Spatial growth rate of emerging SARS-CoV-2 lineages in England, September 2020–December 2021

M. R. Smallman-Raynor1, A. D. Cliff2 and The COVID-19 Genomics UK (COG-UK) Consortium3,†

1School of Geography, University of Nottingham, Nottingham, UK; 2Department of Geography, University of Cambridge, Cambridge, UK and 3https://www.cogconsortium.uk

Abstract
This paper uses a robust method of spatial epidemiological analysis to assess the spatial growth rate of multiple lineages of SARS-CoV-2 in the local authority areas of England, September 2020–December 2021. Using the genomic surveillance records of the COVID-19 Genomics UK (COG-UK) Consortium, the analysis identifies a substantial (7.6-fold) difference in the average rate of spatial growth of 37 sample lineages, from the slowest (Delta AY.4.3) to the fastest (Omicron BA.1). Spatial growth of the Omicron (B.1.1.529 and BA) variant was found to be 2.81× faster than the Delta (B.1.617.2 and AY) variant and 3.76× faster than the Alpha (B.1.1.7 and Q) variant. In addition to AY.4.2 (a designated variant under investigation, VUI-21OCT-01), three Delta sublineages (AY.43, AY.98 and AY.120) were found to display a statistically faster rate of spatial growth than the parent lineage and would seem to merit further investigation. We suggest that the monitoring of spatial growth rates is a potentially valuable adjunct to outbreak response procedures for emerging SARS-CoV-2 variants in a defined population.

Introduction
Emerging lineages of the severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) have the potential to place significant pressure on public health systems due to increased infectivity, transmissibility, virulence, immune escape or other fitness advantage [1, 2]. Global genomic surveillance has identified >1700 SARS-CoV-2 lineages since the beginning of the COVID-19 pandemic [3, 4], of which Alpha (B.1.1.7 and Q), Beta (B.1.351), Gamma (P.1), Delta (B.1.617.2 and AY) and Omicron (B.1.1.529 and BA) have been designated as variants of concern by the World Health Organization (WHO) on account of their global public health significance [5]. Additional lineages are currently classified on the basis of properties that are suggestive of an emerging (variants of interest) or possible future (variants under monitoring) risk to global public health [5]. The risk is well illustrated by the recent and rapid emergence of Omicron as the dominant variant in the UK, South Africa and the USA, among other countries, in late November and December 2021 [6–8].

One important epidemiological facet of an emerging SARS-CoV-2 lineage is its propensity to grow in a defined population [9]. There are well-established methods for assessing the rate of temporal growth by, for example, examining the trajectory of case doubling times or estimating the basic reproduction number, $R_0$, of the agent in question [10, 11]. Viewed from a geographical perspective, these measures are essentially aspatial in that they provide very little information on the geographical growth, or spatial expansion, of the associated infection wave. To extend the examination of SARS-CoV-2 growth rates into the spatial domain, the present paper applies a robust method of spatial epidemiological analysis that is known as the swash-backwash model of the single epidemic wave [12] to the genomic surveillance records of the COVID-19 Genomics UK (COG-UK) Consortium [13]. Using the spatial sequence of detection of sample variants as a proxy for the spatial wave front of infection, our examination yields estimates of the spatial growth rate of multiple SARS-CoV-2 lineages in the local authority areas of England, September 2020–December 2021.

For a total of 37 sample lineages under investigation, we present evidence of a substantial (7.6-fold) difference in the average rate of spatial growth, from the slowest (Delta AY.4.3) to the fastest (Omicron BA.1). Whilst the overall results for the Alpha, Delta and Omicron variants are consistent with the documented growth advantages for these lineages, several emergent Delta sublineages (AY.4.2, AY.43, AY.98 and AY.120) are found to have had a statistically significant growth advantage over the parent lineage. To our knowledge, this is the first comparative study of the spatial growth rate of multiple emerging SARS-CoV-2 lineages at the national level. It is also the first report of a spatial growth advantage for the Delta AY.43, AY.98...
and AY.120 lineages, and the first to document an apparently reduced spatial growth rate for a substantial number of other AY lineages that emerged in the spring and summer of 2021. The modelling of spatial growth rates is equally applicable to the analysis of RT-PCR gene target data, and we suggest it to be a potentially valuable adjunct to outbreak response procedures for SARS-CoV-2 variants in a defined population.

Data and methods

Since September 2020, successive waves of SARS-CoV-2 infection with emerging lineages of the Alpha (September 2020 onset), Delta (March 2021 onset) and Omicron (November 2021 onset) variants have been recorded in England [14–16]. The weekly record of COVID-19 cases to mid-December 2021 is plotted in Figure 1a, whilst the underpinning sequence of variants is depicted in Figure 1b. As Figure 1b shows, Alpha, Delta and Omicron achieved the status of dominant variants in December 2020, May 2021 and December 2021, respectively.

Data

We draw on the integrated national-level SARS-CoV-2 genomic surveillance records of the COG-UK Consortium [13]. These records are based on unselected (random sample) sequencing of positive SARS-CoV-2 test samples that have been identified through standard (‘pillar 2’) diagnostic pathways in the UK. Lineages are assigned using the Phylogenetic Assignment of

Named Global Outbreak Lineages (pangolin) tool, with lineage counts made available by local authority area and week of sample collection. For further information on the data under examination, see COG-UK Consortium, COVID-19 Genomic Surveillance [18].

Lineage counts for England were accessed from the COG-UK website [18] for a 68-week period, September 2020 (epidemiological week 36, ending 5 September) to December 2021 (epidemiological week 50, ending 18 December) (Fig. 1b). The data set included geo-coded information on 979 075 SARS-CoV-2 samples assigned to the 309 Lower Tier Local Authority (LTLA) divisions of England. Here, we define the 309 LTALAs according to their most recent (May 2021) status. Information on the lineage of 20 655 samples (2.1%) was either suppressed (1 105) or not recorded (19 550). Of the remaining 958 420 samples, the majority (93.8%) were classified as belonging to the B.1.1.7 and Q (Alpha, 153 405 samples), B.1.617.2 and AY (Delta, 722 133 samples) and B.1.1.529 and BA (Omicron, 23 137 samples) lineages (Table 1). Samples belonging to these lineages form the basis of all our analysis.

Methods

To assess the spatial growth rate of a given SARS-CoV-2 lineage, we draw on the swash-backwash model of the single epidemic wave [12]. In essence, the model represents a spatial derivative of the generic SIR mass action models of infectious disease transmission [19]. Using the binary (presence/absence) of a disease, the model (i) allows the disaggregation of an infection wave into phases of spatial expansion and retreat and (ii) provides a means of measuring the phase transitions of geographical units from susceptible $S$, through infective $I$ to recovered $R$ status. See, for example, Smallman-Raynor and Cliff [20] and Smallman-Raynor et al. [21].

