing and analysis, and Visualisation:
Luke B Snell [11].

Funding acquisition, Leadership and supervision, Metadata curation, and Samples and logistics:
Catherine Moore [69].

Funding acquisition, Leadership and supervision, Project administration, and Samples and logistics:
Ewan M Harrison [99, 88].

Leadership and supervision, Metadata curation, Project administration, and Samples and logistics:
Sonia Goncalves [99] and Leigh M Jackson [91].

Leadership and supervision, Metadata curation, Samples and logistics, and Sequencing and analysis:
Ian G Goodfellow [24], Derek J Fairley [3, 72], Matthew W Loose [18] and Joanne Watkins [69].

Leadership and supervision, Metadata curation, Samples and logistics, and Software and analysis tools:
Rich Livett [99].

Leadership and supervision, Metadata curation, Samples and logistics, and Visualisation:
Samuel Moses [25, 106].

Leadership and supervision, Metadata curation, Sequencing and analysis, and Software and analysis tools:
Roberto Amato [99], Sam Nicholls [41] and Matthew Bull [69].

Leadership and supervision, Project administration, Samples and logistics, and Sequencing and analysis:
Darren L Smith [37, 58, 105].

Leadership and supervision, Sequencing and analysis, Software and analysis tools, and Visualisation:
Jeff Barrett [99], David M Aanensen [14, 114].

Metadata curation, Project administration, Samples and logistics, and Sequencing and analysis:
Martin D Curran [65], Surendra Parmar [65], Dinesh Aggarwal [95, 99, 64] and James G Shepherd [48].

Metadata curation, Project administration, Sequencing and analysis, and Software and analysis tools:
Matthew D Parker [93].

Metadata curation, Samples and logistics, Sequencing and analysis, and Visualisation:
Sharon Glaysher [61].

Metadata curation, Sequencing and analysis, Software and analysis tools, and Visualisation:
Matthew Bashton [37, 58], Anthony P Underwood [14, 114], Nicole Pacchiarini [69] and Katie F Loveson [77].

Project administration, Sequencing and analysis, Software and analysis tools, and Visualisation:
Alessandro M Carabelli [88].

Funding acquisition, Leadership and supervision, and Metadata curation:
Kate E Templeton [53, 90].

Funding acquisition, Leadership and supervision, and Project administration:
Cordelia F Langford [99], John Sillitoe [99], Thushan I de Silva [93] and Dennis Wang [93].

Funding acquisition, Leadership and supervision, and Sequencing and analysis:
Dominic Kwiatkowski [99, 107], Andrew Rambaut [90], Justin O'Grady [70, 89] and Simon Cottrell [69].

Leadership and supervision, Metadata curation, and Sequencing and analysis:
Matthew T.G. Holden [68] and Emma C Thomson [48].

Leadership and supervision, Project administration, and Samples and logistics:
Husam Osman [64, 36], Monique Andersson [59], Anoop J Chauhan [61] and Mohammed O Hassan-Ibrahim [6].

Leadership and supervision, Project administration, and Sequencing and analysis:
Mara Lawniczak [99].

Leadership and supervision, Samples and logistics, and Sequencing and analysis:
Ravi Kumar Gupta [88, 113], Alex Alderton [99], Meera Chand [66], Chrystala Constantinidou [94], Meera Unnikrishnan [94], Alistair C Darby [92], Julian A Hiscox [92] and Steve Paterson [92].

Leadership and supervision, Sequencing and analysis, and Software and analysis tools:
Inigo Martincorena [99], David L Robertson [48], Erik M Volz [39], Andrew J Page [70] and Oliver G Pybus [23].

Leadership and supervision, Sequencing and analysis, and Visualisation:
Andrew R Bassett [99].

Metadata curation, Project administration, and Samples and logistics:
Cristina V Ariani [99], Michael H Spencer Chapman [99, 88], Kathy K Li [48], Rajiv N Shah [48], Natasha G Jesudason [48] and Yusri Taha [50].

Metadata curation, Project administration, and Sequencing and analysis:
Martin P McHugh [53] and Rebecca Dewar [53].

Metadata curation, Samples and logistics, and Sequencing and analysis:
Aminu S Jahun [24], Claire McMurray [41], Sarojini Pandey [84], James P McKenna [3], Andrew Nelson [58, 105], Gregory R Young [37, 58], Clare M McCann [58, 105] and Scott Elliott [61].

Metadata curation, Samples and logistics, and Visualisation:
Hannah Lowe [25].

