UK Health Security Agency

SARS-CoV-2 variants of concern and variants under investigation in England

Technical briefing 46

7 October 2022

This report provides an update on previous briefings up to 9 September 2022

Contents

Summary

This report has been published to share the detailed variant surveillance analyses which contribute to the variant risk assessments and designation of new SARS-CoV-2 variants. This specialist technical briefing contains early data and analysis on emerging variants and findings have a high level of uncertainty.

Unless stated otherwise, this technical briefing uses a data cut-off of 3 October 2022 to allow time for analyses.

Interpreting variant data

The current testing policy needs to be considered when interpreting all variant data; the targeting of testing at specific groups is likely to delay the detection and characterisation of variants.

Summary and situational risk assessment

1. The genomic surveillance dataset is currently challenging to interpret due to continued changes in testing as well as proliferation of similar variants.

2. A number of new variants have begun to circulate in the United Kingdom (UK) in recent weeks. The development of these additional lineages has been rapid. They have varying Omicron backbones but some convergent receptor-binding domain (RBD) mutations (notably at S:346) which are likely to produce a degree of escape from current immunity in the UK (MODERATE confidence due to predictive and laboratory data).

3. All new variants are currently at relatively low individual prevalence but at least 3 (BA.2.75.X, BF.7, BQ.X) show evidence of a positive growth rate compared to BA.5 (MODERATE confidence due to growth rates estimated by 2 methods). The magnitude of the growth advantages in some cases is in the same range as other variants which have caused increased incidence, such as other Omicron lineages (LOW confidence as early in the growth trajectory).

4. From UK data, BQ.X, BA.2.75.2 and BF.7 are the most concerning variants in terms of both growth and neutralisation data at present; there is also supportive animal model data for BA.2.75. (LOW confidence due to early and incomplete data on other variants). They will be prioritised for vaccine effectiveness assessment.

5. Overall variants may be contributing to the current increase in coronavirus (COVID-19) incidence (LOW confidence), however given the age mix and the timing of the increase in incidence compared to the variant prevalence, it is likely that other factors are contributing.

6. Note is also made of small numbers of genomes from several recombinant lineages with novel spikes, but there are no current indicators of concerning growth.

Horizon scanning

The following variants are new signals in horizon scanning. Sequence and sample counts reported below are as of 4 October 2022.

BF.7 is a sub-lineage of BA.5.2.1 with the spike mutation R346T. Globally, there are 9,809 sequences identified with Belgium representing the most samples (1,752); there are 663 samples with this lineage in the UK.

BQ.1.1 is a BA.5 sub-lineage which contains the spike mutations N460K, K444T, and R346T. Globally, there are 326 sequences uploaded to GISAID from 20 distinct countries; there are currently 60 UK samples.

BJ.1 is a BA.5 sub-lineage with 13 non-synonymous spike mutations, 7 in RBD and including 4 predicted immune escape locations. Globally, there have been 123 samples uploaded to GISAID from 10 distinct countries, including one English case. The majority of cases are associated with India.

BS.1 is a BA.2.3.2 sub-lineage containing the spike mutations R346T, L452R, N460K, and G476S. Globally, 25 sequences are uploaded to GISAID; 15 uploaded from Japan. There are currently no sequences in the UK for this lineage.

Reports from Variant Technical Group members

Epitope mapping from individuals suffering BA.2 infections is currently being undertaken at Oxford University. This shows human monoclonal antibodies (mAb) binding to distinct patches on the RBD. The RBD mutations in recent variants map exactly to these binding sites and knock out or severely knock down neutralisation activity of potent mAb. This mAb escape corresponds with large falls in neutralisation titre of vaccine immune serum.

Performance of genomic surveillance in characterising variants

Modelling undertaken by the Variant Technical Group in early 2022 tested the effects of a reduction in community SARS-COV-2 testing and sequencing on the detection and characterisation of variants. The study assumed that population level symptomatic polymerase chain reaction (PCR) testing would cease, and testing would be limited to 'test to treat', hospital, healthcare worker and surveillance studies.

The first indication of a potential variant of concern and threat to public health from UK data would be a sustained growth advantage of the new variant compared to the circulating variant. Essential characterisation analyses are severity (measured by hospitalisation) and vaccine effectiveness against hospitalisation.

Five surveillance systems were studied by sub-sampling available UK genomes collected during the emergence of Alpha, Delta and Omicron. Without representative national testing and sequencing and if sampling only 200 genomes per week from the community, lag to detection of a high-confidence growth signal was predicted to take between 22 to 26 days (4 weeks). The lag to detection of 22 to 26 days is specifically relative to national community testing that was in place until April 2022.

UKHSA estimates vaccine effectiveness analysis requires at least 500 and ideally more than 1,000 symptomatic community cases of a new variant to be identified to achieve reasonable estimates. Five hundred known Delta admissions were recorded 43 days after first detection in the UK, although updating of the data systems used to analyse this is likely to have taken at least a further week.

Two methods were used to estimate the numbers of cases needed to detect an association between a new variant and hospital admission in a case control study. If only 200 samples are sequenced per week in the community, it would not be possible to reach 80% power within the 6-week simulation period.

Since April 2022, PCR testing in England is targeted at specific groups and not representative of the community. The timescales have recently updated in view of the current data situation. Experience as evidenced by BA.4/BA.5 risk assessment shows that the detection of growth signal with high certainty might be delayed to 6 to 8 weeks or more. Critical number of cases for initial estimates of growth transmissibility/growth rates from case data), initial data on immune evasion and severity (laboratory data) of a new variant took 6 to 15 weeks. Definitive analysis of the relative risk of admission to hospital following presentation to emergency care took longer than 20 weeks although this was also affected by some data supply issues. Between the weeks beginning 4 April 2022 and 3 October 2022 there have been 238,205 SARS-COV-2 sequences submitted from England that have

passed quality control (number of samples per week: minimum 2,130; average 8,822; median 6,696; maximum 32,604).

Published information on variants

On 1 April 2022 UK Health Security Agency (UKHSA) amended its variant classification system. Further details are available in Technical Briefing 39.

SARS-CoV-2 Routine variant data update covers surveillance data and sequencing coverage data on all other variants of concern (VOCs) and variants under investigation (VUIs) up to 25 March 2022.

The collection page gives content on variants, including prior technical briefings. Technical briefings are published periodically. From Technical Briefing 15, briefings include variant diagnoses identified by whole-genome sequencing and a genotyping polymerase chain reaction (PCR) test, including the categorisation of sequenced and genotyped variant results and a rules-based decision algorithm to identify variant and mutation profiles from genotype assay mutation profiles.

The Public Health England repository from 5 March 2021 contains the previous genomic definitions for VOCs and VUIs.

Part 1. Surveillance overview

World Health Organization (WHO) nomenclature from 24 January 2022 is incorporated. Tables 1a and 1b show the current VOCs, variants (V-date-number), and signals in monitoring detected and not detected in the UK incorporating WHO designations with Phylogenetic Assignment of Named Global Outbreak Lineages (Pangolin lineages).