Measuring the spatial growth rate

Full details of the modelling procedure are outlined by Cliff and Haggett [12]. For the purposes of the present analysis, we focus on the spatial expansion phase (i.e. the change of state from $S$ to $I$ across a set geographical units) for a given SARS-CoV-2 lineage. Specifically, let the first week in which the lineage was detected in England be coded as $t = 1$. Subsequent weeks were then coded serially as $t = 2, 3, ..., T$, where $T$ is the number of weekly periods from the beginning to the end of the detected occurrence of the lineage. For any given geographical unit, we refer to the first week in which the lineage was detected as the leading edge (LE) of the infection wave. The average time (in weeks) to the detection of the lineage across the set of units can then be defined by a time-weighted mean, $\overline{t}_{LE}$, of the form

$$\overline{t}_{LE} = \frac{1}{N} \sum_{i=1}^{N} t_{ni}.$$  \hspace{1cm} (1)

Here, $n_i$ is the number of units whose leading edge, $LE$, occurred in week $t$ and $N = \sum n_i$. Formed in this manner, SARS-CoV-2 lineages with relatively high rates of spatial expansion (or rapidly developing $LE$) take on relatively low values of $\overline{t}_{LE}$ (i.e. short average times to detection). Conversely, lineages with relatively low rates of spatial expansion (or slowly developing $LE$) take on relatively high values of $\overline{t}_{LE}$ (i.e. long average times to detection).
Equation (1) was used to estimate the spatial growth rate of sample SARS-CoV-2 lineages for which the earliest detection in England occurred in the time period covered by the dataset (September 2020–December 2021) and for which substantial geographical spread had been documented. To ensure the inclusion

| Variant/lineage | LTLAs (n) | Number of detections\(t\) | Earliest detection\(t = 1\) | \(\bar{t}_{LE}\) (95% CI) (weeks) |
|-----------------|-----------|----------------------------|-------------------------------|----------------------------------|
| Alpha (B.1.1.7 and Q) | 307 | 153405 | 39/2020 (26 Sept.) | 9.90 (9.56–10.23) |
| Delta (B.1.617.2 and AV) | 307 | 722133 | 12/2021 (27 March) | 7.40 (7.12–7.68) |
| B.1.617.2 | 307 | 17504 | 13/2021 (3 April) | 9.11 (8.48–9.74) |
| AY.3 | 168 | 620 | 27/2021 (10 July) | 14.12 (13.26–14.98) |
| AY.4 | 307 | 547403 | 12/2021 (27 March) | 9.45 (9.24–9.66) |
| AY.4.1 | 133 | 416 | 25/2021 (26 June) | 9.50 (8.61–10.40) |
| AY.4.2 | 307 | 77391 | 25/2021 (26 June) | 6.21 (5.86–6.57) |
| AY.4.2.1 | 307 | 11541 | 29/2021 (24 July) | 9.43 (8.91–9.94) |
| AY.4.3 | 188 | 781 | 19/2021 (15 May) | 19.93 (18.92–20.93) |
| AY.4.5 | 158 | 742 | 24/2021 (19 June) | 15.06 (13.76–16.37) |
| AY.5 | 307 | 26111 | 15/2021 (17 April) | 9.83 (9.38–10.28) |
| AY.6 | 306 | 17405 | 17/2021 (1 May) | 8.88 (8.48–9.27) |
| AY.7 | 276 | 3627 | 18/2021 (8 May) | 9.02 (8.39–9.65) |
| AY.8 | 141 | 1241 | 16/2021 (24 April) | 8.60 (7.90–9.29) |
| AY.9 | 304 | 10136 | 14/2021 (10 April) | 10.88 (10.33–11.42) |
| AY.9.2 | 251 | 1990 | 28/2021 (17 July) | 10.44 (9.70–11.17) |
| AY.10 | 118 | 683 | 15/2021 (17 April) | 8.50 (7.96–9.04) |
| AY.25 | 130 | 486 | 31/2021 (7 Aug.) | 13.02 (12.26–13.78) |
| AY.33 | 268 | 3088 | 30/2021 (31 July) | 10.88 (10.19–11.57) |
| AY.36 | 256 | 1999 | 28/2021 (17 July) | 12.69 (12.05–13.32) |
| AY.42 | 115 | 338 | 27/2021 (10 July) | 10.08 (8.81–11.34) |
| AY.43 | 307 | 23068 | 27/2021 (10 July) | 5.56 (5.27–5.86) |
| AY.46 | 305 | 9461 | 21/2021 (29 May) | 11.23 (10.72–11.74) |
| AY.46.5 | 304 | 8649 | 23/2021 (12 June) | 10.13 (9.66–10.61) |
| AY.87 | 172 | 688 | 20/2021 (22 May) | 10.28 (9.28–11.29) |
| AY.89 | 173 | 544 | 21/2021 (29 May) | 13.97 (13.17–14.77) |
| AY.90 | 218 | 1202 | 21/2021 (29 May) | 13.52 (12.54–14.50) |
| AY.98 | 307 | 22465 | 22/2021 (5 June) | 6.34 (5.99–6.70) |
| AY.98.1 | 172 | 621 | 25/2021 (26 June) | 15.77 (14.80–16.75) |
| AY.109 | 116 | 441 | 25/2021 (26 June) | 17.59 (16.67–18.52) |
| AY.111 | 296 | 4649 | 21/2021 (29 May) | 15.25 (14.46–16.04) |
| AY.120 | 306 | 18400 | 20/2021 (22 May) | 6.56 (6.16–6.97) |
| AY.122 | 302 | 5018 | 21/2021 (29 May) | 13.80 (13.07–14.53) |
| AY.124 | 119 | 424 | 25/2021 (26 June) | 12.99 (11.93–14.05) |
| AY.125 | 162 | 534 | 26/2021 (3 July) | 16.65 (15.79–17.51) |
| Omicron (B.1.1.529 and BA) | 304 | 23137 | 47/2021 (27 Nov.)\(c\) | 2.63 (2.56–2.71)\(d\) |
| Average | 233 | 46450 | | 11.27 (10.03–12.50) |

\(a\)Excludes 1105 detections for which lineage data are suppressed and 19550 detections for which lineage data are not available.

\(b\)Epidemiological week/year, with the last day of the week given in parentheses.

\(c\)Excludes a lone detection in week 43 (30 October).

\(d\)Indexed to week 47; \(\bar{t}_{LE} = 6.62 (6.53, 6.70)\) when indexed to week 43.

**Application of the model**

Equation (1) was used to estimate the spatial growth rate of sample SARS-CoV-2 lineages for which the earliest detection in England occurred in the time period covered by the dataset (September 2020–December 2021) and for which substantial geographical spread had been documented. To ensure the inclusion
of sufficient observations for geographical analysis, the sample was limited to lineages that had been detected in at least one-third of the 309 LTLAs by December 2021. Based on these criteria, the sample consisted of 37 lineages. Summary details of the sample, including the number of LTLAs in which each lineage was detected, the total count of detections over the study period and the earliest date of detection, are provided in Table 1.

