UK Health Security Agency

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

## Technical briefing 47

28 October 2022

This report provides an update on previous briefings up to 7 October 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 24 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.

## Situational risk assessment

Whilst there are substantial numbers of genomes still being generated, the demographic composition of the cases sequenced is different from total cases in the population with a greater representation of older individuals. This may affect variant characterisation.

During the small peak that has just been reached, the variants present were BA.5 sub-lineages; receptor-binding domain (RBD) mutations were already present and may have been contributing (low confidence).

There is the potential for BQ.1 (and sub-lineages), XBB and other similar lineages to cause a further near-term increase in transmission. This is based on the growth advantage shown by BQ.1 and sub-lineages in the UK based on more than one model, the international growth of XBB and the available predictive and laboratory data on antigenic change (moderate confidence).

Further Omicron derived variants may continue to contribute to transmission as there is still potential for the RBD to evolve on the current background (moderate confidence).

## Sequencing coverage

Analysis of the mean proportion of sequenced samples originating from different sample groups shows that from the week commencing 18 July 2022 to 3 October 2022, 14.6% of samples were from the Office of National Statistics (ONS) COVID-19 Infection Survey, 0.7%

from VIVALDI (care home study) and 8.0% from AVA (individuals tested because they may be eligible for therapeutics in the community).

Between 26 September to 24 October 2022, the median age of reported COVID-19 cases was 55 years old. However, during the same period the median age of sequenced COVID-19 cases was 73 years old.

# Horizon scanning

The following describes updates to signals/designated variants in horizon scanning. Sequence and sample counts reported below are as of 24 October 2022.

BQ.1 and XBB have been raised from signal in monitoring to designated variant, V-22OCT-01 and V-22OCT-02 respectively.

BQ.1 (V-22OCT-01) is a BA.5 sub-lineage with spike mutations L452R, N460K, and K444T. It has been designated as a variant on the basis of rapid growth. There are 2,490 non-UK BQ.1 sequences on GISAID and 3,207 UK sequences. In the UK, BQ.1 has a logistic regression growth rate of 49.6%.

XBB (V-22OCT-02) is a recombinant lineage derived from BA.2 parent lineages BJ.1 and BM.1.1.1. XBB is characterised by the acquisition of E: T11A, Spike: V83A, H146Q, Q183E, F486S, F490S. It has been designated a variant based on the significant number of RBD mutations in the spike gene and rapid growth in Singapore. Currently there are 1,086 international samples in GISAID (of which 639 are from Singapore) including 18 UK samples.

BF.7 is a sub-lineage of BA.5.2.1 with the spike mutation R346T. BF.7 is a signal in monitoring due to the presence of the R346T mutation in a BA.5 lineage. Globally, there are 11,922 non-UK sequences identified. BF.7 in Belgium and the Netherlands exceeded 20% of their weekly GISAID sequence uploads; there are 2,644 BF.7 samples in the UK.

BQ.1.1 is a BA.5 sub-lineage which contains the spike mutations N460K, K444T, and R346T. This lineage has been raised as a signal in monitoring due to the presence of multiple mutations in the RBD region of the spike gene, in addition to rapid growth rates in multiple western countries, including the UK. Globally, there are 2,304 non-UK sequences uploaded to GISAID from 35 distinct countries; there are currently 1,272 UK samples.

BS.1 is a BA.2.3.2 sub-lineage containing the spike mutations R346T, L452R, N460K, and G476S. This lineage has been raised as a signal in monitoring due to the presence of multiple mutations in the RBD region of the spike gene. Due to the limited geographic spread and growth this lineage has not been designated as a variant. Globally, 79 sequences are

uploaded to GISAID; 35 uploaded from Japan. As of 24 October 2022, 2 sequences from the UK have been uploaded to the GISAID database for this lineage.

BJ.1 has been de-escalated from a signal in monitoring due to reduced incidence.

BA.3 has been de-escalated from a signal in monitoring due to low growth rate.

# Growth rate

BA.5, including all sub-lineages, remains the dominant parent lineage in the UK at greater than 75% of all sequenced samples in the UK.