Table 1a. Variants detected in the UK in the past 12 weeks

Variants of concern	Variants (Vs)	Signals in monitoring
Omicron (B.1.1.529) sub-lineage BA.1 and descendant lineages VOC-21NOV-01	V-22MAY-01 (BA.2.12.1)	BA.3
Omicron (B.1.1.529) sub-lineage BA.2 and descendant lineages VOC-22JAN-01	Delta (B.1.617.2 and sub-lineages) V-21APR-02	Delta and Omicron recombinant lineages (UK), including XBC
Omicron (B.1.1.529) sub-lineage BA.4 VOC-22APR-03	XE Recombinant (BA.1 x BA.2) V-22APR-02	BA.4.7
Omicron (B.1.1.529) sub-lineage BA.5 VOC-22APR-04	V-22JUL-01 (BA.2.75)	BA.2.75.2
	V-22SEP-01 (BA.4.6)	BF.7
		BJ.1
		BQ.1.1
		BQ.1

Table 1b. Variants detected in GISAID, but not in the UK, in the past 12 weeks

Variants of concern	Variants (Vs)	Signals in monitoring
	Alpha (V-20DEC-01/B.1.1.7)	XAW (BA.2 x Delta recombinant)
		BA.2.3.20
		BS.1

VOCs and other variants (V-date-number) are monitored weekly for observations within the last 12 weeks. If variants have not been detected in the UK within this period, they are moved to international status with continued monitoring. If a VOC or variant has not been observed in the UK or international data sets within the preceding 12 weeks, it is designated as provisionally extinct, but monitoring remains in place. Variants and signals in monitoring may also be removed from the grid if they show consistently low growth rates.

1.1 Sequencing coverage

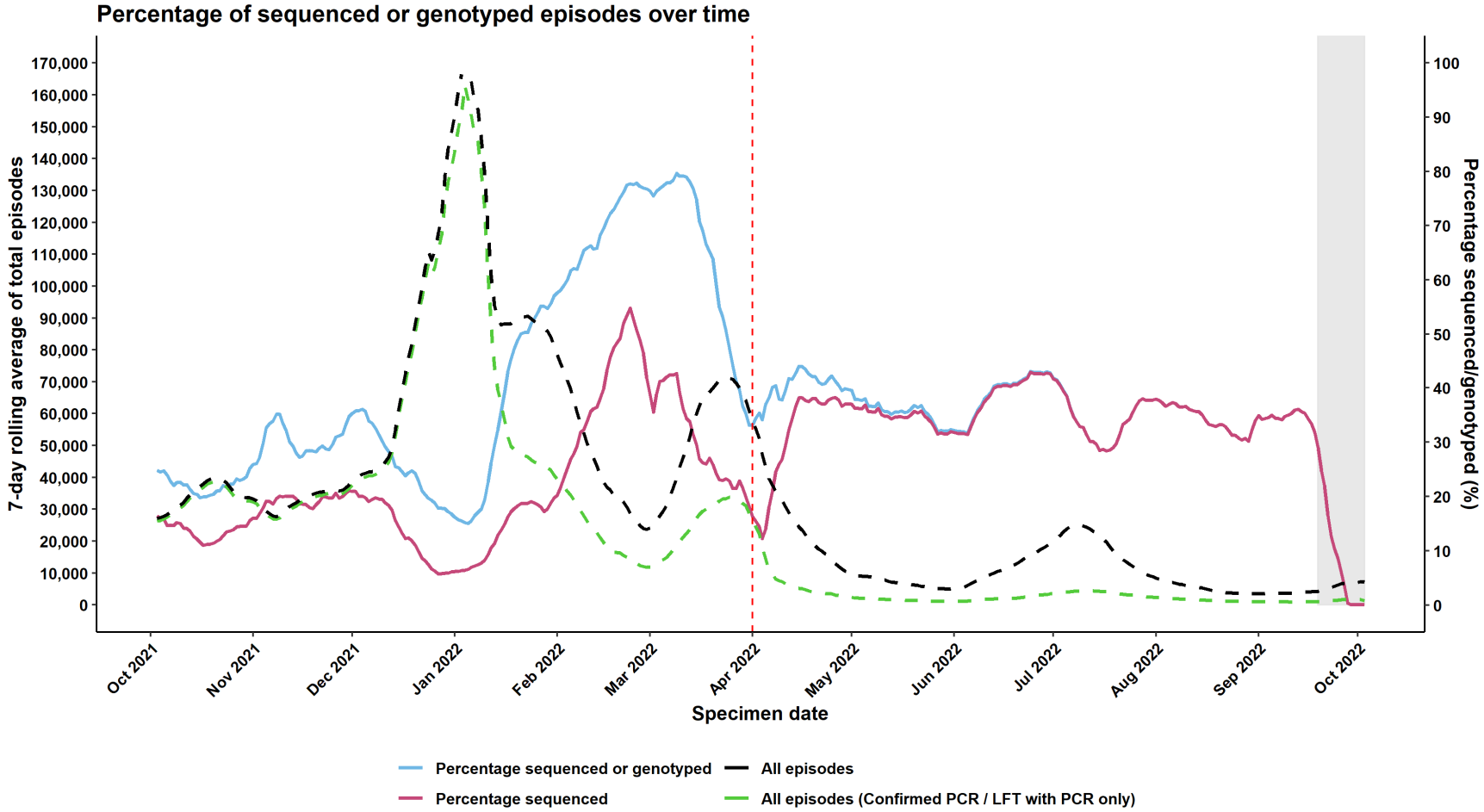
Figure 1 shows the proportion of coronavirus (COVID-19) episodes as detected by PCR that have linked to a valid sequencing result (sequences included have 50% of the genome with sufficient read coverage) or genotyping PCR result over time. Figure 2 shows the proportion of episodes sequenced and genotyped over time by regions. Figure 3 shows the proportion of episodes sequenced and genotyped amongst individuals who tested positive whilst in hospital. The vertical dashed red line indicates the 1 April 2022 when free testing for the general public ended.

Sequencing coverage of PCR confirmed episodes was high during March 2022 (Figure 1). However, this needs to be interpreted with care as PCR tests have declined substantially since mid-February 2022 and case ascertainment is reduced.

Currently, the sequencing strategy prioritises hospitalised cases, patients who are receiving specific antiviral therapy, and national core priority studies.