For each lineage, equation (1) was fitted with \( t = 1 \) set to the week of earliest detection in Table 1. In the instance of Omicron, retrospective analysis has identified a lone detection of the BA.1 lineage in epidemiological week 43 of 2021 (week ending 30 October), 4 weeks prior to the subsequent detection and apparent onset of widespread transmission of the variant in epidemiological week 47 (week ending 27 November). For the purposes of the present analysis, we set week 47 as \( t = 1 \) for Omicron, but we also report the computed value of \( 1/t_{LE} \) based on the earlier detection in week 43. Finally, we exclude two LTLAs (City of London and Isles of Scilly) from all analysis on account of the suppression of lineage data due to their small populations. Data analysis was performed in Minitab®17 (Minitab Inc., Pennsylvania, USA) and data mapping in QGIS 3.10.14-A Coruña (QGIS.org) using Local Authority Districts (May 2021) UK and Regions (December 2020) EN shapefiles from the Office for National Statistics (ONS) [22].

### Results

Table 1 confirms that the 37 sample lineages were geographically extensive in their transmission, with 29 having been detected in >150 LTLAs, 21 in >250 LTLAs, 16 in >300 LTLAs and nine in the complete set of 307 LTLAs under examination. The majority (23) were associated with >1000 detections, 13 with >10,000 detections and three with >100,000 detections. Delta (B.1.617.2 and AY) was the most common lineage (722,133 detections) and AY.4 the most common sublineage (547,403 detections), with AY lineages accounting for 33 of the spread events under examination. In turn, the majority of lineages emerged (as judged by the date of earliest detections) in the spring and summer of 2021, as the Delta infection wave was evolving both domestically and internationally.

### Spatial growth curves and leading edge (LE) maps

The upper graphs in Figure 2 plot the count of LTLAs by week of earliest detection of the Alpha (B.1.1.7 and Q), Delta (B.1.617.2 and AY) and Omicron (B.1.1.529 and BA) variants in England, September 2020–December 2021. The graphs plot, on a weekly basis, the non-cumulative count (upper) and cumulative proportion (lower) of LTLAs in which each of the three variants was first detected. The horizontal (time) axes are indexed to the epidemiological week of first detection (\( t = 2 \)) of the corresponding variant. Average curves, formed across the set of sample lineages in Table 1, are plotted for reference.

Together, the graphs in Figure 2 portray the temporal development of the spatial leading edges (LE) for each variant. The geographical expression of these LE is captured by the choropleth maps in Figure 3 which plot the week of earliest detection of each variant in the set of LTLAs. The sequentially more rapid spatial growth of the variants (Alpha → Delta → Omicron) is evidenced by the sequentially steeper spatial growth curves (Fig. 2) and the sequentially shorter periods to earliest detection (Fig. 3). The latter feature is emphasised when earliest detections are formed as regional averages in Figure 4.
Rates of spatial growth ($\overline{t}_{LE}$)

The right-hand column in Table 1 summarises the results of the application of equation (1) to each of the sample lineages. Computed values of $\overline{t}_{LE}$ and associated 95% confidence intervals (95% CI) are given, along with an overall average value of $\overline{t}_{LE}$ for the entire sample. As noted above, lineages with relatively high rates of spatial expansion (or rapidly developing LE) are represented by relatively low values of $\overline{t}_{LE}$ (i.e. short average times to detection), while lineages with relatively low rates of spatial expansion (or slowly developing LE) take on relatively high values of $\overline{t}_{LE}$ (i.e. long average times to detection). In this manner, the table confirms the sequential increase in the spatial growth rate for Alpha, Delta and Omicron. On average, the earliest detection of the Alpha variant in a given LTLA occurred at $\overline{t}_{LE} = 9.90$ (95% CI 9.56–10.23) weeks after the earliest sampled detection in England. This reduced to 7.40 (95% CI 7.12–7.68) weeks for Delta and 2.63 (95% CI 2.56–2.71) weeks for Omicron.

**Delta AY lineages**

Figure 5 is based on the information in Table 1 and plots the values of $\overline{t}_{LE}$ for B.1.617.2 and AY lineages in order, from the lowest (left, high values of $\overline{t}_{LE}$) to the highest (right, low values of $\overline{t}_{LE}$)
\( t_{LE} \) rates of spatial growth. Values are plotted on an inverted vertical scale to facilitate interpretation. The average value of \( t_{LE} \), formed across the sample set of lineages in Table 1, is indicated for reference as are the \( t_{LE} \) for the Alpha (B.1.1.7 and Q), Delta (B.1.617.2 and AY) and Omicron (B.1.1.529 and BA) variants. Spatial growth curves, formed in the manner of Figure 2, are plotted for a sample of 20 AY lineages with relatively high and low rates of spatial growth in Figure 6.

There is a 7.6-fold difference in the range of values of \( t_{LE} \) in Figure 5, from Delta AY.4.3 with the lowest spatial growth rate (19.93 weeks) to Omicron with the highest (2.63 weeks). A group of four AY lineages (AY.4.2, AY.43, AY.98 and AY.120), first detected in the period from mid-May to mid-July 2021, are positioned between Delta and Omicron in Figure 5 and display rates of spatial growth that are significantly higher (as judged by 95% CI) than the aggregate rate for the Delta variant. In contrast, the overwhelming majority of AY lineages display statistically lower – in many instances substantially lower – spatial growth rates (as judged by 95% CI) than the aggregate rate for the Delta variant.

**Discussion**

Recent experience has underscored the importance of the ongoing tracking, monitoring and analysis of emerging SARS-CoV-2 lineages with a view to mitigating the impacts of the COVID-19 pandemic [23]. We have used a robust model of spatial epidemiological analysis to estimate the rate of spatial growth of multiple lineages of the virus in England over a 68-week period, September 2020–December 2021. We have shown that the Alpha, Delta and Omicron variants took an average of 9.90, 7.40 and 2.63 weeks, respectively, to reach the set of LTLAs under examination (Table 1 and Fig. 5). Expressed in relative terms, the leading spatial edges were 1.34× faster (Delta vs. Alpha), 2.81× faster (Omicron vs. Delta) and 3.76× faster (Omicron vs. Alpha). Our estimates scale to the approximate length of time that Alpha (12 weeks), Delta (8 weeks) and Omicron (3 weeks) took to establish themselves as the dominant variants in England [18], and are consistent with evidence for the fitness advantage of Delta over Alpha and Omicron over Delta [11, 24, 25].

Of the 121 Delta AY lineages detected in England to December 2021 and included in the genomic surveillance records of the COG-UK Consortium, 33 lineages met the geographical criterion for inclusion in the current analysis. In interpreting the results for these lineages, we note that AY designations are phylogenetically defined and do not necessarily denote any fundamental biological differences between the lineages [26]. Moreover, results of the type documented in this paper are context dependent and cannot be interpreted as evidence of a change in biological transmissibility, immune escape or other fitness advantage. Subject to these caveats, we have identified four AY lineages (AY.4.2, AY.43, AY.98 and AY.120) for which the rate of spatial growth exceeded the aggregate rate for the Delta variant. These lineages had been detected in all (AY.4.2, AY.43 and AY.98) or most (AY.120) of the local authority areas under investigation, and each had been associated with considerably more than 10 000 detections (Table 1). Table 2 summarises the global status of these four lineages as of 9 January 2022. With the exception of the AY.43 lineage, which was prevalent in a number of European countries and associated with >267 000 detections worldwide, the majority of detections of these lineages originated from the UK.