In the most recent week, logistic growth of variants with 1, 2 or 3 convergent and antigenically significant RBD mutations was respectively 23%, 47%, and 66% per week. The category with 3 RBD mutations consisted largely of BQ.1.1 (59%) with the remainder consisting primarily of a mixture of BA.2.75 sub-lineages (29%).

# Reports from Variant Technical Group members

Oxford University report neutralisation studies on serum collected 28 days following a third dose of Pfizer BNT162b2 vaccine, and in vaccinated cases infected with BA.1, BA.2 and BA.4/5. Data show significant reductions in neutralisation titres against BA.2.75.2, BA.2.3.20 and BJ.1, compared to BA.2 and BA.4/5, which may suggest they have been selected to escape pre-existing immunity to earlier waves of Omicron infection.

# 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 | Designated Variants (Vs) | Signals in monitoring |
|---|---|---|
| Omicron (B.1.1.529) sub-lineage BA.1 and descendant lineages<br>VOC-21NOV-01 | Omicron BA.2.12.1<br>V-22MAY-01 | Delta and Omicron recombinant XBC |
| Omicron (B.1.1.529) sub-lineage BA.2 and descendant lineages<br>VOC-22JAN-01 | Delta (B.1.617.2 and sub-lineages)<br>V-21APR-02 | BA.4.7 |
| Omicron (B.1.1.529) sub-lineage BA.4<br>VOC-22APR-03 | Omicron XE Recombinant (BA.1 x BA.2)<br>V-22APR-02 | BA.2.75.2 |
| Omicron (B.1.1.529) sub-lineage BA.5<br>VOC-22APR-04 | Omicron BA.2.75<br>V-22JUL-01 | BF.7 (BA.5.2.1) |
| | Omicron BA.4.6<br>V-22SEP-01 | BQ.1.1 |
| | † Omicron BQ.1<br>V-22OCT-01 | BS.1 (BA.2.3.2.1) |
| | † Omicron XBB Recombinant<br>V-22OCT-02 | BA.2.3.20 |

†Newly escalated variants/signals in monitoring since the previous technical briefing.

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

| Variants of concern | Designated Variants (Vs) | Signals in monitoring |
|---|---|---|
| | | Omicron and Delta recombinant XAW |

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) cases 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 cases sequenced and genotyped over time by regions. Figure 3 shows the proportion of cases sequenced and genotyped amongst individuals who tested positive whilst in hospital. The data on people who tested positive whilst in hospital is derived from the Hospital-Onset COVID-19 data set (HO-COVID), which links confirmed COVID-19 episodes to admissions data from the Emergency Care Data Set (ECDS) and Secondary Uses Service (SUS) as provided by NHS Digital (Bhattacharya and others 2021). The vertical dashed red line indicates the 1 April 2022 when free testing for the general public ended.