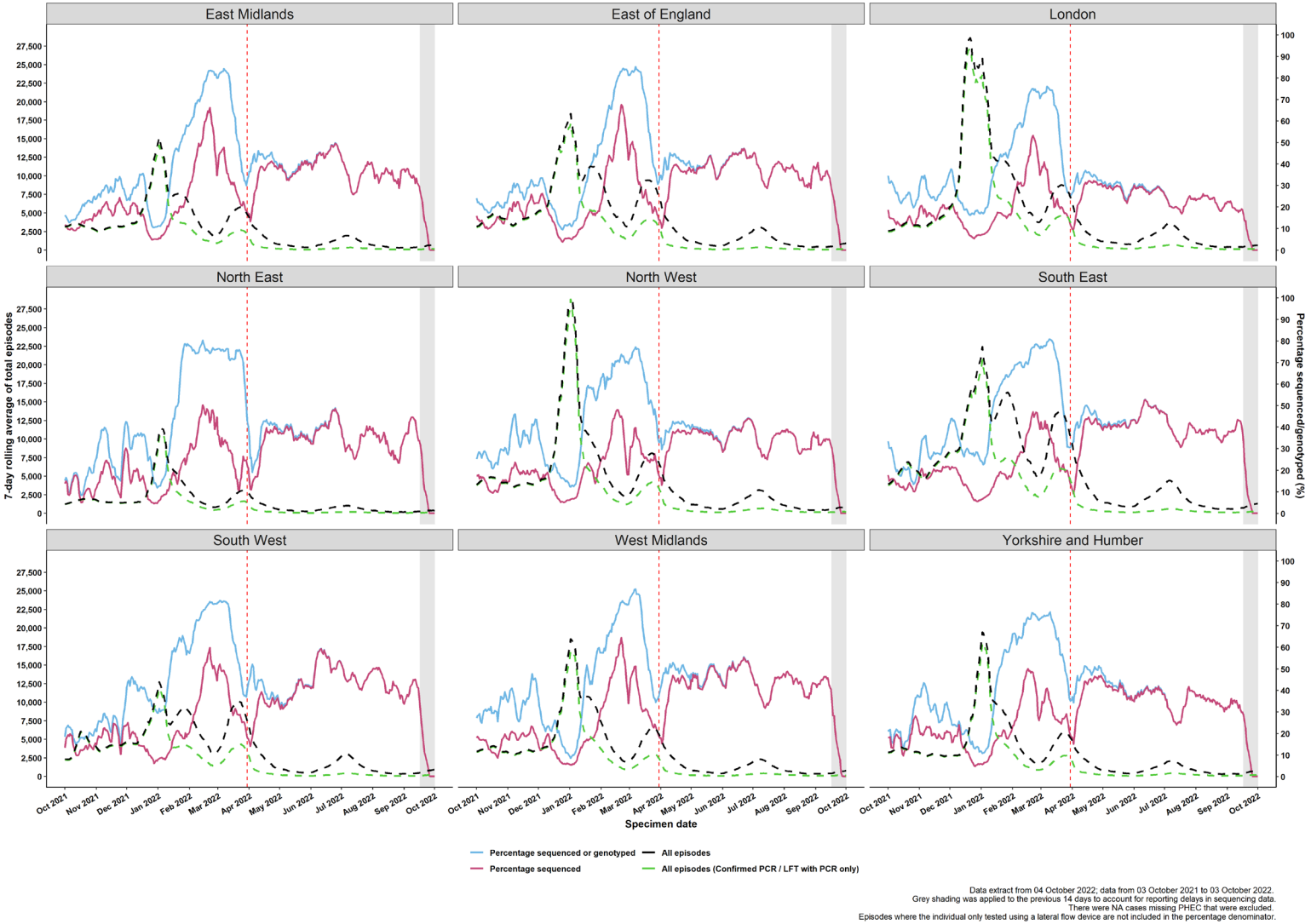
From the week beginning 5 September 2022 up to 2 October 2022, a total of 20,793 SARS-COV-2 sequences have been generated. Of these 2,385 samples are from ONS (11.5%) which is a random sample of community cases and could be used for lineage growth modelling.

Figure 1. Coverage of sequencing with a valid result and genotyping over time (3 October 2021 to 3 October 2022)



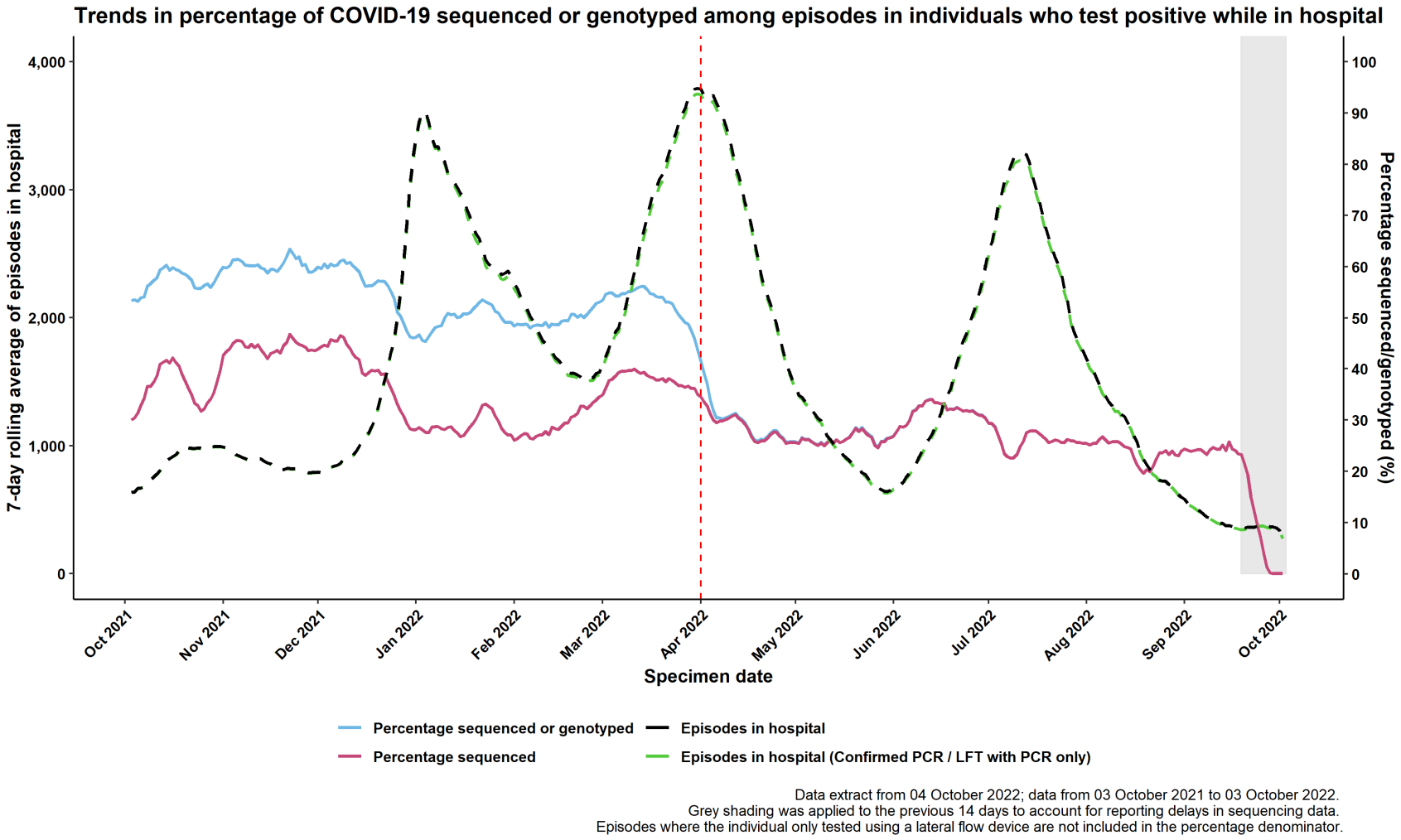
Episodes where the individual only tested using a lateral flow device are excluded. Grey shading was applied to the previous 14 days to account for reporting delays in sequencing data. (The data used in this graph can be found in the accompanying spreadsheet.)

Figure 2. Coverage of sequencing with a valid result and genotyping over time by region (3 October 2021 to 3 October 2022)



Episodes where the individual only tested positive using a lateral flow device are excluded. Grey shading was applied to the previous 14 days to account for reporting delays in sequencing data. (The data used in this graph can be found in the accompanying spreadsheet.)

Figure 3. Coverage of sequencing with valid result and genotyping for cases who test positive in hospital (3 October 2021 to 3 October 2022)



Episodes where the individual only tested positive using a lateral flow device are excluded. Grey shading was applied to the previous 14 days to account for reporting delays in sequencing data. (The data used in this graph can be found in the accompanying spreadsheet.)