Our findings for the AY.4.2 and AY.43 lineages are consistent with their respective designations by the UK Health Security Agency as a distinct variant under investigation (VUI-21OCT-01) and a variant of concern [32, 33]. Preliminary investigations indicated the AY.4.2 lineage to be associated with a higher growth rate and a higher household secondary attack rate, but with no significant reduction in vaccine effectiveness, as compared to the parent lineage [32, 34]. Although the factors underpinning the higher growth rate of AY.4.2 remain to be established [32, 35, 36], we observe that this lineage accounted for a maximum of 24.4% of all detections (week ending 4 December 2021) before being outcompeted by Omicron [18]. Similarly, the status of the AY.43 lineage in terms of transmission advantage and/or immune escape remains to be determined, although further investigation is merited as new AY.43 sublineages have recently been reported from Brazil [37]. Finally, our identification of a rapid rate of spatial growth for the AY.98 and AY.120 lineages, approximating the estimated rates for AY.4.2 and AY.43, is noteworthy. Whilst very little has been documented on the epidemiological facets of these lineages, both have been identified in a number of countries in Europe and elsewhere (Table 2) and would seem to merit further investigation on the basis of the findings presented here.

With the foregoing exceptions, our analysis has shown that many emerging AY lineages in England in the spring and summer of 2021 were associated with spatial growth rates that were lower (in some instances, substantially lower) than the aggregate rate for the Delta variant (Table 1 and Fig. 5). Multiple biological (e.g. reduced infectivity or transmissibility) and contextual (e.g. progressive expansion of the national COVID-19 vaccination programme) factors may account for this observation. Importantly, there is no evidence of a temporal trend in the observed rates of spatial growth that would be suggestive of either (i) a biological selection pressure in favour of a growth advantage of emerging lineages or (ii) a progressive contextual effect in the form of, for example, increasing levels of vaccination coverage or natural immunity that would serve to retard growth rates.

It is important to emphasise the broader societal and epidemiological context to the spread of SARS-CoV-2 lineages that will have influenced our estimates of \( t_{LE} \) in Table 1 and

![Fig. 5. Estimated rate of spatial growth of sample SARS-CoV-2 lineages in England, September 2020–December 2021. The graph plots values of \( t_{LE} \) and associated 95% CI from Table 1. Values are ordered from the lowest (left, high values of \( t_{LE} \)) to the highest (right, low values of \( t_{LE} \)) rates of spatial growth. Values are plotted on an inverted vertical scale to facilitate interpretation. The average value of \( t_{LE} \) for the sample is shown for reference.](image-url)
Figure 5. For the time period covered by the present study, non-pharmaceutical interventions (NPIs) included: a tier system of local lockdown in October 2020; two periods of national lockdown (November–December 2020 and January–March 2021); a phased lifting of national restrictions in the period to July 2021; and the implementation of ‘Plan B’ control guidelines against the Omicron variant in December 2021 [38]. Whilst the phases of national lockdown had significant impacts on population mobility, mixing and associated opportunities for SARS-CoV-2 transmission [39], it is noteworthy that the majority (27) of lineages included in the present analysis were first detected in the period from May to July 2021 (Table 1). This corresponded with the final steps in the Government’s four-stage roadmap for the lifting of lockdown measures and was marked by a substantial...

Fig. 6. Spatial growth curves for sample Delta sublineages in England, March–December 2021. Curves have been formed in the manner of the lower graphs in Figure 2, with the average curve plotted for reference. Lineages are ordered according to the values of $t_{LE}$ in Table 1 and are defined as having relatively high (i.e. low values of $t_{LE}$; upper graphs, a) and relatively low (i.e. high values of $t_{LE}$; lower graphs, b) rates of spatial growth.

Table 2. Worldwide detection of sample Delta AY lineages with relatively high estimated rates of spatial growth (status: 9 January 2022)

| Variant | Countries | Global prevalence (%) | Number of detections | Predominant countries (proportion of global detections) |
|---------|-----------|-----------------------|----------------------|--------------------------------------------------------|
| AY.4.2  | 52        | 1                     | 82 038               | UK (90%), Denmark (4%), Germany (1%), Poland (1%), France (1%) |
| AY.43   | 128       | 4                     | 267 426              | Germany (15%), UK (13%), Denmark (13%), France (12%), Belgium (5%) |
| AY.98   | 66        | 1                     | 41 507               | UK (91%), USA (2%), Germany (1%), Denmark (1%), Ireland (1%) |
| AY.120  | 74        | 1                     | 30 320               | UK (85%), India (3%), USA (3%), Germany (2%), France (1%) |

Sources: data from cov-lineages.org [27] and Latif et al. [28–31].
easing and eventual removal of restrictions on social mixing [38].
To set against this easing of restrictions, lineage growth rates will
have been retarded to an unknown extent by the immunity
afforded by prior infection with antigenically similar SARS-CoV-2 variants (B.1.617.2 and AY sublineages, in particu-
lar) and by the phased rollout of the national COVID-19 vaccin-
ation programme [40].

The results we have presented are subject to the limitations of the
available lineage data. Although the COG-UK Consortium
genomic surveillance data are recognised for their extent and reli-
bility [41], the data are formed as a sample of positive SARS-CoV-2 test results and are subject to the limitations and
biases of sample data. In this context, we note that the cumulative
coverage of the COG-UK records for England was estimated at
13.7% of people with positive SARS-CoV-2 test results to
October 2021 [42]. We also note that the sample test data are
derived from a laboratory system with testing capabilities that
vary by region and time period [9]. Such space-time variations
have potentially important implications for analyses, of the type
outlined in the present paper, that are dependent on the dates of
first detection of SARS-CoV-2 lineages in a multi-region
setting.

Our results are also subject to the underpinning assumptions
of the analytical procedure. In particular, the computation of
$L_{LE}$ is dependent on the specification of the index week (i.e. the
week that a given lineage was first detected in England) and
the degree to which this reflects the date of actual emergence of
the lineage in England. The extent to which the sample data
accurately track the spatial expansion of the $LE$ for a given
lineage, the variable contributions of international travel- and
community-related transmission to the development of the $LE$,
and the geographical starting point(s) of a given lineage in the
national transmission network, will also have influenced our
results in unknown ways. For example, the early involvement
and high degree of geographical connectivity of London and the
South East may have served to accelerate the spatial transmission
of the Alpha variant in the latter months of 2020 [14]. The
observed rapid spread of the Delta variant may reflect inter-
national importations and onwards transmission from multiple
different geographical locations in the spring of 2021 [15, 43],
whilst early cases of the Omicron variant were observed in highly
connected regions at a time of reduced NPIs in November and
December 2021 [44].

For the purposes of the present analysis, our application of the
swash-backwash model has utilised genomic surveillance data. We
note, however, that the modelling approach is equally applicable
to the analysis of RT-PCR gene target data. As such, the approach
may be used to facilitate timely assessments of the spatial growth
of emerging SARS-CoV-2 variants and thereby contribute to
rapid outbreak responses [9, 45].