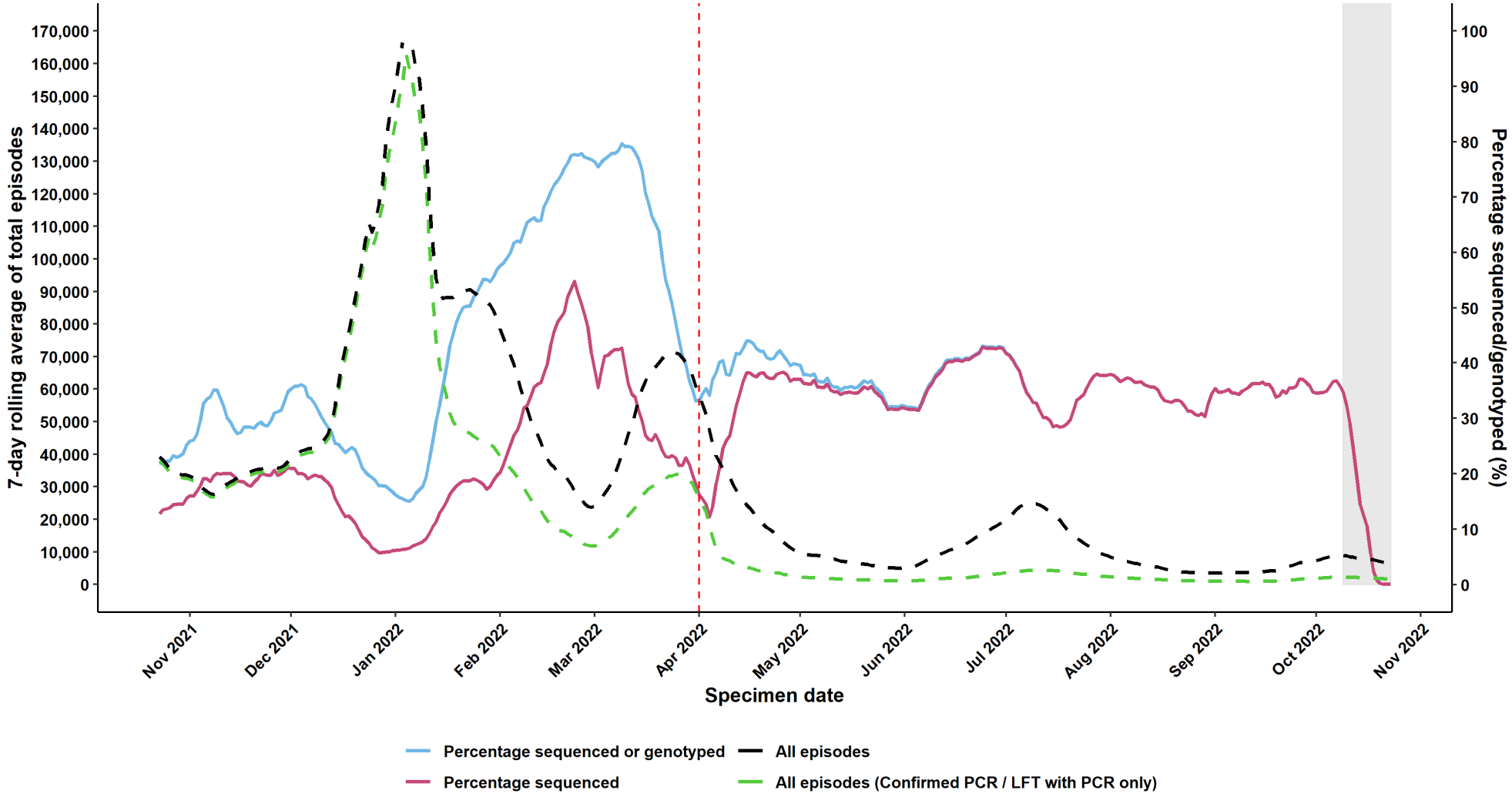
Sequencing coverage of PCR confirmed cases was high during March 2022 (Figure 1). However, this needs to be interpreted with care as PCR tests have declined substantially since mid-February 2022 and case ascertainment is reduced.

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

Figure 4 shows the proportion of sequenced samples that originate from different sample groups, assigned as “AVA” (individuals tested because they may be eligible for therapeutics in the community), “VIVALDI” (care homes study) and “ONS” (Office of National Statistics COVID-19 Infection Survey) over time. Samples not assigned to a specific group are denoted as “other” and will include samples submitted from NHS and UKHSA routine laboratories and routine testing from care homes. Over the past 12 complete weeks of data (week commencing 18 July 2022 to 3 October 2022), mean proportions were 14.6% ONS, 0.7% VIVALDI and 8.0% AVA.