Metadata curation, Sequencing and analysis, and Software and analysis tools:
Ben Temperton [91], Sunando Roy [82], Anna Price [10], Sara Rey [69] and Matthew Wyles [93].

Metadata curation, Sequencing and analysis, and Visualisation:
Stefan Rooke [90] and Sharif Shaaban [68].

Project administration, Samples and logistics, Sequencing and analysis:
Mariateresa de Cesare [98].

Project administration, Samples and logistics, and Software and analysis tools:
Laura Letchford [99].

Project administration, Samples and logistics, and Visualisation:
Siona Silveira [81], Emanuela Pelosi [81] and Eleri Wilson-Davies [81].

Samples and logistics, Sequencing and analysis, and Software and analysis tools:
Myra Hosmillo [24].

Sequencing and analysis, Software and analysis tools, and Visualisation:
Áine O'Toole [90], Andrew R Hesketh [87], Richard Stark [94], Louis du Plessis [23], Chris Ruis [88], Helen Adams [4] and Yann Bourgeois [76].

Funding acquisition, and Leadership and supervision:
Stephen L Michell [91], Dimitris Grammatopoulos[84, 112], Jonathan Edgeworth [12], Judith Breuer [30, 82], John A Todd [98] and Christophe Fraser [5].

Funding acquisition, and Project administration:
David Buck [98] and Michaela John [9].

Leadership and supervision, and Metadata curation:
Gemma L Kay [70].

Leadership and supervision, and Project administration:
Steve Palmer [99], Sharon J Peacock [88, 64] and David Heyburn [69].

Leadership and supervision, and Samples and logistics:
Danni Weldon [99], Esther Robinson [64, 36], Alan McNally [41, 86], Peter Muir [64], Ian B Vipond [64], John BoYes [29], Venkat Sivaprakasam [46], Tranprit Salluja [75], Samir Dervisevic [54] and Emma J Meader [54].

Leadership and supervision, and Sequencing and analysis:
Naomi R Park [99], Karen Oliver [99], Aaron R Jeffries [91], Sascha Ott [94], Ana da Silva Filipe [48], David A Simpson [72] and Chris Williams [69].

Leadership and supervision, and Visualisation:
Jane AH Masoli [73, 91].

Metadata curation, and Samples and logistics:
Bridget A Knight [73, 91], Christopher R Jones [73, 91], Cherian Koshy [1], Amy Ash [1], Anna Casey [71], Andrew Bosworth [64, 36], Liz Ratcliffe [71], Li Xu-McCrae [36], Hannah M Pymont [64], Stephanie Hutchings [64], Lisa Berry [84], Katie Jones [84], Fenella Halstead [46], Thomas Davis [21], Christopher Holmes [16], Miren Iturriza-Gomara [92], Anita O Lucaci [92], Paul Anthony Randell [38, 104], Alison Cox [38, 104], Pinglawathee Madona [38, 104], Kathryn Ann Harris [30], Julianne Rose Brown [30], Tabitha W Mahungu [74], Dianne Irish-Tavares [74], Tanzina Haque [74], Jennifer Hart [74], Eric Witele [74], Melisa Louise Fenton [75], Steven Liggett [79], Clive Graham [56], Emma Swindells [57], Jennifer Collins [50], Gary Eltringham [50], Sharon Campbell [17], Patrick C McClure [97], Gemma Clark [15], Tim J Sloan [60], Carl Jones [15] and Jessica Lynch [2, 111].

Metadata curation, and Sequencing and analysis:
Ben Warne [8], Steven Leonard [99], Jillian Durham [99], Thomas Williams [90], Sam T Haldenby [92], Nathaniel Storey [30], Nabil-Fareed Alikhan [70], Nadine Holmes [18], Christopher Moore [18], Matthew Carlile [18], Malorie Perry [69], Noel Craine [69], Ronan A Lyons [80], Angela H Beckett [13], Salman Goudarzi [77], Christopher Fearn [77], Kate Cook [77], Hannah Dent [77] and Hannah Paul [77].

Metadata curation, and Software and analysis tools:
Robert Davies [99].

Project administration, and Samples and logistics:
Beth Blane [88], Sophia T Girgis [88], Mathew A Beale [99], Katherine L Bellis [99, 88], Matthew J Dorman [99], Eleanor Drury [99], Leanne Kane [99], Sally Kay [99], Samantha McGuigan [99], Rachel Nelson [99], Liam Prestwood [99], Shavanthi Rajatileka [99], Rahul Batra [12], Rachel J Williams [82], Mark Kristiansen [82], Angie Green [98], Anita Justice [59], Adhyana I.K Mahanama [81, 102] and Buddhini Samaraweera [81, 102].

