Changes in symptomatology, re-infection and transmissibility associated with SARS-CoV-2 variant B.1.1.7: an ecological study

Mark S. Graham PhD1*, Carole H. Sudre PhD1,2,3*, Anna May MA4, Michela Antonelli PhD1, Benjamin Murray MSc1, Thomas Varsavsky MSc1, Kerstin Kläser MSc1, Liane S. Canas PhD1, Erika Molteni PhD1, Marc Modat PhD1, David A. Drew PhD5, Long H. Nguyen MD5, Lorenzo Polidori MSc4, Somesh Selvachandran MSc4, Christina Hu MA4, Joan Capdevila PhD4, The COVID-19 Genomics UK (COG-UK) consortium6+, Professor Alexander Hammers PhD1, Professor Andrew T. Chan MD5, Jonathan Wolf MA4, Professor Tim D. Spector PhD7, Claire J. Steves PhD7+, Professor Sebastien Ourselin PhD1+

1. School of Biomedical Engineering & Imaging Sciences, King’s College London, London, UK
2. MRC Unit for Lifelong Health and Ageing, Department of Population Science and Experimental Medicine, University College London, UK
3. Centre for Medical Image Computing, Department of Computer Science, University College London, UK
5. Clinical and Translational Epidemiology Unit, Massachusetts General Hospital and Harvard Medical School, Boston, MA, USA
6. https://www.cogconsortium.uk
7. Department of Twin Research and Genetic Epidemiology, King’s College London, London, UK

* Equal contribution
† Equal contribution
+ Full list of consortium names and affiliations are in the appendix

Corresponding Author:
Mark Graham, PhD
School of Biomedical Engineering and Imaging Sciences
King’s College London,
Lambeth Palace Road,
SE1 7EH
mark.graham@kcl.ac.uk

NOTE: This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.
Abstract

Background
SARS-CoV-2 variant B.1.1.7 was first identified in December 2020 in England. It is not known if the new variant presents with variation in symptoms or disease course, if previously infected individuals may become reinfected with the new variant, or how the variant’s increased transmissibility affects measures to reduce its spread.

Methods
Using longitudinal symptom reports from 36,920 users of the COVID Symptom Study app testing positive for Covid-19 between 28 September and 27 December 2020, we performed an ecological study to examine the association between the regional proportion of B.1.1.7 and reported symptoms, disease course, rates of reinfection, and transmissibility.

Findings
We found no evidence for changes in reported symptoms or disease duration associated with B.1.1.7. We found a likely reinfection rate of 0.7% (95% CI 0.6-0.8), but no evidence that this was higher compared to older strains. We found an increase in R(t) by a factor of 1.35 (95% CI 1.02-1.69). Despite this, we found that R(t) fell below 1 during regional and national lockdowns, even in regions with high proportions of B.1.1.7.

Interpretation
The lack of change in symptoms indicates existing testing and surveillance infrastructure do not need to change specifically for the new variant, and the reinfection findings suggest that vaccines are likely to remain effective against the new variant.

Funding
Zoe Global Limited, Department of Health, Wellcome Trust, EPSRC, NIHR, MRC, Alzheimer’s Society.
Research in context

Evidence before this study

To identify existing evidence on SARS-CoV-2 variant B.1.1.7 we searched PubMed and Google Scholar for articles between 1 December 2020 and 1 February 2021 using the keywords Covid-19 AND B.1.1.7, finding 281 results. We did not find any studies that investigated B.1.1.7-associated changes in the symptoms experienced, their severity and duration, but found one study showing B.1.1.7 did not change the ratio of symptomatic to asymptomatic infections. We found six articles describing laboratory-based investigations of the responses of B.1.1.7 to vaccine-induced immunity to B.1.1.7, but no work investigating what this means for natural immunity and the likelihood of reinfection outside of the lab. We found five articles demonstrating the increased transmissibility of B.1.1.7.

Added value of this study

To our knowledge, this is the first study to explore changes in symptom type and duration, as well as community reinfection rates, associated with B.1.1.7. The work uses self-reported symptom logs from 36,920 users of the COVID Symptom Study app reporting positive test results between 28 September and 27 December 2020. We find that B.1.1.7 is not associated with changes in the symptoms experienced in Covid-19, nor their duration. Building on existing lab studies, our work suggests that natural immunity developed from previous infection provides similar levels of protection to B.1.1.7. We add to the emerging consensus that B.1.1.7 exhibits increased transmissibility.

Implications of all the available evidence

Our findings suggest that existing criteria for obtaining a Covid-19 test in the community need not change for the rise of B.1.1.7. The fact that immunity developed from infection by wild type variants protects against B.1.1.7 provides an indication that vaccines will remain effective against B.1.1.7. R(t) fell below 1 during the UK’s national lockdown, even in regions with high levels of B.1.1.7, but further investigation is required to establish the factors that enabled this, to facilitate countries seeking to control the spread of B.1.1.7.
Introduction

In early December 2020, a phylogenetically distinct cluster of SARS-CoV-2 was genetically characterised in the South-East of England. The majority of cases had been detected in November with a small number detected as early as September. Genomic surveillance reveals that this new variant, termed B.1.1.7, has a number of mutations of immunologic significance and is growing rapidly in frequency and spread. It is important that we understand how these mutations may affect the presentation and spread of Covid-19, so that we can formulate effective public health responses.