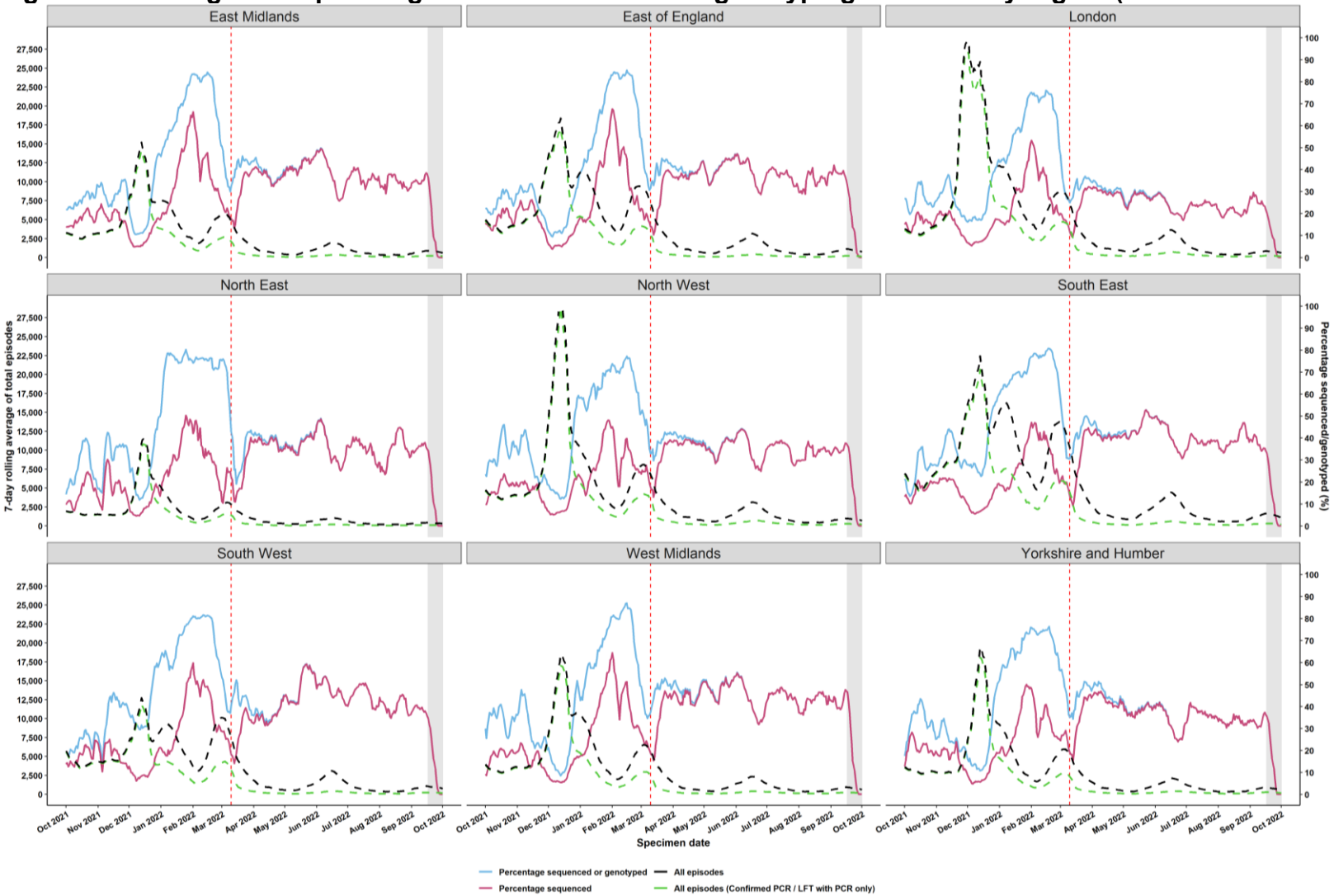
From the week beginning 3 October 2022 up to 23 October 2022, a total of 21,358 SARS-COV-2 sequences have been generated. Of these 4,233 samples are from ONS (19.8%) which is a random sample of community cases and could be used for lineage growth modelling.

**Figure 1. Coverage of sequenced cases with a valid result and genotyping over time (23 October 2021 to 23 October 2022)**