1.2 Variant prevalence

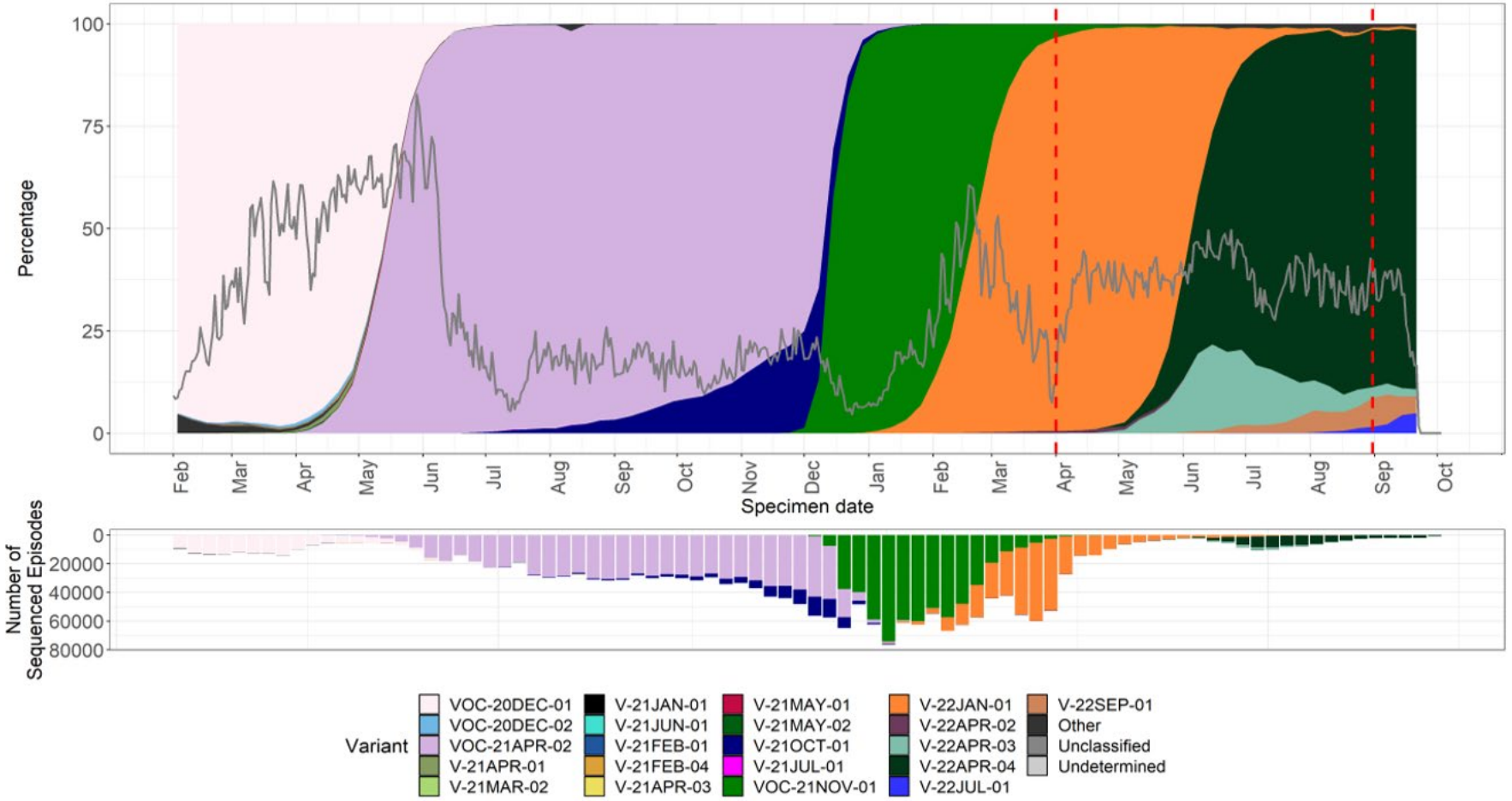
The prevalence of different UKHSA-designated variants amongst sequenced episodes is presented in Figure 4. UKHSA designated variants are those assigned for more comprehensive epidemiological studies and may incorporate multiple sub-lineages.

Of the sequenced episodes from 11 September 2022 to 17 September 2022, 0.6% were BA.2 (VOC-22JAN-01), 2.1% BA.4 (VOC-22APR-03), 87.8% BA.5 (VOC-22APR-04), 4.5% BA.2.75 (V-22JUL-01), 4.5% BA.4.6 (V-22SEP-01) and 0.5% were classified as other.

The prevalence of lineages amongst sequences by Pangolin designation is presented in Figure 5. This provides a greater resolution showing the breakdown of sub-lineages. Lineages are shown if there are more than or equal to 5,000 sequences since 4 April 2022 or if they are more than or equal to 1% of sequences within a single week over the last 6 weeks. Lineages that do not meet these criteria are combined with their parent lineage (for example, BA.2.4 is combine with BA.2).

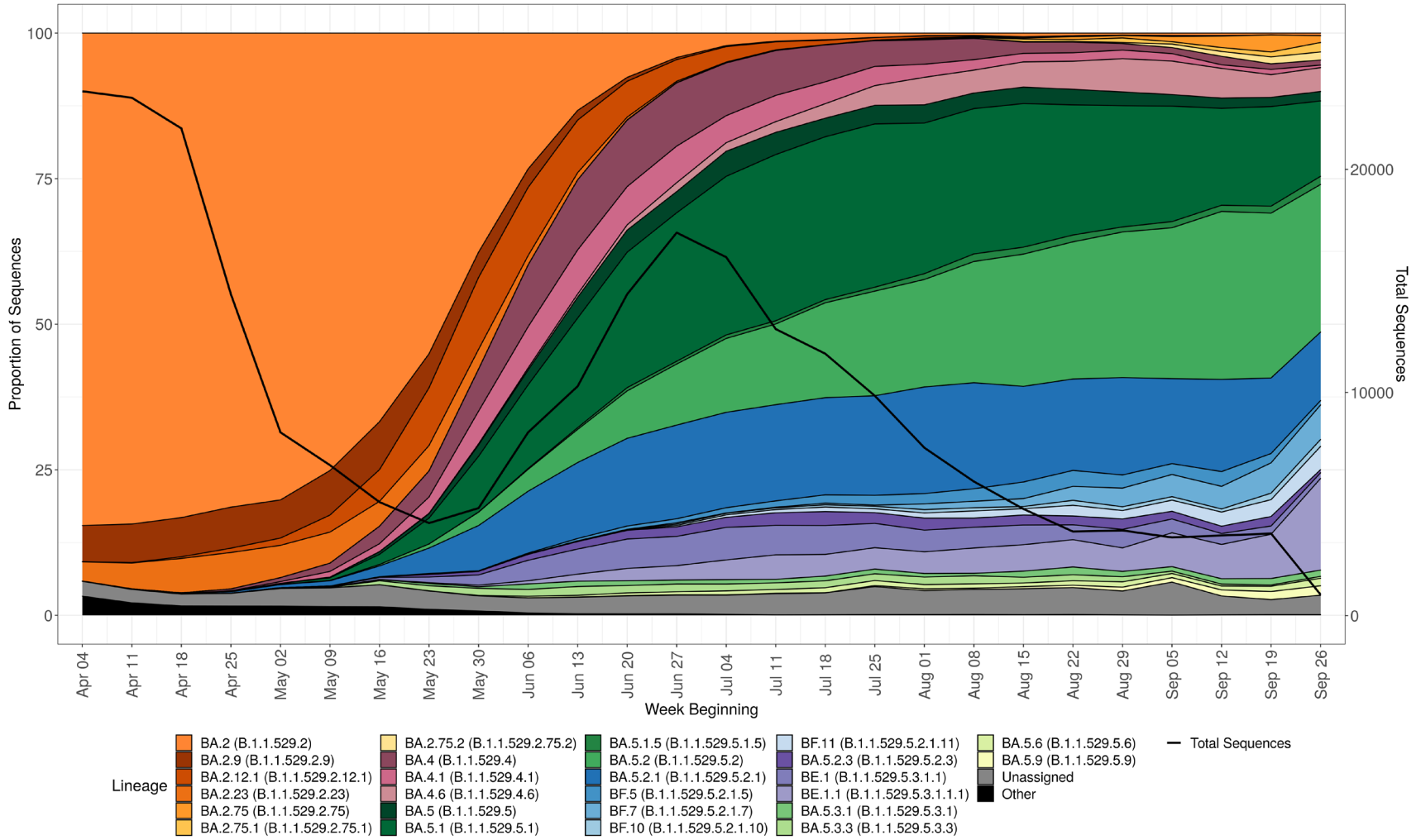
The lineages have been assigned using the accurate Ultrafast Sample placement on Existing tRee (UShER) mode and version 1.14 of the Pangolin data. The UShER mode identifies lineages based on their phylogenetic placement, rather than by specific mutation profiles. This allows sequences with reduced coverage to be assigned to lineages and easier separation of sub-lineages that are distinguished by a small number of mutations. This version does not include recently designated lineages such as BQ.1. The BQ.1 sequences are contained within BE.1.1 in Figure 5 and for the week beginning 12 September 2022 accounts for 23.22% of BE.1.1 sequences (1.37% of all sequences).