Preliminary evidence from epidemiological studies suggests the new strain is more transmissible. Davies et al. found the new strain is 56% (95% CI 50-74) more transmissible and Volz et al. found the new strain increases the effective reproduction number R(t) by a factor of 1.4-1.8. There is evidence to suggest B.1.1.7 increases risk of hospitalisation and death. However, there is much that is still unknown. From a public health perspective, it is crucial to understand if B.1.1.17 necessitates changes to existing measures for disease monitoring and containment. For instance, changes to symptomatology could require modifications to symptomatic testing programmes to ensure new cases are identified, and changes to disease duration could require changes in the amount of time infected individuals are required to isolate for. It is important for modelling and forecasting models to understand if B.1.1.7 alters the rate of reinfection. Early estimates of the new transmissibility of B.1.1.7 are uncertain and there is a need for additional estimates using independent data sources; furthermore it is important to understand how these findings will affect measures to control the pandemic’s spread using non-pharmaceutical interventions, such as lockdowns.

We make use of data from the COVID Symptom Study (CSS) to investigate the symptomatology, disease course, and transmissibility of the new variant. The longitudinal dataset provides symptom reports and test results from a population of over 4 million adults living in the UK using the mobile application. By combining these data with surveillance data from the Covid-19 UK Genetics Consortium (COG-UK) and a spike-gene target failure correlate in community testing data, we performed associative, ecological studies to assess the symptoms, disease course, rates of reinfection, and transmissibility of the new variant.

Methods

Symptom study data

Longitudinal data were prospectively collected using the CSS app, developed by Zoe Global with input from King’s College London (London, UK), the Massachusetts General Hospital (Boston, MA, USA), and Lund and Uppsala Universities (Sweden). The app guides users through a set of enrolment questions, establishing baseline demographic and health information. Users are asked to record each day whether they feel physically normal, and if not, to log any symptoms. After a user reports any symptoms, they are asked “Where are you right now?”, with the options “At home”, “At hospital with suspected Covid-19 Symptoms”, or “Back
from hospital”. Users are also asked to maintain a record of any Covid-19 tests, their date, type, and result in the app. Users are able to record the same data on behalf of others, such as family members, to increase data coverage amongst those unlikely to use mobile applications, such as the elderly. More details about the app can be found in a study by Drew and colleagues. We included users living in the UK who had logged in the app at least once in the period between 28 September to 27 December 2020.

Genomic data

We used data released on 13 January 2021 from COG-UK to extract time-series of the percentage of daily cases that came from the B.1.1.7 lineage in Scotland, Wales, and each of the seven National Health Service (NHS) regions in England. Northern Ireland was excluded due to the low number of samples in the COG-UK dataset. These data are produced by sequencing a sample of polymerase chain reaction (PCR) tests carried out in the community. Due to the delay of approximately two weeks between PCR and genomic sequencing, we only used data from samples taken up to 31 December to avoid censoring effects.

Additionally, we used data from Public Health England (PHE) on the probable new variant captured in community cases in England using spike gene target failure (SGTF). It has been observed that one of the spike gene mutations in B.1.1.7 causes an SGTF in the test used in three of England’s large laboratories used for analysis of community cases. This failure results in a marker that is sensitive to B.1.1.7, but not necessarily specific, as other circulating variants also contain the mutation leading to an SGTF. Comparison to genomic data finds that from 30 November 2020 onwards more than 96% of cases with the SGTF were from lineage B.1.1.7. The proportion of SGTF cases is made available in England for each of the 316 “Lower Tier” Local Authorities. We grouped these data to each NHS region using a population-weighted average to enable integration with other data sources.

Disease symptoms and course

In order to assess whether the symptomatology of infection from B.1.1.7 differed from previous variants, we investigated the change in symptom reporting from 28 September to 27 December 2020, covering 13 complete weeks over the period when the proportion of B.1.1.7 grew most notably in London, South East and East of England. For each week in every region considered, we calculated the proportion of users reporting each symptom. Users were included in a week if they had reported a positive swab test (PCR or lateral flow) in the period 14 days before or after that week. For each region and symptom we performed a linear regression, examining the association between the proportion of B.1.1.7 in that region (independent variable) and the proportion of users reporting the symptom (dependent variable) over the 13 weeks considered. We adjusted for the age and sex of users, as well as two seasonal environmental confounders: regional temperature and humidity. Seasonal confounders were calculated each day the temperature and relative humidity at two meters above the surface, averaged across each region considered.

We also examined the association between proportion of B.1.1.7 and disease burden, measured here as the total number of different symptoms reported over a period of two weeks.
before and two weeks after the test, and the relation with asymptomatic infection, defined as users reporting a positive test result but no symptoms in the two weeks before or after the test. We also investigated the rate of self-reported hospital visits, including both users who reported being in hospital with suspected Covid-19 symptoms, and being back from hospital. We also investigated the proportion of individuals reporting long symptom duration using a previously published definition of continuous symptoms reported for at least 28 days.\textsuperscript{11} To avoid censoring effects, both hospitalisation and long duration analyses included symptom reports extended up to 18 January, and the long duration analysis only considered reports of positive tests up to 21 December. All analyses adjusted for sex, age, temperature and humidity. We controlled for the false discovery rate to account for multiple comparisons.