Further insights into the spatial growth and decay of
SARS-CoV-2 lineages may be gained by application of the full
swash-backwash model, but this is dependent on the substantial
spatial retreat of any given lineage from the population. Here we
note that, with the exception of AY.10 (last detected in July 2021)
and AY.8 and Alpha (B.1.1.7 and Q) (both last detected in
August/September 2021), there is evidence of the circulation of all
the lineages included in Table 1 in the weeks to December 2021.

We have demonstrated, for the first time, a robust method for
assessing and comparing the rate of spatial growth of multiple
SARS-CoV-2 lineages in a set of geographical areas. We suggest
that this approach represents a potentially valuable adjunct to
outbreak response procedures for emerging SARS-CoV-2 variants
in a defined population.

Acknowledgements. 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) [grant code: MC_PC_19027],
and Genome Research Limited, operating as the Wellcome Sanger Institute.
The authors acknowledge use of data generated through the COVID-19
Genomics Programme funded by the Department of Health and Social Care.
The views expressed are those of the author and not necessarily those
of the Department of Health and Social Care or PHE or UKHSA.

Conflict of interest. None.

Data availability statement. The data that support the findings of this
study are available at Wellcome Sanger Institute COVID–19 Genomic
Surveillance (https://covid19.sanger.ac.uk/lineages/raw).

References

1. Dubey A et al. (2021) Emerging SARS-CoV-2 variants: genetic variability and
clinical implications. Current Microbiology 79, 20.

2. Mukherjee R and Satardekar R (2021) Why are some coronavirus
variants more infectious? Journal of Biosciences 46, 101.

3. O’Ttoole À et al. (2021) Assignment of epidemiological lineages in an
emerging pandemic using the pangolin tool. Virus Evolution 7, veab064.

4. GISAID. Available at https://www.gisaid.org/hcov19-variants (Accessed
11 January 2022).

5. World Health Organization. Tracking SARS-CoV-2 variants. Available at
https://www.who.int/en/activities/tracking-SARS-CoV-2-variants/ (Accessed
11 January 2022).

6. UK Health Security Agency. Variants of concern or under investigation:
data up to 22 December 2021. Available at https://www.gov.uk/
government/publications/covid-19-variants-genomically-confirmed-case-
numbers/variants/distribution-of-case-data-23-december-2021 (Accessed
11 January 2022).

7. Centers for Disease Control and Prevention. COVID data tracker. Available at
https://covid.cdc.gov/covid-data-tracker/#/nowcast-heading
(Accessed 11 January 2022).

8. Latif AA et al. (2022) BA.1 Lineage Report. Outbreak.info. Available at
https://outbreak.info/situation-reports?pango=BA.1 (Accessed 11 January
2022).

9. Ward T et al. (2021) Growth, reproduction numbers and factors affecting
the spread of SARS-CoV-2 novel variants of concern in the UK from
October 2020 to July 2021: a modelling analysis. BMJ Open 11, e056636.

10. UK Health Security Agency (2021) SARS-CoV-2 Variants of Concern and
Variants under Investigation in England Technical Briefing 31. London:
UK Health Security Agency, p. 43.

11. UK Health Security Agency (2021) SARS-CoV-2 Variants of Concern and
Variants under Investigation in England Technical Briefing 33. London:
UK Health Security Agency, p. 42.

12. Cliff AD and Haggett P (2006) A swash-backwash model of the single
epidemic wave. Journal of Geographical Systems 8, 227–252.

13. COVID-19 Genomics UK (COG-UK) Consortium (2020) An integrated
national scale SARS-CoV-2 genomic surveillance network. Lancet Microbe 1,
e99–e100.

14. Davies NG et al. (2021) Estimated transmissibility and impact of
SARS-CoV-2 lineage B.1.1.7 in England. Science 372, eabg3055.

15. Mishra S et al. (2021) Changing composition of SARS-CoV-2 lineages and
rise of Delta variant in England. EClinicalMedicine 39, 101064.

16. UK Health Security Agency (2021) SARS-CoV-2 Variants of Concern and
Variants under Investigation in England Technical Briefing: Update on
Hospitalisation and Vaccine Effectiveness for Omicron VOC-21NOV-01
(B.1.1.529). London: UK Health Security Agency, p. 17.