Project administration, and Sequencing and analysis:
Nazreen F Hadjirin [88] and Joshua Quick [41].

Project administration, and Software and analysis tools:
Radoslaw Poplawski [41].

Samples and logistics, and Sequencing and analysis:
Leanne M Kermack [88], Nicola Reynolds [7], Grant Hall [24], Yasmin Chaudhry [24], Malte L Pinckert [24], Iliana Georgana [24], Robin J Moll [99], Alicia Thornton [66], Richard Myers [66], Joanne Stockton [41], Charlotte A Williams [82], Wen C Yew [58], Alexander J Trotter [70], Amy Trebes [98], George MacIntyre-Cockett [98], Alec Birchley [69], Alexander Adams [69], Amy Plimmer [69], Bree Gatica-Wilcox [69], Caoimhe McKerr [69], Ember Hilvers [69], Hannah Jones [69], Hibo Asad [69], Jason Coombes [69], Johnathan M Evans [69], Laia Fina [69], Lauren Gilbert [69], Lee Graham [69], Michelle Cronin [69], Sara Kumziene-SummerhaYes [69], Sarah Taylor [69], Sophie Jones [69], Danielle C Groves [93], Peijun Zhang [93], Marta Gallis [93] and Stavroula F Louka [93].

Samples and logistics, and Software and analysis tools:
Igor Starinskij [48].

Sequencing and analysis, and Software and analysis tools:
Chris J Illingworth [47], Chris Jackson [47], Marina Gourtovaia [99], Gerry Tonkin-Hill [99], Kevin Lewis [99], Jaime M Tovar-Corona [99], Keith James [99], Laura Baxter [94], Mohammad T. Alam [94], Richard J Orton [48], Joseph Hughes [48], Sreenu Vattipally [48], Manon Ragonnet-Cronin [39], Fabricia F. Nascimento [39], David Jorgensen [39], Olivia Boyd [39], Lily Geidelberg [39], Alex E Zarebski [23], Jayna Raghwani [23], Moritz UG Kraemer [23], Joel Southgate [10, 69], Benjamin B Lindsey [93] and Timothy M Freeman [93].

Software and analysis tools, and Visualisation:
Jon-Paul Keatley [99], Joshua B Singer [48], Leonardo de Oliveira Martins [70], Corin A Yeats [14], Khalil Abudahab [14, 114], Ben EW Taylor [14, 114] and Mirko Menegazzo [14].

Leadership and supervision:
John Danesh [99], Wendy Hogsden [46], Sahar Eldirdiri [21], Anita Kenyon [21], Jenifer Mason [43], Trevor I Robinson [43], Alison Holmes [38, 103], James Price [38, 103], John A Hartley [82], Tanya Curran [3], Alison E Mather [70], Giri Shankar [69], Rachel Jones [69], Robin Howe [69] and Sian Morgan [9].

Metadata curation:
Elizabeth Wastenge [53], Michael R Chapman [34, 88, 99], Siddharth Mookerjee [38, 103], Rachael Stanley [54], Wendy Smith [15], Timothy Peto [59], David Eyre [59], Derrick Crook [59], Gabrielle Vernet [33], Christine Kitchen [10], Huw Gulliver [10], Ian Merrick [10], Martyn Guest [10], Robert Munn [10], Declan T Bradley [63, 72] and Tim Wyatt [63].

Project administration:
Charlotte Beaver [99], Luke Foulser [99], Sophie Palmer [88], Carol M Churcher [88], Ellena Brooks [88], Kim S Smith [88], Katerina Galai [88], Georgina M McManus [88], Frances Bolt [38, 103], Francesc Coll [19], Lizzie Meadows [70], Stephen W Attwood [23], Alisha Davies [69], Elen De Lacy [69], Fatima Downing [69], Sue Edwards [69], Garry P Scarlett [76], Sarah Jeremiah [83] and Nikki Smith [93].