Reinfection

We defined possible reinfection as the presence of two reported positive tests separated by more than 90 days with a period of reporting no symptoms for more than seven days before the second positive test. We calculated the proportion of possible reinfections among individuals reporting their first positive test before 1 October 2020. To assess whether the risk of reinfection was stronger in the presence of the new variant, in every region we performed ecological studies, examining the Spearman correlation between the proportion of B.1.1.7 cases and the number of reinfections over time, and between the proportion of positive tests reported through the app and the number of reinfections. We compared these two correlations in each region using the Mann-Whitney U test.

Transmissibility

Daily incidence for Scotland, Wales, and each of the seven NHS regions in England were produced from the period 1 October 2020 to 27 December 2020 using data from the CSS app and previously described methodology.\textsuperscript{12} Using the COG-UK data to estimate the proportion of B.1.1.7 in circulation in each region per day, these incidence estimates were decomposed into two incidence time-series per region, one for the old variants and and one for B.1.1.7, with the constraint that the two time-series should sum to match the total incidence. \(R(t)\) was estimated separately for the old and new variants using previously described methodology; briefly, we used the relationship \(I_{t+1} = I_t \exp(\frac{\mu}{\mu/1} (R(t)/I_t - 1))\), where \(1/\mu\) is the serial interval and \(I_t\) the incidence on day \(t\). We modelled the system as a Poisson process and assumed that the serial interval was drawn from a gamma distribution with \(\alpha=6\) and \(\beta=1.5\), and used Markov Chain Monte-Carlo to estimate \(R(t)\). We compared both multiplicative and additive differences of the new and old \(R(t)\) values for days when the proportion of B.1.1.7 in a region was greater than 3%. While data is not available for the proportion of B.1.1.7 in January, we also computed total incidence and \(R(t)\) from 1 October 2020 to 16 January 2021 to see how they changed during national lockdown in England.

Role of the funding source

Zoe Global developed the app for data collection. The funders had no role in the study design, data collection, data analysis, data interpretation, or writing of the report. All authors had full
access to all the data and the corresponding author had final responsibility for the decision to submit for publication.

Results

Symptom study data

Table 1 shows the demographic data for the cohort studied. From 24 March to 27 December 2020, 4,327,245 participants from the UK signed up to use the app. We excluded users living in Northern Ireland due to the low number of sign-ups (38,976), 383,352 users lacking information on sex, and 2,175,979 who had not logged in the app during the period 28 September to 27 December 2020, leaving a total of 1,767,914 users. Between them, these users recorded 65,606,869 logs in the app between 28 September and 27 December. In this period, 497,989 users reported a swab test. 55,192 of these reported a positive test, and we investigated the symptom reports of 36,920 of those whose region was known and who reported as healthy on app sign-up.

<table>
<thead>
<tr>
<th>Overall</th>
<th>Tested</th>
<th>Tested positive**</th>
<th>Signed up healthy with reporting around positive test</th>
</tr>
</thead>
<tbody>
<tr>
<td>N</td>
<td>%</td>
<td>N</td>
<td>%</td>
</tr>
<tr>
<td>Users</td>
<td>1,767,914</td>
<td>---</td>
<td>497,989</td>
</tr>
<tr>
<td>Daily reports*</td>
<td>65,613,697</td>
<td>---</td>
<td>19,154,601</td>
</tr>
<tr>
<td>Age in years</td>
<td>48.4</td>
<td>(19.3)</td>
<td>46.06 (17.8)</td>
</tr>
<tr>
<td>≤18</td>
<td>163,112</td>
<td>9.2</td>
<td>40,717</td>
</tr>
<tr>
<td>19 - 64</td>
<td>1,234,259</td>
<td>69.8</td>
<td>381,900</td>
</tr>
<tr>
<td>≥ 65</td>
<td>370,543</td>
<td>20.9</td>
<td>72,741</td>
</tr>
<tr>
<td>Invalid</td>
<td>5,576</td>
<td>0.3</td>
<td>2,631</td>
</tr>
<tr>
<td>Sex</td>
<td>Female</td>
<td>1,046,074</td>
<td>59.2</td>
</tr>
<tr>
<td></td>
<td>Male</td>
<td>40.8</td>
<td>181,110</td>
</tr>
<tr>
<td>-----------</td>
<td>----------</td>
<td>------</td>
<td>---------</td>
</tr>
<tr>
<td>Intersex</td>
<td>79</td>
<td>&lt;0.1</td>
<td>21</td>
</tr>
<tr>
<td>Prefer not to say</td>
<td>1,199</td>
<td>0.1</td>
<td>983</td>
</tr>
<tr>
<td>Region</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>South East</td>
<td>342,881</td>
<td>19.4</td>
<td>97,143</td>
</tr>
<tr>
<td>East of England</td>
<td>196,063</td>
<td>11.1</td>
<td>57,680</td>
</tr>
<tr>
<td>London</td>
<td>227,004</td>
<td>12.8</td>
<td>81,940</td>
</tr>
<tr>
<td>Midlands</td>
<td>198,350</td>
<td>11.2</td>
<td>57,582</td>
</tr>
<tr>
<td>North East and Yorkshire</td>
<td>156,999</td>
<td>8.9</td>
<td>42,986</td>
</tr>
<tr>
<td>North West</td>
<td>123,201</td>
<td>7.0</td>
<td>45,156</td>
</tr>
<tr>
<td>South West</td>
<td>186,372</td>
<td>10.5</td>
<td>46,780</td>
</tr>
<tr>
<td>Scotland</td>
<td>87,263</td>
<td>4.9</td>
<td>13,793</td>
</tr>
<tr>
<td>Wales</td>
<td>82,886</td>
<td>4.7</td>
<td>16,471</td>
</tr>
<tr>
<td>Not known</td>
<td>165,164</td>
<td>9.3</td>
<td>38,458</td>
</tr>
</tbody>
</table>