Figure 4. Variant prevalence (UKHSA designated variant definitions only) of available sequenced episodes for England from 1 February 2021 as of 4 October 2022



Find accessible data used in this graph in underlying data. Dashed lines indicate period incorporating issue at a sequencing site. Grey line indicates proportion of cases sequenced. The first red dashed line denotes the start of England's 'Living with COVID' Plan at the start of April 2022 and the second indicates the pause of asymptomatic testing for high-risk settings at the end of August 2022. Note recombinants, such as XD, are not specified but are largely within the 'other' group currently as numbers are too small.

Figure 5. Prevalence of Pangolin lineages in the UK with sequence data from 4 April 2022 to 2 October 2022



The total number of valid sequence results per week is shown by the black line. The “Other” category in this plot contains all lineages that do not meet the relevant criteria after combining smaller sub-lineages. “Unassigned" are sequences that could not be assigned a lineage by Pangolin. Find accessible data used in this graph in underlying data.

1.3 Variant modelling

Multiple models are used to estimate the growth advantage of emerging lineages relative to currently circulating lineages. By comparing outputs from multiple models, it is possible to balance strengths and weaknesses of both and provide a more robust perspective on a given lineage's growth. Here we describe lineages using a multinomial model and a logistic regression model.

Variant growth rates were estimated using 3 models in comparison to different background reference data sets (Table 2). The multinomial model (MM) is fitted to Sanger mart sequenced data and growth rates are estimated with respect to a given reference lineage (in this case BA.5). The logistic regression and generalised additive models are fitted with respect to a geographically matched data set reflecting growth with respect to the mixture of lineages co-circulating with a given variant. All reported growth rates are in logistic units and reflect growth in frequency of a given variant, not growth in cases or numbers of samples.

Multinomial model

A Bayesian multinomial model was fit to an English sequenced cases from 1 January 2022 and 25 September 2022, to model the relative fitness advantages of Omicron lineages. The model is fit at the regional level.

The data is sourced from the Sanger mart, where the version of Pangolin used to classify lineages will differ to the in-house UKHSA definitions outlined above. Some lineage classes are combinations of several lineages; any sub-lineage with a small number of samples is folded in with its parent lineage (for example, BA.2 + BA.2.X includes all BA.2 lineages not explicitly modelled). A small percentage (less than 5%) of samples that could not be assigned a probable parent lineage. It should be noted that this model uses Pangolin calls that are subject to change and are not as robust as the UKHSA variant calls from elsewhere in this report. As the Omicron lineage BQ.1 is concerning and not yet correctly classified by Pangolin, the relative growth rate of BE.1.1 samples (the parent lineage of BQ) are presented here from 1 September 2022 onwards.

The modelled percentage representation is shown in Figure 6, with relative growth rates compared in Figure 7. BA.4.6 has a 5.66% (credible interval, CrI: 5.07 to 6.22) larger estimated relative growth rate than BA.5 at a prevalence of 6.59% (CrI: 4.13 to 10.2). BE.1.1 lineages have a 28.82% (CrI: 26.61 to 31.09) larger relative growth rate at an 18.0% (CrI: 11.7 to 26.7) representation in the sample. BA.2.75.X lineages (including BA.2.75.2) is at a lower prevalence of 5.18% (CrI: 3.05 to 8.39) and had a smaller estimated advantage of 18.78% (CrI: 17.10 to 20.52). We had insufficient numbers of BA.2.75.2 to include it separately in the model. BF.7 has a relative advantage of 17.95% (CrI: 16.58 to 19.44) and a prevalence of 7.26% (CrI: 4.53 to 11.1).

Logistic Regression and Generalised Additive Models

The growth rate is estimated by logistic regression of variant of each sample unit on time of sample selection. Growth rates were based on sequences sampled through Pillar 1 testing and from Office for National Statistics (ONS) testing in England. To decorrelate Pillar 1 testing, the data were subsampled so that at most one sequence came from a given combination of hospital, day of sampling, and upper tier local authority (UTLA).