Samples and logistics:
Danielle Leek [88], Sushmita Sridhar [88, 99], Sally Forrest [88], Claire Cormie [88], Harmeet K Gill [88], Joana Dias [88], Ellen E Higginson [88], Mailis Maes [88], Jamie Young [88], Michelle Wantoch [7], Sanger Covid Team (www.sanger.ac.uk/covid-team) [99], Dorota Jamrozy [99], Stephanie Lo [99], Minal Patel [99], Verity Hill [90], Claire M Bewshea [91], Sian Ellard [73, 91], Cressida Auckland [73], Ian Harrison [66], Chloe Bishop [66], Vicki Chalker [66], Alex Richter [85], Andrew Beggs [85], Angus Best [86], Benita Percival [86], Jeremy Mirza [86], Oliver Megram [86], Megan Mayhew [86], Liam Crawford [86], Fiona Ashcroft [86], Emma Moles-Garcia [86], Nicola Cumley [86], Richard Hopes [64], Patawee Asamaphan [48], Marc O Niebel [48], Rory N Gunson [100], Amanda Bradley [52], Alasdair Maclean [52], Guy Mollett [52], Rachel Blacow [52], Paul Bird [16], Thomas Helmer [16], Karlie Fallon [16], Julian Tang [16], Antony D Hale [49], Louissa R Macfarlane-Smith [49], Katherine L Harper [49], Holli Carden [49], Nicholas W Machin [45, 64], Kathryn A Jackson [92], Shazaad S Y Ahmad [45, 64], Ryan P George [45], Lance Turtle [92], Elaine O'Toole [43], Joanne Watts [43], Cassie Breen [43], Angela Cowell [43], Adela Alcolea-Medina [32, 96], Themoula Charalampous [12, 42], Amita Patel [11], Lisa J Levett [35], Judith Heaney [35], Aileen Rowan [39], Graham P Taylor [39], Divya Shah [30], Laura Atkinson [30], Jack CD Lee [30], Adam P Westhorpe [82], Riaz Jannoo [82], Helen L Lowe [82], Angeliki Karamani [82], Leah Ensell [82], Wendy Chatterton [35], Monika Pusok [35], Ashok Dadrah [75], Amanda Symmonds [75], Graciela Sluga [44], Zoltan Molnar [72], Paul Baker [79], Stephen Bonner [79], Sarah Essex [79], Edward Barton [56], Debra Padgett [56], Garren Scott [56], Jane Greenaway [57], Brendan AI Payne [50], Shirelle Burton-Fanning [50], Sheila Waugh [50], Veena Raviprakash [17], Nicola Sheriff [17], Victoria Blakey [17], Lesley-Anne Williams [17], Jonathan Moore [27], Susanne Stonehouse [27], Louise Smith [55], Rose K Davidson [89], Luke Bedford [26], Lindsay Coupland [54], Victoria Wright [18], Joseph G Chappell [97], Theocharis Tsoleridis [97], Jonathan Ball [97], Manjinder Khakh [15], Vicki M Fleming [15], Michelle M Lister [15], Hannah C Howson-Wells [15], Louise Berry [15], Tim Boswell [15], Amelia Joseph [15], Iona Willingham [15], Nichola Duckworth [60], Sarah Walsh [60], Emma Wise [2, 111], Nathan Moore [2, 111], Matilde Mori [2, 108, 111], Nick Cortes [2, 111], Stephen Kidd [2, 111], Rebecca Williams [33], Laura Gifford [69], Kelly Bicknell [61], Sarah Wyllie [61], Allyson Lloyd [61], Robert Impey [61], Cassandra S Malone [6], Benjamin J Cogger [6], Nick Levene [62], Lynn Monaghan [62], Alexander J Keeley [93], David G Partridge [78, 93], Mohammad Raza [78, 93], Cariad Evans [78, 93] and Kate Johnson [78, 93].