Table 1. Characteristics of GB app users active in the period 28 September - 27 December 2020

* Reports logged between 28 September - 27 December. For some analyses we took further reports from an extended time period 14 September 2020 - 18 January 2021
**May be more than one test per individual as the overall number contains failed tests and unknown results
Genomic data

In the period between 27 September 2020 and 31 December 2020, 98,170 sequences were made available by COG-UK, corresponding to 4.4% of the 2,207,476 cases recorded in this period. 16,224 sequences (16.5%) were variant B.1.1.7. Considering the mean of the rolling average across December, the three regions with the largest proportion of B.1.1.7 are the South East, London, and East of England. The three regions with the lowest proportion are Wales, the North East and Yorkshire, and the North West. SGTF data was made available in England on a weekly basis from 10 November 2020 to 29 December 2020. Of the 700,590 cases reported in this period, 295,404 (42.2%) caused an SGTF. Examining the COG-UK data from England in the same time period, we find 34.6% cases are B.1.1.7. The difference is in part attributable to the SGTF being a nonspecific marker of B.1.1.7: in the week from 9-15 November 81% of cases with an SGTF were B.1.1.7, while from 30 November at least 96% of cases with the SGTF were from B.1.1.7. Figure 1 shows how the proportion of the new variant changed over time in regions of the UK using COG-UK and the SGTF data.

Figure 1. Presence of B.1.1.7 in each of the 7 NHS regions in England, and Scotland and Wales, as measured using genomic surveillance data (COG-UK) and SGTF data. SGTF data are not available for Scotland or Wales.

Disease symptoms and course

Figure 2 illustrates the variation of symptom occurrence over time considering a one-week window smoothed over 3 time points as a function of time, and Supplementary Figure 1 shows...
how these symptoms vary as a function of the proportion of B.1.1.7. Qualitatively, these results show no change in the proportion of users reporting each symptom with the new variant. Linear regression (both unadjusted and adjusted for age, sex, temperature and humidity) did not find evidence of association between the proportion of B.1.1.7 and symptoms reported after controlling the false discovery rate. Both adjusted and unadjusted results are shown in Supplementary Figure 4.

Figure 3 shows the variation of total number of symptoms reported, the total number of asymptomatic infections, self-reported hospital visits, and instances of long symptom duration over time, and Supplementary Figure 2 shows how these plots vary with proportion of B.1.1.7. Visually, we see no change in any of these outcomes with an increasing proportion of B.1.1.7. When correcting for mean age, sex, ambient temperature and humidity there was no evidence of an association between B.1.1.7 and either the number of symptoms reported over a 4-week window, the number of hospitalisations, long symptom duration, or proportion of asymptomatic cases, see Supplementary Table 1.
**Figure 2.** Regional plots of the frequency of reporting of symptoms over time for each reported symptom. Drop in fever reporting in early November was caused by a change in the question wording; this wording was subsequently reverted a week later.

**Figure 3.** Regional plot of hospitalisation reports, proportion of asymptomatic positives, instances of long symptom duration and the total number of different symptoms reported. For the study of long symptom duration, tests are only considered up to 21 December, symptom reports up to 18 January 2021 to limit right censoring effects. Only symptomatic individuals for which duration can be ascertained are included.
Reinfection

Overall, we identified 304 individuals reporting two positive tests with more than 90 days between the two. Among these individuals, symptom reporting allowed us to identify 249 for which there is a period of at least 7 symptom-free days in between positive tests among the 36,509 individuals having reported a positive swab test before 1 October 2020 (0.7%, 95% CI 0.6-0.8). Among those 249, daily reports were available in the periods around both of the positive tests for 173. There was no difference in reinfection reporting rates across the different NHS regions (p=0.1). Figure 4 shows the evolution in the number of possible reinfections along with reported positive cases (green line) and proportion of B.1.1.7 (red line). For all regions (except Scotland), reinfection occurrences were more positively correlated with the overall regional rise in cases rather than the regional rise in the new variant percentage (Number of cases:reinfection, Spearman rho 0.55 to 0.69 [p<0.05] for South East, London and East of England; % new variant:reinfection, Spearman rho 0.37 to 0.55 in the same regions).

Supplementary Table 2 shows the bootstrapped median values of correlation compared across the different regions and the outcome of a Mann-Whitney U test across the bootstrapped distributions.
Figure 4. Number of reinfection reports by region according to week of the second infection, along with the total number of positive tests reported through the app (green line) and the proportion of B.1.1.7 in circulation (red line).