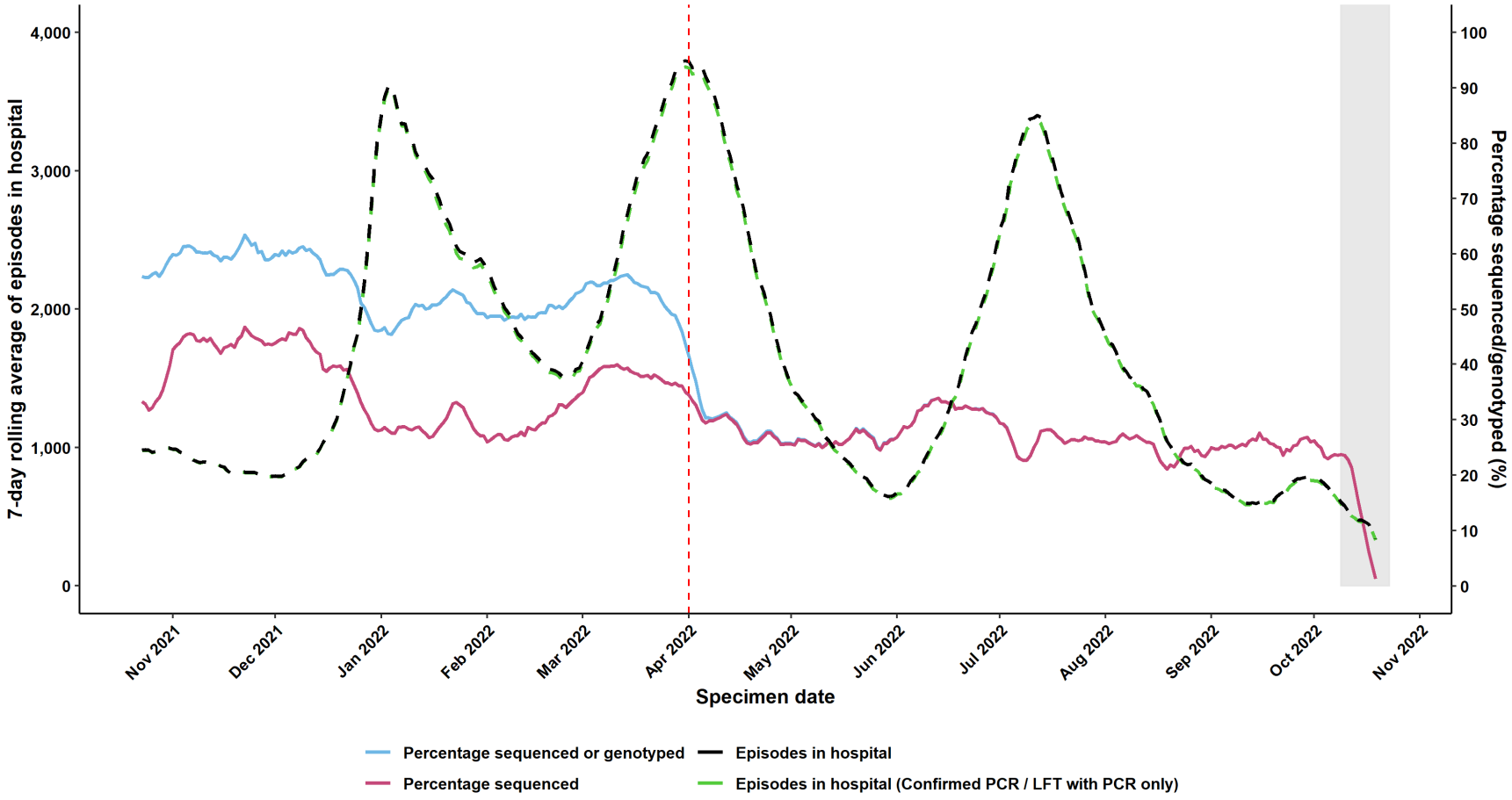

Cases where the individual only tested using a lateral flow device are excluded. Grey shading was applied to the previous 14 days to account for reporting delays in sequencing data. (The data used in this graph can be found in the accompanying spreadsheet.)

**Figure 2. Coverage of sequencing with a valid result and genotyping over time by region (23 October 2021 to 23 October 2022)**