17. GOV.UK. Coronavirus (COVID-19) in the UK. Available at https://cor-
onavirus.data.gov.uk/ (Accessed 11 January 2022).
18. COVID-19 Genomics UK (COG-UK) Consortium. COVID-19 Genomic Surveillance. Available at https://covid19.sanger.ac.uk/lineages/raw (Accessed 11 January 2022).
19. Anderson RM and May R (1991) Infectious Diseases of Humans: Dynamics and Control. Oxford: Oxford University Press, p. 757.
20. Smallman-Raynor MR and Cliff AD (2014) Abrupt transition to heightened poliomyelitis epidemicity in England and Wales, 1947–1957, associated with a pronounced increase in the geographical rate of disease propagation. Epidemiology and Infection 142, 577–591.
21. Smallman-Raynor MR, Cliff AD and Stickler PJ (2022) Meningococcal meningitis and coal mining in provincial England: geographical perspectives on a major epidemic, 1929–33. Geographical Analysis 54, 197–216. doi: 10.1111/gean.12272.
22. Office for National Statistics (ONS). Open geography portal. Available at https://geoportal.statistics.gov.uk/ (Accessed 11 January 2022).
23. Angeletti S et al. (2022) SARS-CoV-2 AY.4.2 variant circulating in Italy: genomic preliminary insight. Journal of Medical Virology 94, 1689–1692. doi: 10.1002/jmv.27451.
24. He X et al. (2021) The challenges of COVID-19 Delta variant: prevention and vaccine development. MedComm (2020) 2, 846–854.
25. Mahase E (2021) Covid-19: do vaccines work against omicron—and other questions answered. British Medical Journal 375, n3062.
26. Public Health England (2021) SARS-CoV-2 Variants of Concern and Variants under Investigation in England Technical Briefing 23. London: Public Health England, p. 61. cov.lineages.org. PANGO lineages: latest epidemiological lineages of SARS-CoV-2. Available at https://cov-lineages.org/ (Accessed 11 January 2022).
27. Latif AA et al. (2022) AY.4.2 Lineage Report. Outbreak.info. Available at https://outbreak.info/situation-reports/pango=BA.1 (Accessed 11 January 2022).
28. Latif AA et al. (2022) AY.43 Lineage Report. Outbreak.info. Available at https://outbreak.info/situation-reports/pango=BA.1 (Accessed 11 January 2022).
29. Latif AA et al. (2022) AY.98 Lineage Report. Outbreak.info. Available at https://outbreak.info/situation-reports/pango=BA.1 (Accessed 11 January 2022).
30. Latif AA et al. (2022) AY.120 Lineage Report. Outbreak.info. Available at https://outbreak.info/situation-reports/pango=BA.1 (Accessed 11 January 2022).
31. UK Health Security Agency (2021) SARS-CoV-2 Variants of Concern and Variants under Investigation in England Technical Briefing 27. London: UK Health Security Agency, p. 63.
32. UK Health Security Agency (2021) SARS-CoV-2 Variants of Concern and Variants under Investigation in England Technical Briefing 29. London: UK Health Security Agency, p. 45.
33. Le Page M (2021) New variant gains ground. New Scientist 252, 8.
34. UK Health Security Agency (2021) SARS-CoV-2 Variants of Concern and Variants under Investigation in England Technical Briefing 26. London: UK Health Security Agency, p. 31.
35. UK Health Security Agency. Variants of concern or under investigation: data up to 5 January 2022. Available at https://www.gov.uk/government/publications/covid-19-variants-genomically-confirmed-case-numbers/variants-distribution-of-case-data-7-january-2022 (Accessed 11 January 2022).
36. Lima ARJ et al. (2021) SARS-CoV-2 genomic monitoring in the São Paulo State unveils new sublineages of the AY.43 strain. medRxiv. https://doi.org/10.1101/2021.11.29.21266819.
37. Brown J and Kirk-Wade E (2021) Coronavirus: A History of 'Lockdown Laws' in England. London: House of Commons Library, p. 44.
38. Shepherd HER et al. (2021) Domestic and international mobility trends in the United Kingdom during the COVID-19 pandemic: an analysis of Facebook data. International Journal of Health Geographies 20, 46.
39. NHS England. COVID-19 vaccinations archive. Available at https://www.england.nhs.uk/statistics/statistical-work-areas/covid-19-vaccinations/covid-19-vaccinations-archive/ (Accessed 11 January 2022).
40. Robishaw JD et al. (2021) Genomic surveillance to combat COVID-19: challenges and opportunities. Lancet Microbe 2, e481–e484.
41. COVID-19 Genomics UK (COG-UK) Consortium (2020) Coverage reports. Available at https://www.cogconsortium.uk/news-reports/coverage-reports/ (Accessed 11 January 2022).
42. McConne JT et al. (2021) Context-specific emergence and growth of the SARS-CoV-2 Delta variant. medRxiv. doi:10.1101/2021.12.14.21267606.
43. UK Health Security Agency (2021) SARS-CoV-2 Variants of Concern and Variants under Investigation in England. Variant of Concern: Omicron, VOC21NOV-01 (B.1.1.529) Technical Briefing 30. London: UK Health Security Agency, p. 40.
44. European Centre for Disease Prevention and Control/World Health Organization Regional Office for Europe (2021) Methods for the detection and characterisation of SARS-CoV-2 variants – first update. Available at https://www.ecdc.europa.eu/sites/default/files/documents/Methods-for-the-detection-and-characterisation-of-SARS-CoV-2-variants-first-update.pdf (Accessed 23 April 2022).

Appendix: The COVID-19 Genomics UK (COG-UK) Consortium

- 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, 84
- Funding acquisition, Leadership and supervision, Metadata curation, Project administration, Samples and logistics, Sequencing and analysis, Software and analysis tools, and Visualisation: Tanya Golubchik 5
- Funding acquisition, Leadership and supervision, Metadata curation, Samples and logistics, Sequencing and analysis, and Visualisation: Rocío T Martínez Nunez 46
- Funding acquisition, Leadership and supervision, Project administration, Samples and logistics, Sequencing and analysis, and Software and analysis tools: Luke B Snell 12
- Funding acquisition, Leadership and supervision, Project administration, Samples and logistics, Software and analysis tools, and Visualisation: Rich Livett 116
- Funding acquisition, Leadership and supervision, Metadata curation, Project administration, and Samples and logistics: Catherine Ludden 20, 70
- Funding acquisition, Leadership and supervision, Metadata curation, Samples and logistics, and Sequencing and analysis: Sally Corden 74 and Eleni Nastouli 96, 95, 30
- Funding acquisition, Leadership and supervision, Metadata curation, Sequencing and analysis, and Software and analysis tools: Gaia Nebbia 12
- Funding acquisition, Leadership and supervision, Project administration, Samples and logistics, and Sequencing and analysis: Ian Johnston 116
Leadership and supervision, Metadata curation, Project administration, Samples and logistics, and Sequencing and analysis: Katrina Lythgoe 7, M. Estee Torok 19, 20 and Ian G Goodfellow 24

Leadership and supervision, Metadata curation, Project administration, Samples and logistics, and Visualisation: Jacqui A Prieto 97, 82 and Kordo Saeed 97, 83

Leadership and supervision, Metadata curation, Project administration, Sequencing and analysis, and Software and analysis tools: David K Jackson 116

Leadership and supervision, Metadata curation, Samples and logistics, Sequencing and analysis, and Visualisation: Catherine Houlihan 96, 94

Leadership and supervision, Metadata curation, Sequencing and analysis, Software and analysis tools, and Visualisation: Dan Frampton 94, 95

Metadata curation, Project administration, Samples and logistics, Sequencing and analysis, and Software and analysis tools: William L Hamilton 19 and Adam A Witney 41

Funding acquisition, Samples and logistics, Sequencing and analysis, and Visualisation: Giselda Bucca 101

Funding acquisition, Leadership and supervision, Metadata curation, and Project administration: Cassie F Pope 40, 41

Funding acquisition, Leadership and supervision, Metadata curation, and Samples and logistics: Catherine Moore 74

Funding acquisition, Leadership and supervision, Metadata curation, Sequencing and analysis, and Visualisation: Emma C Thomson 92

Funding acquisition, Leadership and supervision, Project administration, and Samples and logistics: Ewan M Harrison 116, 102

Funding acquisition, Leadership and supervision, Sequencing and analysis, and Visualisation: Colin P Smith 101

Leadership and supervision, Metadata curation, Project administration, and Sequencing and analysis: Fiona Rogan 77

Leadership and supervision, Metadata curation, Project administration, and Samples and logistics: Shaun M Beckwith 6, Abigail Murray 6, Dawn Singleton 6, Kirstine Eastick 37, Liz A Sheridan 98, Paul Randell 99, Leigh M Jackson 1005, Cristina V Ariani 116 and Sónia Gonçalves 116

Leadership and supervision, Metadata curation, Samples and logistics, and Sequencing and analysis: Derek J Fairley 3, 5, Matthew W Loose 18 and Joanne Watkins 74

Leadership and supervision, Metadata curation, Samples and logistics, and Sequencing and analysis, and Software and analysis tools: Samuel Moses 25, 106

Leadership and supervision, Metadata curation, Sequencing and analysis, and Software and analysis tools: Sam Nicholls 41, Matthew Bull 24 and Roberto Amato 116

Leadership and supervision, Project administration, Samples and logistics, and Sequencing and analysis: Darren L Smith 80, 85, 86

Leadership and supervision, Sequencing and analysis, Software and analysis tools, and Visualisation: David M Aanensen 14, 116 and Jeffrey C Barrett 116

Metadata curation, Project administration, Samples and logistics, and Sequencing and analysis: Dinesh Aggarwal 20, 116, 70, James G Shepherd 83, Martin D Curran 71 and Surendra Parmar 71