Transmissibility

Figure 5 shows incidence and $R(t)$ for the old and new variants in the three regions in England with the highest proportions of the new variant. Results consistently show the $R(t)$ of B.1.1.7 to be greater than that of other variants. The mean (95% CI) of the additive increase in $R$ for B.1.1.7 was 0.34 (0.02-0.66), and the multiplicative increase was 1.35 (1.02-1.69). England exited its second national lockdown on 2 December, leading to a change in behaviour and $R(t)$. When considering only the period after the second lockdown ended, we find 0.28 (0.01-0.61) for the additive and 1.28 (1.02-1.61) for the multiplicative increases. Supplementary Figure 3 shows the same using the SGTF data, with analysis limited to the period after 1 December when at least 95% of all SGTF cases were B.1.1.7. These data are provided weekly, and linear interpolation was used to obtain daily estimates, leading to smoother estimates for variant-
specific incidence and $R(t)$. Using these values, we find $R(t)$ of B.1.1.7 has an additive increase of 0.26 (0.15-0.37) and a multiplicative increase of 1.25 (1.17-1.34).

On 19 December 2020 London and much of the South East and East of England were placed in ‘Tier 4’ restrictions, enforcing stricter rules for social distancing and decreased human-to-human contact that stopped short of nationwide measures. On 5 January 2021 the whole of England was placed in national lockdown. Figure 6 shows overall incidence and $R(t)$ for the longer period from 1 October 2020 to 16 January 2021 in the three regions with the largest proportion of B.1.1.7. The proportion of B.1.1.7 in these regions in January is at least 80%, assuming the proportion has not decreased from the end of December. $R(t)$ fell to ~0.8 in all three of these regions during national lockdown. An extended plot including Scotland, Wales, and all regions in England, is shown in Supplementary Figure 5.

Figure 5. Incidence and $R(t)$ for the old and new variants, along with the ratio between these $R$ values, for the three regions in England with the largest proportion of B.1.1.7. Dark grey regions indicate national lockdowns, light grey the period where London and much of the South East and East of England were placed in Tier 4 restrictions.
Figure 6. Total incidence and R(t) for the three regions with the highest proportion of B.1.1.7 in December, extended to capture the third national lockdown beginning 5 January 2021. Dark grey regions indicate national lockdowns, light grey indicate the period where London and much of the South East and East of England were placed in Tier 4 restrictions.

Discussion

Using data collected through community reporting of symptoms and tests via the COVID Symptom Study app, we performed an ecological study to investigate whether the appearance of the variant B.1.1.7, first detected in a sample from England in September 2020, was associated with differences in symptoms experienced, disease duration, hospitalisation, asymptomatic infection, risks of reinfection, and transmissibility for users reporting a positive test result between 28 September and 27 December 2020.

We did not find associations between the proportion of B.1.1.7 in circulation and the type of symptoms experienced by our app users. We also did not find evidence for any change in the total number of symptoms experienced by individuals associated with B.1.1.7, nor the proportion of individuals experiencing long disease duration, defined as recording symptoms for more than 28 days without a break of more than seven days. We were also able to assess rates of asymptomatic disease and hospitalisation associated with B.1.1.7. We found the proportion of users with asymptomatic disease did not significantly change as B.1.1.7 increased in prevalence, in agreement with other work on the subject.\textsuperscript{14} We did not find any changes in hospitalisation, but other work has found evidence that B.1.1.7 increases hospitalisation rates.\textsuperscript{6} Limitations to the assessments of asymptomatic rates and hospitalisation should be noted: the majority of our users only get tested when they have symptoms and so we have relatively few asymptomatic infections recorded, and the self-reported nature of our hospitalisation data means we likely miss cases of more severe hospitalisation, when the individual is unlikely to self-report. There is also evidence that infection with B.1.1.7 carries increased risk of mortality,\textsuperscript{6} our data does not allow for us to assess this.
A recent study by the UK Office for National Statistics Community Infection Survey (CIS) reported that individuals infected with B.1.1.7 were more likely to report a cough, sore throat, fatigue, myalgia and fever in the seven days preceding the test, and less likely to report a loss of taste or smell.\textsuperscript{15} It is not clear if this report adjusted for age, sex, and environmental factors, though our analysis did not find this adjustment affected our findings (Supplementary Figure 4). The discrepancy may be explained by sampling at different points in the disease course. Our users predominantly seek testing at symptom onset, whilst the CIS design means the test may be administered at any point during the disease. Recent evidence has shown that B.1.1.7 causes longer infections\textsuperscript{16} which means the CIS symptom reports from users infected with B.1.1.7 may be sampled from later in the disease course than non-B.1.1.7 cases, causing apparent differences. The periods considered also differed: we considered symptoms reported both two weeks before and after the positive test result, compared to the one week before the positive test considered by the CIS. Further opportunity to study symptoms with B.1.1.7 in different contexts are required to be definitive.

We observed, based on 249 potential cases, a very low rate of possible reinfection of 0.7% (95% CI 0.6- 0.8). This rate is consistent with another study of 6614 healthcare workers that had previously tested positive for Covid-19, finding 44 possible reinfections (0.66%).\textsuperscript{17} Our reinfection rate did not vary consistently across regions or time, which would be consistent with the hypothesis that reinfection is no more likely in the context of B.1.1.7. This may mean that if adequate immunity is built over the first infection it may be sufficient to protect against reinfection in the presence of B.1.1.7. Ultimately this is a positive sign that the immunity built through vaccination against the old variants could also be effective against B.1.1.7. This is in line with initial, laboratory-based studies regarding the efficacy of vaccines designed for early strains against this newer variant.\textsuperscript{18-20}