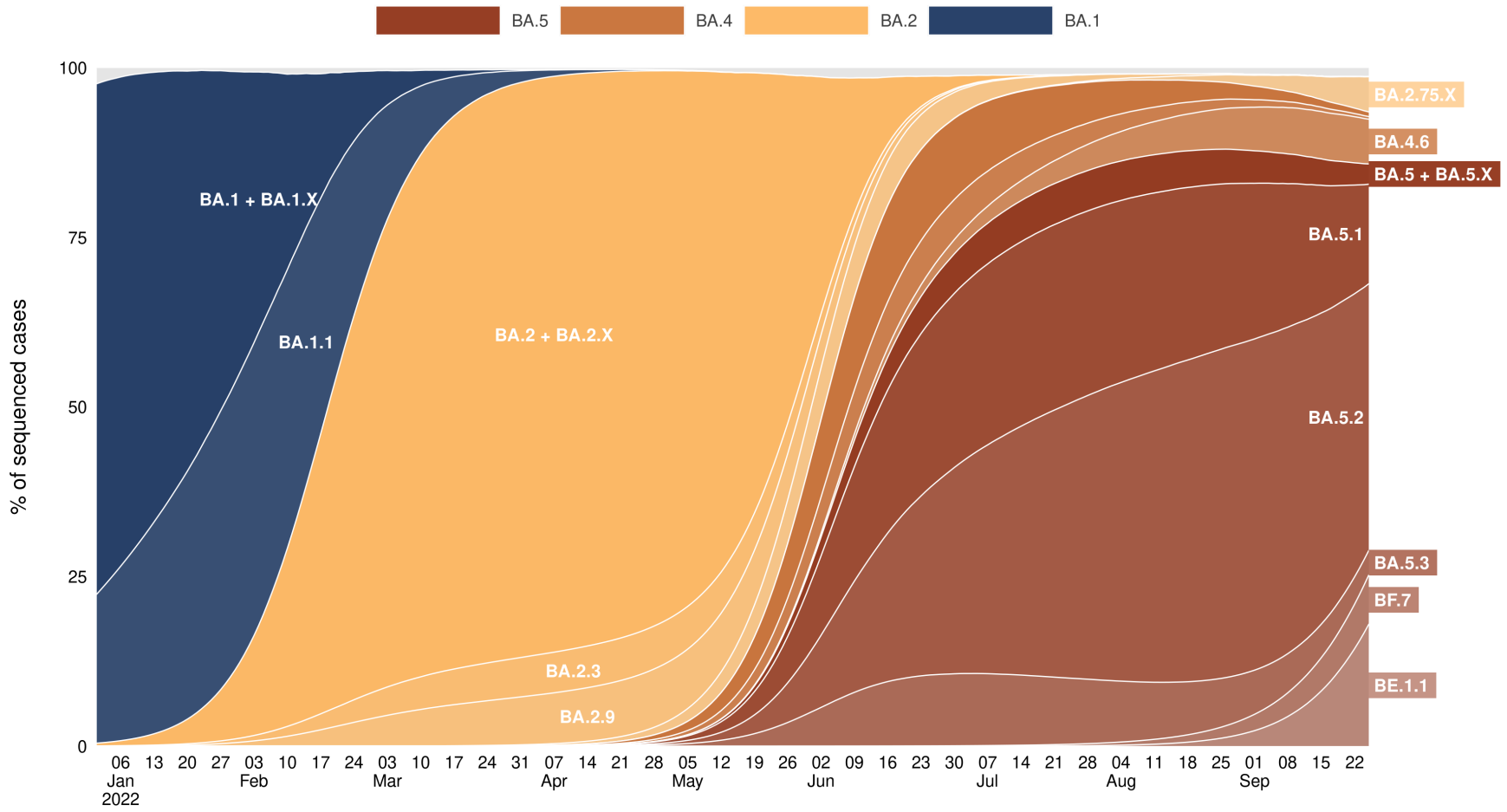
To characterise how growth rates change through time, a generalised additive model is also fitted which allows the growth rate to vary over time. To adjust for geographic variation in case growth rates and differences in sampling intensity, lineage growth rates were estimated relative to a geographically matched sample of genomes. A logistic growth rate of zero would indicate no difference in growth rates between a given lineage and other variants.

Table 2. Growth rate (GR) of variants and signals under monitoring as of 23 September 2022

Lineage	English Sequences used in MM*	MM England estimated prevalence*	MM estimate for the weekly growth relative to BA.5*	English Sequences Counts used in Logistic Regression and Generalised Additive Model	Logistic Regression GR (1/week)	Generalised Additive Model most recent GR (1/week)
BE.1.1 (includes BQ.X)	491	18.0% (CrI: 11.7 to 26.7)	28.82% (CrI: 26.61 to 31.09)	-	-	-
BQ.X	-	-	-	145	60%	34%
BF.7	350	7.26% (CrI: 4.53 to 11.1)	17.95% (CrI: 16.58 to 19.44)	511	32%	-4%
BA.4.6	1,586	6.59% (CrI: 4.13 to 10.2)	5.66% (CrI: 5.07 to 6.22)	3,112	8%	-7%
BA.2.75.X	254	5.18% (CrI: 3.05 to 8.39)	18.78% (CrI: 17.10 to 20.52)	286	41%	40%
BA.2.75.2	-	NA	-	52	-	-

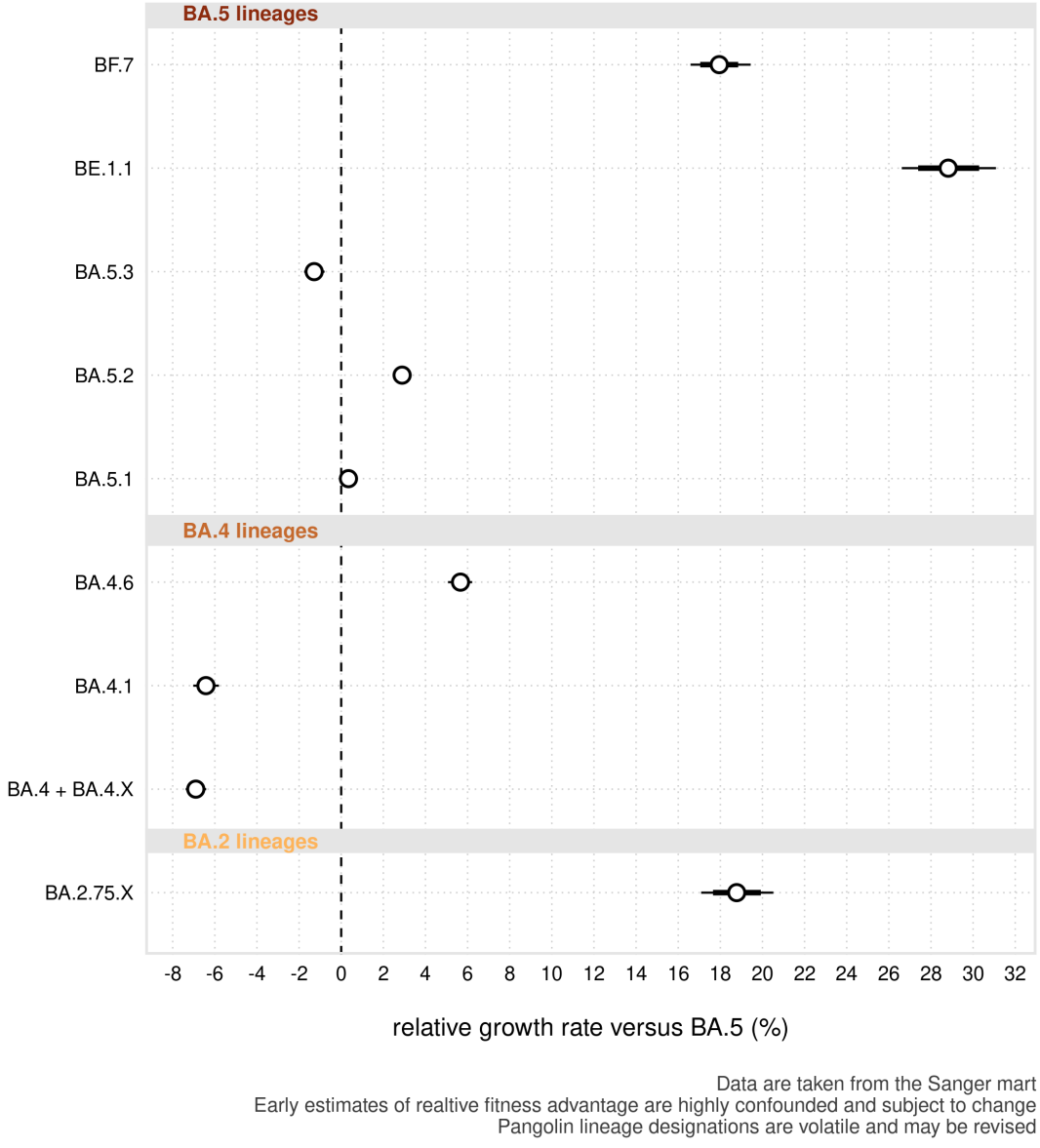
*As of 25 September 2022, for MM estimates

Figure 6. Area plot showing the predicted representation of each lineage of the multinomial model of Sanger mart samples



This figure shows the predicted representation of different lineages from the multinomial model. The grey region denotes other non-Omicron or recombinant lineages. Supplementary data is not available for this figure.