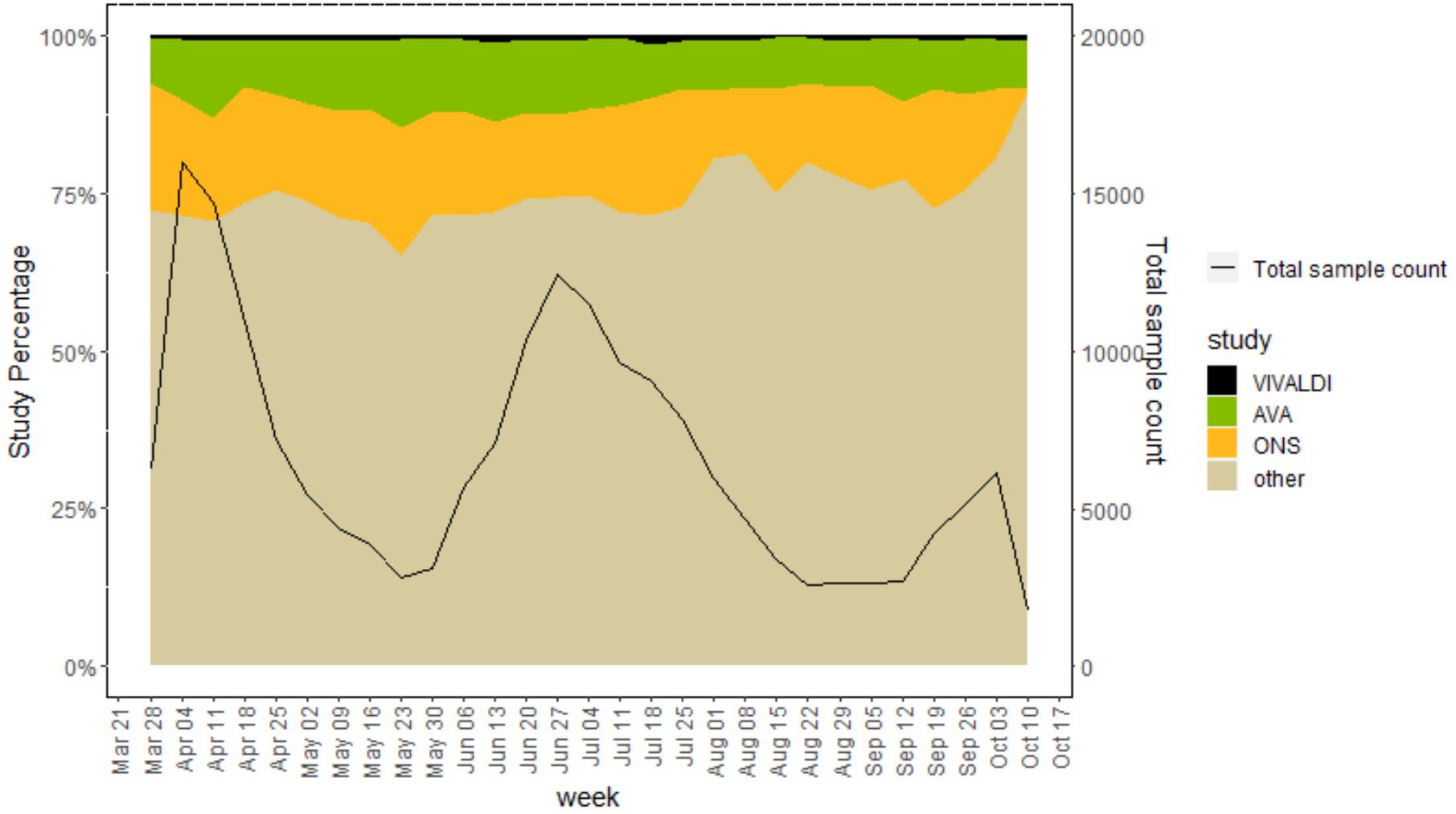

Cases where the individual only tested positive using a lateral flow device are excluded. Grey shading was applied to the previous 14 days to account for reporting delays in sequencing data. (The data used in this graph can be found in the accompanying spreadsheet.)

**Figure 3. Coverage of sequencing with valid result and genotyping for cases who test positive in hospital (23 October 2021 to 23 October 2022)**




Cases 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 4. Proportion of sequenced sample assigned to studies AVA (community therapeutics), VIVALDI (care home study) and ONS (COVID-19 Infection Survey) from 28 March 2022 to 10 October 2022**




Other category denotes all sequenced samples that are not assigned to one of the listed studies (ONS, AVA, VIVALDI)
The data used in this graph can be found in the accompanying spreadsheet.

Due to prioritisation of samples for sequencing from hospitalised patients and care homes, sequenced cases are significantly older than reported cases. Between 26 September to 24 October 2022, the median age of reported COVID-19 cases was 55 years old. However, during the same period the median age of sequenced COVID-19 cases was 73 years old (Figures 5A and 5B).