Metadata curation, Project administration, Sequencing and analysis, and Software and analysis tools: Matthew D Parker 109

Metadata curation, Samples and logistics, Sequencing and analysis, and Software and analysis tools: Catryn Williams 74

Metadata curation, Samples and logistics, Sequencing and analysis, and Visualisation: Sharon Glaysher 68

Metadata curation, Sequencing and analysis, Software and analysis tools, and Visualisation: Anthony P Underwood 14, 116, Matthew Bashton 36, 65, Nicole Pacchiari 74, Katie F Loveson 84 and Matthew Byott 85, 96

Project administration, Sequencing and analysis, Software and analysis tools, and Visualisation: Alessandro M Carabelli 20

Funding acquisition, Leadership and supervision, and Metadata curation: Kate E Templeton 56, 104

Funding acquisition, Leadership and supervision, and Project administration: Thushan I de Silva 109, Dennis Wang 109, Cordelia F Langford 116 and John Sillitoe 116

Funding acquisition, Leadership and supervision, and Samples and logistics: Rory N Gunson 55

Funding acquisition, Leadership and supervision, and Sequencing and analysis: Simon Cottrell 74, Justin O’Grady 75, 102 and Dominic Kwiatkowski 116, 108

Leadership and supervision, Metadata curation, and Project administration: Patrick J Lillie 37

Leadership and supervision, Metadata curation, and Samples and logistics: Nicholas Cortes 33, Nathan Moore 33, Claire Thomas 33, Phillippa J Burns 37, Tabitha W Mahungu 80 and Steven Liggett 86

Leadership and supervision, Metadata curation, and Sequencing and analysis: Angela H Beckett 13, 81 and Matthew TG Holden 73

Leadership and supervision, Project administration, and Samples and logistics: Lisa J Levet 34, Husam Osman 70, 35 and Mohammed O Hassan-Ibrahim 99

Leadership and supervision, Project administration, and Sequencing and analysis: David A Simpson 77

Leadership and supervision, Samples and logistics, and Sequencing and analysis: Meera Chang 72, Ravi K Gupta 102, Alistair C Darby 107 and Steve Paterson 107

Leadership and supervision, Sequencing and analysis, and Software and analysis tools: Oliver G Pybus 23, Erik M Volz 95, Daniela de Angelis 52, David L. Robertson 53, Andrew J Page 22 and Inigo Martincorena 146
Leadership and supervision, Sequencing and analysis, and Visualisation: Louise Agrain 116 and Andrew R Bassett 116

Metadata curation, Project administration, and Samples and logistics: Nick Wong 50, Yusri Taha 89, Michelle J Erkier 99 and Michael H Spencer Chapman 116, 102

Metadata curation, Project administration, and Sequencing and analysis: Rebecca Dewar 56 and Martin P McHugh 56, 111

Metadata curation, Project administration, and Software and tool administration: Siddharth Mookerjee 38, 57

Metadata curation, Project administration, and Visualisation: Stephen Aplin 97, Matthew Harvey 97, Thea Sass 97, Helen Uplemeby 97 and Helen Wheeler 97

Metadata curation, Samples and logistics, and Sequencing and analysis: James P McKenna 3, Ben Warne 3, Joshua F Taylor 22, Yasmin Chaudhry 24, Rhys Izzagbe 24, Aminu S Jahun 24, Gregory R Young 26, 65, Claire McMurray 43, Clare M McCann 65, 66, Andrew Nelson 65, 66 and Scott Elliott 58

Metadata curation, Samples and logistics, and Visualisation: Hannah Lowe 25

Metadata curation, Sequencing and analysis, and Software and analysis tools: Anna Price 11, Matthew R Crown 65, Sara Rey 74, Sunando Roy 96 and Ben Temperton 105

Metadata curation, Sequencing and analysis, and Visualisation: Sharif Shaaban 73 and Andrew R Hesketh 101

Project administration, Samples and logistics, and Sequencing and analysis: Kenneth G Laing 41, Irene M Monahan 41 and Judith Heaney 95, 96, 34

Project administration, Samples and logistics, and Sequencing and analysis: Emanuela Pelosi 71, Siona Silveria 77 and Eleri Wilson-Davies 95

Samples and logistics, Software and analysis tools, and Visualisation: Helen Fryer 5

Sequencing and analysis, Software and analysis tools, and Visualisation: Helen Adams 4, Louis du Plessis 23, Rob Johnson 39, William T Harvey 53, 42, Joseph Hughes 53, Richard J Orton 53, Lewis G Spurgin 53, Yann Bourgeois 53, Chris Ruis 102, Áine O’Toole 101, Martina Gourtovia 101 and Theo Sanderson 146

Funding acquisition, and Leadership and supervision: Christophe Fraser 5, Jonathan Edgeworth 12, Judith Breuer 96, 29, Stephen L Michell 105 and John A Todd 115

Funding acquisition, and Project administration: Michaela John 10 and David Buck 115

Leadership and supervision, and Metadata curation: Kavitha Gaje 37 and Gemma L Kay 78

Leadership and supervision, and Project administration: Sharon J Peacock 26, 70 and David Heyburn 74

Leadership and supervision, and Samples and logistics: Katie Kitchman 3, Alan McLarty 50, 65, 93, David T Pritchard 39, Samir Dervisevic 58, Peter Muiu 70, Esther Robinson 70, 35, Barry B Vipond 109, Newara A Rahman 78, Christopher Jeanes 90, Danni Weldon 116, Jana Catalan 118 and Neil Jones 118

Leadership and supervision, and Sequencing and analysis: Ana da Silva Filipe 56, Chris Williams 57, Marc Fuchs 77, Julia Miskelly 77, Aaron R Jeffries 109, Karen Oliver 116 and Naomi R Park 116

Metadata curation, and Samples and logistics: Amy Ash 1, Cherian Koshy 1, Magdalena Barrow 7, Sarah I. Buchanan 7, Anna Mantzouratou 7, Gemma Clark 15, Christopher W Holmes 16, Sharon Campbell 17, Thomas Davis 21, Nge Keong Tan 22, Julianne R Brown 29, Kathryn A Harris 29, 7, Stephen P Kidd 33, Paul R Grant 34, Li Xu-McCrea 35, Alison Cox 38, 63, Pinglawatee Madona 38, 63, Marcus Pond 38, 63, Paul A Randell 38, 63, Karen T Withell 63, Cheryl Williams 31, Clive Graham 66, Rebecca Denton-Smith 52, Emma Swindells 52, Robyn Turnbull 62, Tim J Sloan 62, Andrew Bosworth 70, 81, Stephanie Hutchings 79, Hannah M Pymont 70, Anna Casey 76, Liz Ratcliffe 79, Christopher R Jones 79, 103, Bridget A Knight 79, 103, Tanzania Haque 103, Jennifer Hart 96, Dianne Irish-Tavares 90, Eric Witte 90, Craig Mower 86, Louisa K Watson 86, Jennifer Collins 89, Gary Eltringham 89, Donn Cradginton 98, Ben Macklin 98, Miren Irlunza-Gomara 107, Anita O Lucaci 107 and Patrick C McClure 113