Figure 7. Comparison of the estimated relative growth rates for BA.5 and BA.4 lineages versus that for specifically BA.5 and other low count BA.5 lineages (BA.5 + BA.5.X)



The relative growth rates are taken from a multinomial model of sequenced Pillar 2 cases in England, described above. Supplementary data is not available for this figure.

Part 2. VOC-22APR-04 (BA.5)

Omicron sub-lineage BA.5 was identified as part of horizon scanning on 4 April 2022. On 6 April 2022, the Variant Technical Group classified Omicron sub-lineage BA.5 as V-22APR-04. On 18 May 2022, UKHSA re-classified V-22APR-04 as VOC-22APR-04.

The revised genomic case definition for V-22APR-04 is available in Technical Briefing 41.

2.1 Genomic diversity

Diversity in Spike

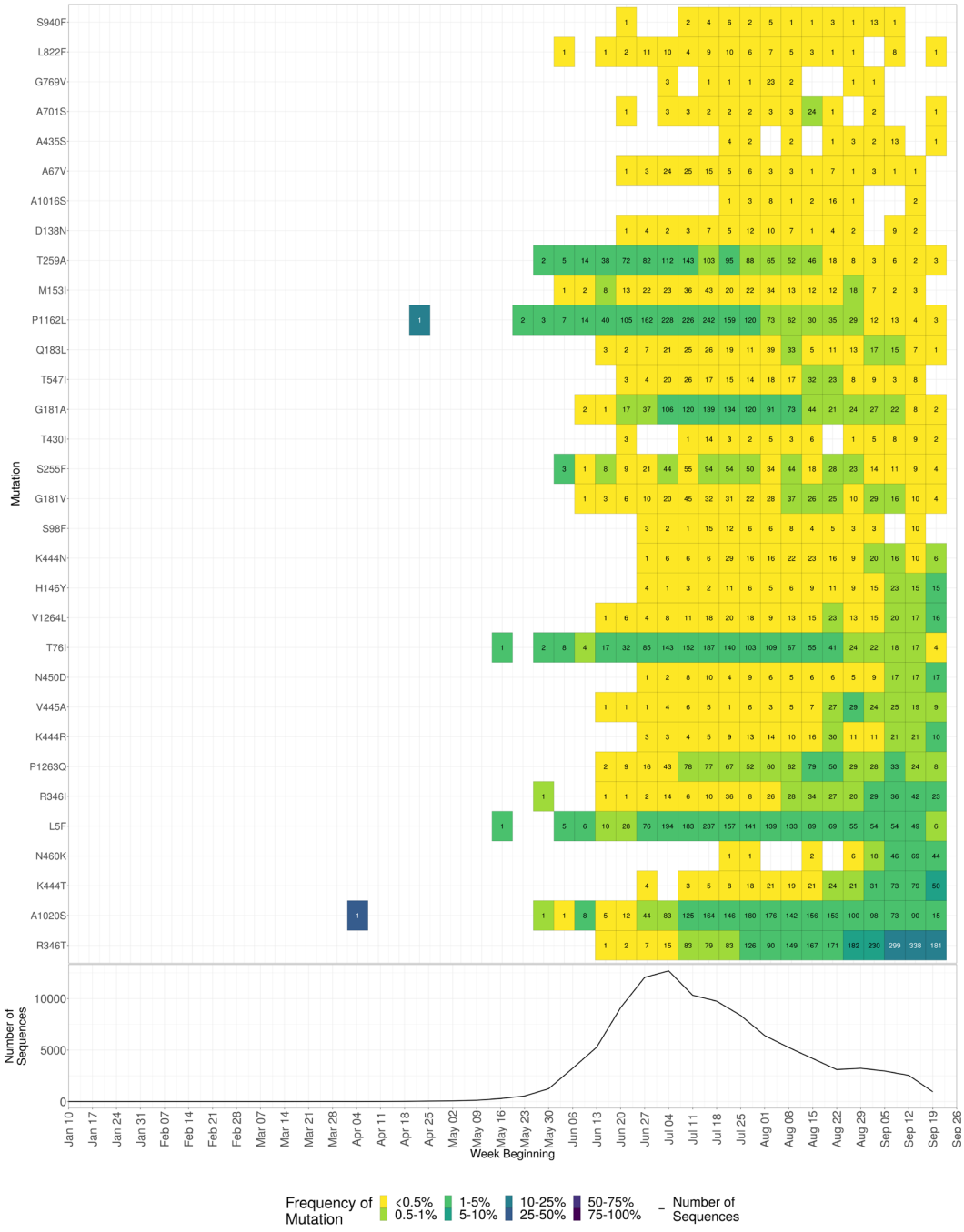
Spike mutations are monitored within VOC-22APR-04 using 4 criteria (Table 4). A mutation is investigated further if it meets more than one of these criteria and is present in at least 10 sequences. Thirty-two additional mutations have been observed in VOC-22APR-04 sequences according to the criteria in Table 4 (Figure 8). The criteria for mutation monitoring are currently being reviewed and amended.

Table 4. Criteria used to assess emerging mutations

Criteria	Threshold
Cumulative count	Running total for the number of sequences containing mutation is at least 50
Proportion	1% of sequences classified as this variant contain this mutation within a single week
Difference in proportion	The difference in the proportion of sequences in 2 consecutive weeks is at least 0.25%
Percentage change in the number of sequences	The percentage change between the number of sequences containing the mutation in 2 consecutive weeks is at least 5%

The sequences included in the plot in Figure 8 are all VOC-22APR-04 sequences; therefore, any BA.5 sub-lineages that meet the definition requirements will be included and the defining mutations for these sub-lineages will be included if they meet the criteria described above. Mutations that are expected to be present in all VOC-22APR-04 sequences (T19I, G142D, V213G, G339D, S371F, S373P, S375F, T376A, D405N, R408S, K417N, N440K, L452R, S477N, T478K, E484A, F486V, Q498R, N501Y, Y505H, D614G, H655Y, N679K, P681H, N764K, D796Y, Q954H, N969K) are not included in the plot but are monitored and any significant changes in the proportions of these mutations (for example a reversion) will be reported as required.