We found an increase in the reproduction number $R(t)$ in association with the B.1.1.7 variant: we found a multiplicative increase in $R(t)$ of $\sim 1.35$ (95% 1.02-1.69), compatible with estimates from Volz et al, of 1.4-1.8, Davies et al. who estimated a transmissibility increase of 1.56 (95% CI 1.50-1.74), and Walker et al who found an increase in growth rate that corresponds to a transmissibility increase of 1.33 (95% CI 1.21 - 1.53) if we assume a generation time of 4.7 days.\textsuperscript{4,5,14,21} These increases in transmissivity have worrying implications for the ability of lockdown measures to control B.1.1.7, given $R(t)$ was estimated to be 0.7-0.9 during the first national lockdown in England.\textsuperscript{22} Despite this, we found $R(t)$ to be $\sim 0.8$ in the three regions in England with at least 80% of B.1.1.7. during the national lockdown beginning on 5 January 2021. There are several potential explanations for this. It could be that adherence to this lockdown is greater than previous lockdowns, helping to reduce $R(t)$. It may also indicate the true increase in transmissivity is at the lower end of the available estimates, or that the increase in transmissivity estimated outside of lockdown cannot be extrapolated to lockdown, perhaps due to B.1.1.7 responding differently to lockdown measures than the old variants. Another possible explanation is that there is now sufficient community immunity to reduce $R(t)$ further than seen during previous lockdowns. One serology study estimates that in the period 21 December 2020 - 18 January 2021, 15.3% (95% CI 14.7% to 15.9%) individuals in England would have tested positive for Covid-19 antibodies.\textsuperscript{23} Many countries have now detected cases...
of B.1.1.7 and work to better understand the factors that helped suppress it in the UK will help other countries to formulate their public health responses.24

Our study has several strengths. The large, longitudinal nature of the CSS data, with good coverage of the UK population, provides a unique opportunity to study potential changes in symptomatology and disease duration. The ability to match tests and symptom reports over long periods further allows us to measure possible reinfection rates. Our data also offers the ability to provide a valuable complementary measure to existing measurements of the increased transmissibility of B.1.1.7: we were able to use real-time, representative incidence estimates to measure R(t), whilst other studies have relied on deaths and hospitalisations, which are lagged, or community case numbers which do not reflect true infection numbers.

We acknowledge several limitations to this work. As we lack information on the disease strain of individual positive infections reported through the app, we performed an ecological study, assessing the association between the proportion of B.1.1.7 and population-level measures. This design does not allow for causal interpretation of the effect of B.1.1.7 on the measures we investigate. Our work assumes that all non-B.1.1.7 variants in circulation at the time of study give rise to the same symptomatology and immune response, and have the same transmissibility. Genomic surveillance has detected a very low number of non-B.1.1.7 variants of concern in circulation25, supporting the validity of this assumption, but it cannot be ruled out that other variants with different characteristics are circulating undetected. Data obtained from participatory, digital platforms have well-documented26 biases in demographics. Whilst we were able to correct for some of these factors in our analysis, such as age and sex, there are others that are more difficult to characterise and correct for. For example, respondents signing up to a participatory platform such as the CSS app are likely to be more interested in health and COVID-19 than the wider population, and may exhibit different behaviour. Participatory studies may also suffer from ascertainment or collider bias.27 Self report also carries the risk of data input errors, although the design of the app seeks to minimise this; for example, each time a user submits a log in the app they are shown the full history of their test results and are given the option to amend incorrect entries. Previous publications from our group have found that population-level estimates of disease prevalence from our app triangulate well with those obtained from studies designed to be representative of the population.12 . We make the assumption that testing positive for SARS-CoV-2 after an interval of 90 days with at least a seven day period with an absence of symptoms is consistent with reinfection. Repeated positive testing has been reported shortly after hospital discharge28 and showed that PCR positivity could be detected up to 28 days post symptom resolution. While the chosen cut-off of 90 days between two positive tests is unlikely to be due to prolonged PCR positivity, this cannot be ruled out, but would only affect a small number of cases. Viral sequencing of the two infections would ideally be required to confirm reinfection. Despite correcting for changes in temperature and humidity, a possible limitation in the study is that comparisons in symptoms are made over time, and seasonal effects (e.g. on symptoms) may not have been fully taken into account.29
Conclusions

We examined the effect of SARS-CoV-2 variant B.1.1.7 on the symptoms, disease course, rates of reinfection, and transmissibility in the UK. We found no change in symptoms or their duration. We found a low rate of reinfection (0.7%) and no evidence of increased rates associated with B.1.1.7. We found an increase in R(t) of ~ 1.38 (95% CI 1.06-1.71), but evidence that R(t) fell below 1 during lockdown even in regions with very high (>80%) proportions of B.1.1.7.

Ethics

Ethics has been approved by KCL Ethics Committee REMAS ID 18210, review reference LRS-19/20-18210 and all participants provided consent.

Data sharing

Data collected in the COVID Symptom Study smartphone application are being shared with other health researchers through the UK National Health Service-funded Health Data Research UK (HDRUK) and Secure Anonymised Information Linkage consortium, housed in the UK Secure Research Platform (Swansea, UK). Anonymised data are available to be shared with researchers according to their protocols in the public interest (https://web.www.healthdatagateway.org/dataset/fddcb382-3051-4394-8436-b92295f14259). US investigators are encouraged to coordinate data requests through the Coronavirus Pandemic Epidemiology Consortium (https://www.monganinstitute.org/cope-consortium).