**Figure 5A. Age-sex distribution of all COVID-19 cases for the past 4 weeks (26 September 2022 to 24 October 2022)**

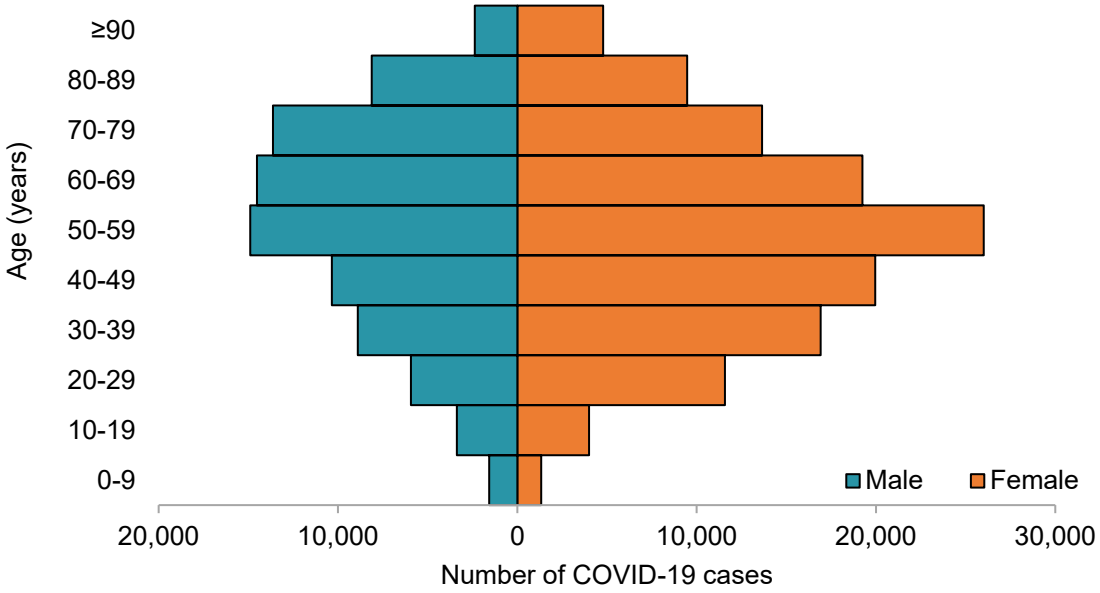
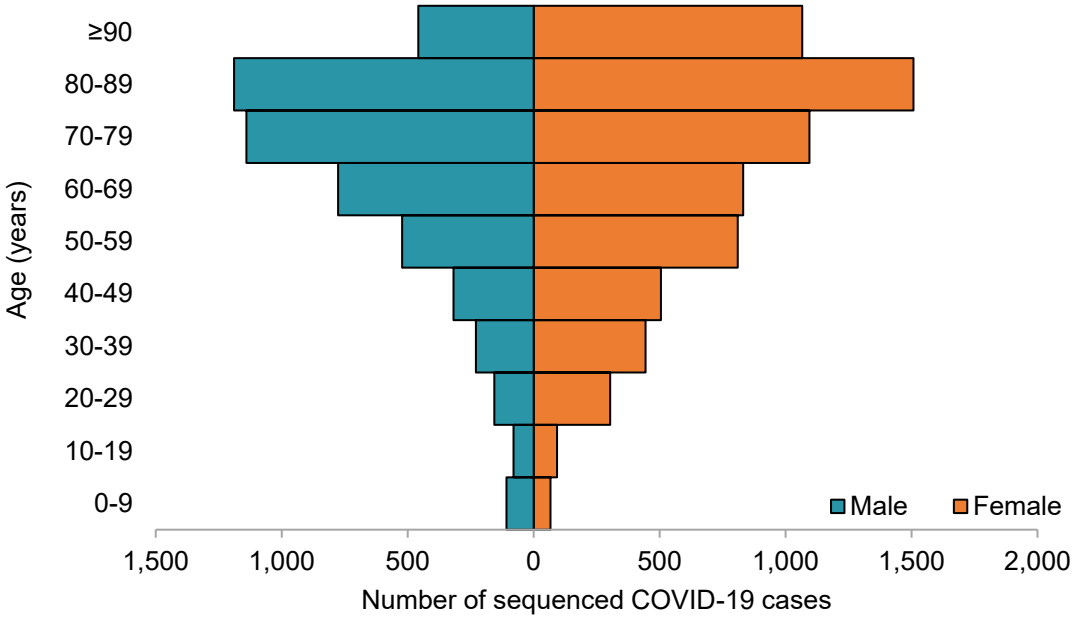



**Figure 5B. Age-sex distribution of sequenced COVID-19 cases for the past 4 weeks (26 September 2022 to 24 October 2022)**




The data used in this graph can be found in the accompanying spreadsheet.