Figure 8. Spike mutations found in VOC-22APR-04 genomes in the UK dataset relative to the Wuhan sequence NC_045512.2 between 10 January 2022 and 25 September 2022

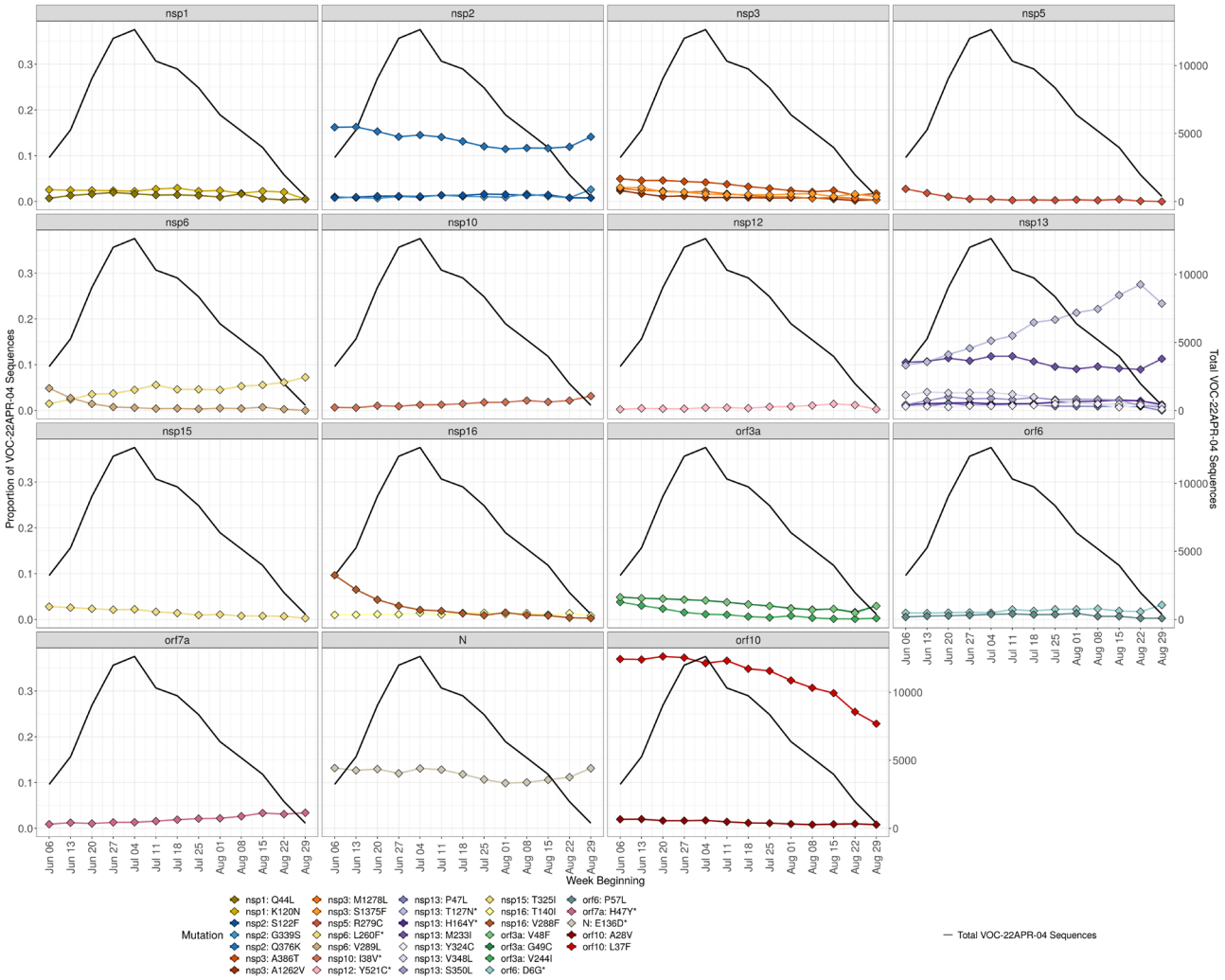


Find accessible data used in this graph in underlying data.

NB: mutations expected to be in all VOC-22APR-04 sequences are not included in this plot to increase readability.

Outside of Spike, there are 36 mutations that are present in at least 1% of VOC-22APR-04 sequences for at least 3 consecutive weeks (Figure 9).

Figure 9. Mutations acquired by VOC-22APR-04 outside Spike, shown as a proportion of total VOC-22APR-04 sequences (1 June 2022 and 25 September 2022)



The total number of VOC-22APR-04 sequences per week are indicated by the black line.

Mutations for each genome are called relative to reference Wuhan NC_045512.2 and acquired mutations are those additional to the ancestral BA.5 mutation set. Those that are considered additional, and that are present in at least 1% of BA.5 sequences for at least 3 consecutive weeks in the UK dataset, are included in Figure 9 as a proportion of total BA.5 sequences. The plot includes all sequences designated as one of the BA.5 sub-lineages that meet the VOC-22APR-04 definition.

Mutations labelled with (*) are those that have been increasing as a proportion of VOC-22APR-04 sequences for at least 3 consecutive weeks within the previous 6 weeks.

Find accessible data used in this graph in underlying data.

Sources and acknowledgments

Data sources

Data used in this investigation is derived from the COG-UK and UKHSA genomic programme data set, ONS COVID-19 Infection Survey, the UKHSA Second Generation Surveillance System, the Secondary Uses Service data set, Emergency Care Data Set and the UKHSA Case and Incident Management System.

Authors of this report

UKHSA Genomics Public Health Analysis Team

UKHSA COVID-19 Vaccines and Epidemiology Team

UKHSA Immunisations Team

UKHSA Surveillance Team

UKHSA Public Health Incident Directors

UKHSA Data, Analytics and Surveillance

UKHSA Infectious Disease Modelling Team

Contributions from the Variant Technical Group

Variant Technical Group members

Person	**Institution**
Meera Chand (Chair)	UKHSA
Genomics and bioinformatics	
Andrew Rambaut	University of Edinburgh
Thomas Peacock	UKHSA / Imperial College London
Matt Holden	Public Health Scotland
Nicholas Loman	UKHSA / University of Birmingham
Richard Myers	UKHSA
Ewan Harrison	Sanger Institute

Person	Institution
Virology and Immunology	
Bassam Hallis	UKHSA
Gavin Screaton	University of Oxford
Lance Turtle	University of Liverpool
Maria Zambon	UKHSA
Ravi Gupta	University of Cambridge
Susanna Dunachie	University of Oxford
Tim Wyatt	Northern Ireland Public Health Agency
Thushan da Silva	University of Sheffield
Wendy Barclay	Imperial College London
Emma Thomson	University of Glasgow / London School of Hygiene and Tropical Medicine
Epidemiology and modelling	
Charlotte Anderson	UKHSA
Chris Williams	Public Health Wales
Daniela de Angelis	University of Cambridge
Erik Volz	UKHSA / Imperial College London
Jamie Lopez-Bernal	UKHSA
John Edmunds	London School of Hygiene and Tropical Medicine
Julia Gog	Scientific Pandemic Influenza Group on Modelling / University of Cambridge
Maria Rossi	Public Health Scotland
Neil Ferguson	Imperial College London
Simon Thelwall	UKHSA
Susan Hopkins	UKHSA
Paula Blomquist	UKHSA
Thomas Finnie	UKHSA
Thomas Ward	UKHSA
International epidemiology	
Chris Lewis	Foreign, Commonwealth and Development Office
Katherine Russell	UKHSA
Leena Inamdar	UKHSA

Acknowledgements

The authors are grateful to those teams and groups providing data for these analyses including: the Lighthouse Laboratories, National Health Service, COG-UK, the Wellcome Sanger Institute, Health Protection Data Science teams, the University of Oxford, the Genotype to Phenotype Consortium, Medical Research Council Biostatistics Unit, Cambridge and Imperial College, London.

About the UK Health Security Agency

UKHSA is responsible for protecting every member of every community from the impact of infectious diseases, chemical, biological, radiological and nuclear incidents and other health threats. We provide intellectual, scientific and operational leadership at national and local level, as well as on the global stage, to make the nation heath secure.

UKHSA is an executive agency, sponsored by the Department of Health and Social Care.

Version 1.0

Published: October 2022
Publishing reference: GOV-13388