Author Contributions

MSG, CHS, ATC, JW, TDS, CJS, SO contributed to study concept and design. CHS, AM, BM, DAD, LHN, LP, SS, CH, JCP, The COVID-19 Genomics UK (COG-UK) consortium, ATC, JW, TDS, CJS, SO contributed to acquisition of data. MSG, CHS, AM, TV, contributed to data analysis and have verified the underlying data. MSG, CHS contributed to initial drafting of the manuscript. MGS, CHS, CJS and SO were responsible for the decision to submit the manuscript. All authors contributed to interpretation of data and critical revision of the manuscript. ATC, CJS, TDS, SO contributed to study supervision.

Acknowledgements

ZOE Global provided in kind support for all aspects of building, running and supporting the app and service to all users worldwide. 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.
Support for this study was provided by the NIHR-funded Biomedical Research Centre based at GSTT NHS Foundation Trust. Investigators also received support from the Wellcome Trust (212904/Z/18/Z, WT203148/Z/16/Z), the MRC/BHF (MR/M016560/1), Alzheimer’s Society, EU, NIHR, CDRF, and the NIHR-funded BioResource, Clinical Research Facility and BRC based at GSTT NHS Foundation Trust in partnership with KCL, the UK Research and Innovation London Medical Imaging & Artificial Intelligence Centre for Value Based Healthcare, the Wellcome Flagship Programme (WT213038/Z/18/Z), the Chronic Disease Research Foundation, and DHSC. ATC was supported in this work through a Stuart and Suzanne Steele MGH Research Scholar Award. The Massachusetts Consortium on Pathogen Readiness (MassCPR) and Mark and Lisa Schwartz supported MGH investigators (DAD, LHN, ATC).

Declaration of interests

AM, LP, SS, JCP, CH, JW are employees of Zoe Global Ltd. TDS is a consultant to Zoe Global Ltd. DAD and ATC previously served as investigators on a clinical trial of diet and lifestyle using a separate smartphone application that was supported by Zoe Global. ATC reports grants from Massachusetts Consortium on Pathogen Readiness, during the conduct of the study; personal fees from Pfizer Inc., personal fees from Boehringer Ingelheim, personal fees from Bayer Pharma AG, outside the submitted work. DAD reports grants from National Institutes of Health, grants from MassCPR, grants from American Gastroenterological Association during the conduct of the study.

References


6 Peter Horby, Catherine Huntley, Nick Davies, John Edmunds, Neil Ferguson, Graham Medley, Andrew Hayward, Muge Cevik, Calum Semple. Update note on B.1.1.7 severity. NERVTAG, 2021.


19 Wu K, Werner AP, Moliva JI, et al. mRNA-1273 vaccine induces neutralizing antibodies against spike mutants from global SARS-CoV-2 variants. DOI:10.1101/2021.01.25.427948.


24 Grubaugh ND, Hodcroft EB, Fauver JR, Phelan AL, Cevik M. Public health actions to control new SARS-CoV-2 variants. *Cell* 2021; published online Jan 29. DOI:10.1016/j.cell.2021.01.044.


Supplementary material

Supplementary tables and figures

- Severe fatigue
- Abdominal pain
- Chest pain
- Sore throat
- Severe shortness of breath
- Skipped meals
- Loss of smell
- Unusual muscle pains
- Headache
- Hoarse voice
- Delirium
- Diarrhoea
- Fever
- Persistent cough

- South East
- East of England
- London
- Midlands
- North East and Yorkshire
- North West
- South West
- Scotland
- Wales
Supplementary Figure 1. Regional plots of the frequency of reporting of symptoms over time for each reported symptom, against the proportion of B.1.1.7. Drop in fever reporting in early November was caused by a change in the question wording; this wording was subsequently reverted a week later.
Supplementary Figure 2. Regional plot of hospitalisation reports, proportion of asymptomatic positives, instances of long symptom duration and the total number of different experienced symptoms against proportion of B.1.1.7. For the study of long symptom duration, tests are only considered up to 21 December, and symptom reports up to 18 January 2021 to limit right censoring effects. Only symptomatic individuals for which duration can be ascertained are included.
Supplementary Figure 3. Incidence and $R(t)$ for the old and new variants, along with the ratio between these $R$ values, for the three regions in England with the largest proportion of B.1.1.7, using SGTF data. Dark grey regions indicate national lockdowns, light grey shaded the period where London and much of the South East and East of England were placed in Tier 4 restrictions.
**Supplementary Figure 4:** Colour plot of beta values and associated p-values for each region and symptom when investigating association between symptom report (in a 4 week window around the test) and proportion of variant B.1.1.7. Top row shows results for uncorrected model, and the bottom row shows results for the model corrected for age, sex, temperature and humidity. Note that the p-values are capped at 0.1. Beta-values are presented for an increase of 0.1 in the proportion of variant B.1.1.7.