Sequencing and analysis:
Emma Betteridge [99], Ben W Farr [99], Scott Goodwin [99], Michael A Quail [99], Carol Scott [99], Lesley Shirley [99], Scott AJ Thurston [99], Diana Rajan [99], Iraad F Bronner [99], Louise Aigrain [99], Nicholas M Redshaw [99], Stefanie V Lensing [99], Shane McCarthy [99], Alex Makunin [99], Carlos E Balcazar [90], Michael D Gallagher [90], Kathleen A Williamson [90], Thomas D Stanton [90], Michelle L Michelsen [91], Joanna Warwick-Dugdale [91], Robin Manley [91], Audrey Farbos [91], James W Harrison [91], Christine M Sambles [91], David J Studholme [91], Angie Lackenby [66], Tamyo Mbisa [66], Steven Platt [66], Shahjahan Miah [66], David Bibby [66], Carmen Manso [66], Jonathan Hubb [66], Gavin Dabrera [66], Mary Ramsay [66], Daniel Bradshaw [66], Ulf Schaefer [66], Natalie Groves [66], Eileen Gallagher [66], David Lee [66], David Williams [66], Nicholas Ellaby [66], Hassan Hartman [66], Nikos Manesis [66], Vineet Patel [66], Juan Ledesma [67], Katherine A Twohig [67], Elias Allara [64, 88], Clare Pearson [64, 88], Jeffrey K. J. Cheng [94], Hannah E. Bridgewater [94], Lucy R. Frost [94], Grace Taylor-Joyce [94], Paul E Brown [94], Lily Tong [48], Alice Broos [48], Daniel Mair [48], Jenna Nichols [48], Stephen N Carmichael [48], Katherine L Smollett [40], Kyriaki Nomikou [48], Elihu Aranday-Cortes [48], Natasha Johnson [48], Seema Nickbakhsh [48, 68], Edith E Vamos [92], Margaret Hughes [92], Lucille Rainbow [92], Richard Eccles [92], Charlotte Nelson [92], Mark Whitehead [92], Richard Gregory [92], Matthew Gemmell [92], Claudia Wierzbicki [92], Hermione J Webster [92], Chloe L Fisher [28], Adrian W Signell [20], Gilberto Betancor [20], Harry D Wilson [20], Gaia Nebbia [12], Flavia Flaviani [31], Alberto C Cerda [96], Tammy V Merrill [96], Rebekah E Wilson [96], Marius Cotic [82], Nadua Bayzid [82], Thomas Thompson [72], Erwan Acheson [72], Steven Rushton [51], Sarah O'Brien [51], David J Baker [70], Steven Rudder [70], Alp Aydin [70], Fei Sang [18], Johnny Debebe [18], Sarah Francois [23], Tetyana I Vasylyeva [23], Marina Escalera Zamudio [23], Bernardo Gutierrez [23], Angela Marchbank [10], Joshua Maksimovic [9], Karla Spellman [9], Kathryn McCluggage [9], Mari Morgan [69], Robert Beer [9], Safiah Afifi [9], Trudy Workman [10], William Fuller [10], Catherine Bresner [10], Adrienn Angyal [93], Luke R Green [93], Paul J Parsons [93], Rachel M Tucker [93], Rebecca Brown [93] and Max Whiteley [93].

Software and analysis tools:
James Bonfield [99], Christoph Puethe [99], Andrew Whitwham [99], Jennifier Liddle [99], Will Rowe [41], Igor Siveroni [39], Thanh Le-Viet [70] and Amy Gaskin [69].

Visualisation:
Rob Johnson [39].

1 Barking, Havering and Redbridge University Hospitals NHS Trust, **2** Basingstoke Hospital, **3** Belfast Health & Social Care Trust, **4** Betsi Cadwaladr University Health Board, **5** Big Data Institute, Nuffield Department of Medicine, University of Oxford, **6** Brighton and Sussex University Hospitals NHS Trust, **7** Cambridge Stem Cell Institute, University of Cambridge, **8** Cambridge University Hospitals NHS Foundation Trust, **9** Cardiff and Vale University Health Board, **10** Cardiff University, **11** Centre for Clinical Infection & Diagnostics Research, St. Thomas' Hospital and Kings College London, **12** Centre for Clinical Infection and Diagnostics Research, Department of Infectious Diseases, Guy's and St Thomas' NHS Foundation Trust, **13** Centre for Enzyme Innovation, University of Portsmouth (PORT), **14** Centre for Genomic Pathogen Surveillance, University of Oxford, **15** Clinical Microbiology Department, Queens Medical Centre, **16** Clinical Microbiology, University Hospitals of Leicester NHS Trust, **17** County Durham and Darlington NHS Foundation Trust, **18** Deep Seq, School of Life Sciences, Queens Medical Centre, University of Nottingham, **19** Department of Infection Biology, Faculty of Infectious & Tropical Diseases, London School of Hygiene & Tropical Medicine, **20** Department of