Metadata curation, and Sequencing and analysis: Matthew Carlile 14, Nadine Holmes 18, Christopher Moore 18, Nathaniel Storey 29, Stefano Rooke 73, Gonzalo Yebra 74, Joel Craine 74, Malorie Perry 74, Nabih-Fareed Alkakh 73, Stephen Bridgett 73, Kate F Cook 84, Christopher Fearn 84, Salmon Goudarzi 85, Ronan A Lyons 86, Thomas Williams 86, Sam T Haldenby 101, Jillian Durham 116 and Steven Leonard 116

Metadata curation, and Software and analysis tools: Robert M Davies 116

Project administration, and Samples and logistics: Rahul Batra 12, Beth Blane 26, Moira J Spyres 56, 93, 96, Perminder Smith 32, 112, Mehmet Yavuz 85, 109, Rachel J Williams 86, Adhyanata I Mahanama 95, Buddhini Samaraweera 86, Sophia T Gergis 102, Samantha E Hansford 109, Angie Green 106, Charlotte Beaver 106, Katherine I Bellis 116, 101, Matthew J Dorman 116, Sally Kay 116, Liam Prestwood 116 and Shavanthi Rajatikela 116

Project administration, and Sequencing and analysis: Joshua Quick 14

Project administration, and Software and analysis tools: Radoslaw Poplawski 14

Samples and logistics, and Sequencing and analysis: Nicola Reynolds 8, Arthur Morris 11, Arthur Morris 11, Thammasalee Muanap 36, 37, Pratul Patel 12, Urooj Islam 24, Myra Hossmilo 24, Malte L Pinckert 24, Joanne Stockton 43, John H Henderson 63, Amy Hollis 65, William Stanley 65, Wen C Yew 65, Richard Myers 72, Alicia Thornton 72, Alexander Adams 74, Tara Annett 74, Hibo Asad 74, Alec Birdsey 74, Jason Coombes 74, Johnathan M Evans 74, Laia Fina 74, Bree Gatica-Whitaker 74, Lauren Gilbert 74, Lee Graham 74, Jessica Hey 74, Emir Hliver 74, Sophie 74, Hannah Jones 74, Sara Kumi-Zumen-Synchroneter 74, Caomh McKeever 74, Jessica Powell 74, Georgia Pugh 74, Sarah Taylor 74, Alexander J Trotter 74, Charlotte A Williams 74, Leanne M Kerrmac 74, Benjamin H Foules 74, Marta Gall 106, Hailey R Horneby 109, Stavroula F Louka 109, Manoj Pohare 109, Paige Wolverson 109, Pojyun Zhang 109, George MacIntyre-Cockett 113, Amy Trebes 115, Robin J Moll 116, Lynne Ferguson 117, Emily J Goldstein 117, Alasdair Maclean 117 and Rachad Tomb 117

Samples and logistics, and Software and analysis tools: Igor Starinskij 53

Sequencing and analysis, and Software and analysis tools: Laura Thomson 5, Joel Southgate 7, 84, Moritz UC Krammer 23, Jayan Raghwani 23, Alex E Zarebski 23, Olivia Boyd 39, Lily Geidelberg 39, Chris J Illingworth 52, Chris Jackson 52, David Pascall 52, Sreenu Vattipally 53, Timothy M Freeman 109, Sharon N Hsu 109, Benjamin B Lindsey 109, Keith James 116, Kevin Lewis 116, Gerry Tonkin-Hill 116 and Jaime M Tovar-Corona 116

Sequencing and analysis, and Visualisation: MacGregor Cox 29

Software and analysis tools, and Visualisation: Khalil Abudahab 14, 116, Mirko Menegazzo 14, Ben EW Taylor MEng 14, 116, Corin A Yeats 14, Afrida Mukaddas 53, Derek W Wright 53, Leonardo de Oliveira Martins 55, Rachel Colquhoun 104, Verity Hill 104, Ben Jackson 104, JT McCrone 104, Nathan Medd 104, Emily Scher 104 and Jon-Paul Keatley 116
College London, 58 Norfolk and Norwich University Hospitals NHS Foundation Trust, 59 Norfolk County Council, 60 North Cumbria Integrated Care NHS Foundation Trust, 61 North Middlesex University Hospital NHS Trust, 62 North Tees and Hartlepool NHS Foundation Trust, 63 North West London Pathology, 64 Northumbria Healthcare NHS Foundation Trust, 65 Northumbria University, 66 NU-OMICS, Northumbria University, 67 Path Links, Northern Lincolnshire and Goole NHS Foundation Trust, 68 Portsmouth Hospitals University NHS Trust, 69 Public Health Agency, Northern Ireland, 70 Public Health England, 71 Public Health England, Cambridge, 72 Public Health England, Colindale, 73 Public Health Scotland, 74 Public Health Wales, 75 Quadram Institute Bioscience, 76 Queen Elizabeth Hospital, Birmingham, 77 Queen’s University Belfast, 78 Royal Brompton and Harefield Hospitals, 79 Royal Devon and Exeter NHS Foundation Trust, 80 Royal Free London NHS Foundation Trust, 81 School of Biological Sciences, University of Portsmouth, 82 School of Health Sciences, University of Southampton, 83 School of Medicine, University of Southampton, 84 School of Pharmacy & Biomedical Sciences, University of Portsmouth, 85 Sheffield Teaching Hospitals NHS Foundation Trust, 86 South Tees Hospitals NHS Foundation Trust, 87 Southwest Pathology Services, 88 Swansea University, 89 The Newcastle upon Tyne Hospitals NHS Foundation Trust, 90 The Queen Elizabeth Hospital King’s Lynn NHS Foundation Trust, 91 The Royal Marsden NHS Foundation Trust, 92 The Royal Wolverhampton NHS Trust, 93 Turnkey Laboratory, University of Birmingham, 94 University College London Division of Infection and Immunity, 95 University College London Hospital Advanced Pathogen Diagnostics Unit, 96 University College London Hospitals NHS Foundation Trust, 97 University Hospital Southampton NHS Foundation Trust, 98 University Hospitals Dorset NHS Foundation Trust, 99 University Hospitals Sussex NHS Foundation Trust, 100 University of Birmingham, 101 University of Brighton, 102 University of Cambridge, 103 University of East Anglia, 104 University of Edinburgh, 105 University of Exeter, 106 University of Kent, 107 University of Liverpool, 108 University of Oxford, 109 University of Sheffield, 110 University of Southampton, 111 University of St Andrews, 112 Viapath, Guy’s and St Thomas’ NHS Foundation Trust, and King’s College Hospital NHS Foundation Trust, 113 Virology, School of Life Sciences, Queens Medical Centre, University of Nottingham, 114 Watford General Hospital, 115 Wellcome Centre for Human Genetics, Nuffield Department of Medicine, University of Oxford, 116 Wellcome Sanger Institute, 117 West of Scotland Specialist Virology Centre, NHS Greater Glasgow and Clyde, 118 Whittington Health NHS Trust.