**Supplementary Figure 5:** Total incidence and R(t) for all regions, extended to capture the third national lockdown beginning 5 January 2021. Dark grey regions indicate national lockdowns, light grey indicate the period where London and much of the South East and East of England were placed in Tier 4 restrictions.
<table>
<thead>
<tr>
<th>Region</th>
<th>Proportion of fully asymptomatic</th>
<th>Number of symptoms reported over 4 weeks around test</th>
<th>Proportion of hospital reports</th>
<th>Proportion of individuals with symptom duration &gt;= 28 days</th>
</tr>
</thead>
<tbody>
<tr>
<td>South East</td>
<td>0.001 [-0.015;0.017] ; 0.901</td>
<td>-0.021 [-0.163;0.121] ; 0.733</td>
<td>-0.002 [-0.011;0.007] ; 0.624</td>
<td>-0.003 [-0.009;0.004] ; 0.37</td>
</tr>
<tr>
<td>East of England</td>
<td>0.002 [-0.008;0.012] ; 0.588</td>
<td>-0.012 [-0.153;0.13] ; 0.851</td>
<td>-0.002 [-0.01;0.006] ; 0.52</td>
<td>-0.002 [-0.015;0.01] ; 0.689</td>
</tr>
<tr>
<td>London</td>
<td>-0.005 [-0.014;0.005] ; 0.298</td>
<td>0.031 [-0.055;0.116] ; 0.423</td>
<td>-0.002 [-0.007;0.003] ; 0.298</td>
<td>-0.002 [-0.013;0.009] ; 0.682</td>
</tr>
<tr>
<td>Midlands</td>
<td>-0.016 [-0.028;0.004] ; 0.014</td>
<td>0.02 [-0.133;0.173] ; 0.766</td>
<td>-0.002 [-0.007;0.003] ; 0.328</td>
<td>0.002 [-0.01;0.015] ; 0.671</td>
</tr>
<tr>
<td>North East and Yorkshire</td>
<td>-0.011 [-0.046;0.023] ; 0.462</td>
<td>-0.086 [-0.444;0.272] ; 0.586</td>
<td>-0.011 [-0.04;0.019] ; 0.426</td>
<td>0.015 [-0.022;0.052] ; 0.349</td>
</tr>
<tr>
<td>North West</td>
<td>-0.005 [-0.023;0.013] ; 0.512</td>
<td>-0.053 [-0.218;0.111] ; 0.468</td>
<td>-0.009 [-0.015;0.004] ; 0.005</td>
<td>0 [-0.031;0.031] ; 0.98</td>
</tr>
<tr>
<td>South West</td>
<td>0.015 [-0.011;0.04] ; 0.217</td>
<td>-0.261 [-0.437;0.085] ; 0.01</td>
<td>-0.001 [-0.02;0.018] ; 0.902</td>
<td>-0.048 [-0.091;0.004] ; 0.036</td>
</tr>
<tr>
<td>Scotland</td>
<td>0.022 [-0.013;0.058] ; 0.177</td>
<td>-0.4 [-0.711;0.088] ; 0.019</td>
<td>-0.018 [-0.037;0.002] ; 0.073</td>
<td>-0.012 [-0.027;0.003] ; 0.107</td>
</tr>
<tr>
<td>Wales</td>
<td>-0.002 [-0.05;0.047] ; 0.943</td>
<td>-0.041 [-0.683;0.601] ; 0.884</td>
<td>-0.008 [-0.045;0.028] ; 0.602</td>
<td>-0.053 [-0.141;0.035] ; 0.192</td>
</tr>
</tbody>
</table>
**Supplementary Table 1:** Beta coefficient of the variant proportion when evaluating association with number of reported symptoms, asymptomatic rate, proportion of hospital report and proportion of individuals with duration >28 days (among symptomatic) across the different regions when correcting for age, sex, temperature and humidity. All values are presented for an increase in 0.1 in the proportion of variant B.1.1.7. All results are presented in the form mean [CI]; p-value

<table>
<thead>
<tr>
<th>Region</th>
<th>Correlation Variant/Reinfection</th>
<th>Correlation New cases/Reinfection</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>South East</td>
<td>0.55</td>
<td>0.69</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>East of England</td>
<td>0.51</td>
<td>0.56</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>London</td>
<td>0.46</td>
<td>0.62</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Midlands</td>
<td>0.28</td>
<td>0.75</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>North East and Yorkshire</td>
<td>-0.02</td>
<td>0.30</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>North West</td>
<td>0.06</td>
<td>0.43</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>South West</td>
<td>-0.35</td>
<td>0.05</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Scotland</td>
<td>0.59</td>
<td>-0.15</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Wales</td>
<td>0.07</td>
<td>0.26</td>
<td>&lt;0.001</td>
</tr>
</tbody>
</table>

**Supplementary Table 2:** Comparison of regional correlation over time between proportion of B.1.1.7 and number of possible reinfections and between new reported cases and number of possible reinfections. Medians over 100 bootstrapped samples are calculated for each and compared using a Mann-Whitney U test.
COG-UK authorship 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:

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.

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:
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:

Leadership and supervision, Sequencing and analysis, and Software and analysis tools:

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

Metadata curation, Project administration, and Samples and logistics:

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, Claire 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 88.

Project administration, Samples and logistics, Sequencing and analysis:
Mariateresa de Cesare 88.

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 68, Helen Adams 4 and Yann Bourgeois 76.

Funding acquisition, and Leadership and supervision:

Funding acquisition, and Project administration:
David Buck 96 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:

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, Melissa 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:

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:

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

Sequencing and analysis, and Software and analysis tools:

Software and analysis tools, and Visualisation:

Leadership and supervision:

Metadata curation:

Project administration:

Samples and logistics:

Sequencing and analysis:
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:

Visualisation:
Rob Johnson 39.
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 and Institute of Precision Diagnostics, Pathology, UHCW NHS Trust, 113 Wellcome Africa Health Research Institute Durban and 114 Wellcome Genome Campus.