Infectious Diseases, King's College London, **21** Department of Microbiology, Kettering General Hospital, **22** Departments of Infectious Diseases and Microbiology, Cambridge University Hospitals NHS Foundation Trust; Cambridge, UK, **23** Department of Zoology, University of Oxford, **24** Division of Virology, Department of Pathology, University of Cambridge, **25** East Kent Hospitals University NHS Foundation Trust, **26** East Suffolk and North Essex NHS Foundation Trust, **27** Gateshead Health NHS Foundation Trust, **28** Genomics Innovation Unit, Guy's and St. Thomas' NHS Foundation Trust, **29** Gloucestershire Hospitals NHS Foundation Trust, **30** Great Ormond Street Hospital for Children NHS Foundation Trust, **31** Guy's and St. Thomas' BRC, **32** Guy's and St. Thomas' Hospitals, **33** Hampshire Hospitals NHS Foundation Trust, **34** Health Data Research UK Cambridge, **35** Health Services Laboratories, **36** Heartlands Hospital, Birmingham, **37** Hub for Biotechnology in the Built Environment, Northumbria University, **38** Imperial College Hospitals NHS Trust, **39** Imperial College London, **40** Institute of Biodiversity, Animal Health & Comparative Medicine, **41** Institute of Microbiology and Infection, University of Birmingham, **42** King's College London, **43** Liverpool Clinical Laboratories, **44** Maidstone and Tunbridge Wells NHS Trust, **45** Manchester University NHS Foundation Trust, **46** Microbiology Department, Wye Valley NHS Trust, Hereford, **47** MRC Biostatistics Unit, University of Cambridge, **48** MRC-University of Glasgow Centre for Virus Research, **49** National Infection Service, PHE and Leeds Teaching Hospitals Trust, **50** Newcastle Hospitals NHS Foundation Trust, **51** Newcastle University, **52** NHS Greater Glasgow and Clyde, **53** NHS Lothian, **54** Norfolk and Norwich University Hospital, **55** Norfolk County Council, **56** North Cumbria Integrated Care NHS Foundation Trust, **57** North Tees and Hartlepool NHS Foundation Trust, **58** Northumbria University, **59** Oxford University Hospitals NHS Foundation Trust, **60** PathLinks, Northern Lincolnshire & Goole NHS Foundation Trust, **61** Portsmouth Hospitals University NHS Trust, **62** Princess Alexandra Hospital Microbiology Dept., **63** Public Health Agency, **64** Public Health England, **65** Public Health England, Clinical Microbiology and Public Health Laboratory, Cambridge, UK, **66** Public Health England, Colindale, **67** Public Health England, Colindale, **68** Public Health Scotland, **69** Public Health Wales NHS Trust, **70** Quadram Institute Bioscience, **71** Queen Elizabeth Hospital, **72** Queen's University Belfast, **73** Royal Devon and Exeter NHS Foundation Trust, **74** Royal Free NHS Trust, **75** Sandwell and West Birmingham NHS Trust, **76** School of Biological Sciences, University of Portsmouth (PORT), **77** School of Pharmacy and Biomedical Sciences, University of Portsmouth (PORT), **78** Sheffield Teaching Hospitals, **79** South Tees Hospitals NHS Foundation Trust, **80** Swansea University, **81** University Hospitals Southampton NHS Foundation Trust, **82** University College London, **83** University Hospital Southampton NHS Foundation Trust, **84** University Hospitals Coventry and Warwickshire, **85** University of Birmingham, **86** University of Birmingham Turnkey Laboratory, **87** University of Brighton, **88** University of Cambridge, **89** University of East Anglia, **90** University of Edinburgh, **91** University of Exeter, **92** University of Liverpool, **93** University of Sheffield, **94** University of Warwick, **95** University of Cambridge, **96** Viapath, Guy's and St Thomas' NHS Foundation Trust, and King's College Hospital NHS Foundation Trust, **97** Virology, School of Life Sciences, Queens Medical Centre, University of Nottingham, **98** Wellcome Centre for Human Genetics, Nuffield Department of Medicine, University of Oxford, **99** Wellcome Sanger Institute, **100** West of Scotland Specialist Virology Centre, NHS Greater Glasgow and Clyde, **101** Department of Medicine, University of Cambridge, **102** Ministry of Health, Sri Lanka, **103** NIHR Health Protection Research Unit in HCAI and AMR, Imperial College London, **104** North West London Pathology, **105** NU-OMICS, Northumbria University, **106** University of Kent, **107** University of Oxford, **108** University of Southampton, **109** University of Southampton School of Health Sciences, **110** University of Southampton School of Medicine, **111** University of Surrey, **112** Warwick Medical School

DOI: https://doi.org/10.25561/88876

and Institute of Precision Diagnostics, Pathology, UHCW NHS Trust, **113** Wellcome Africa Health Research Institute Durban and **114** Wellcome Genome Campus.

DOI: https://doi.org/10.25561/88876