# 1.2 Variant prevalence

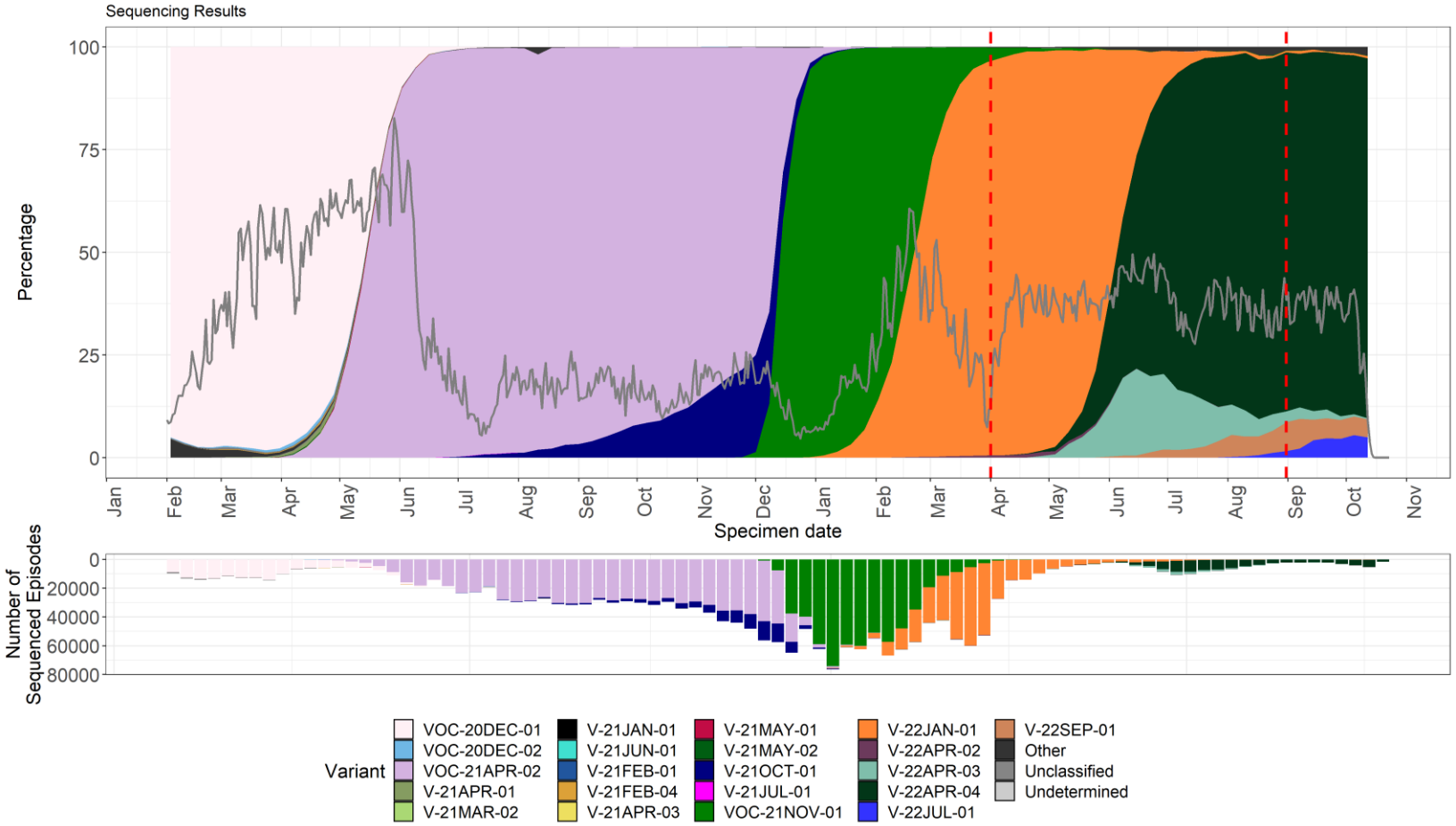
The prevalence of different UKHSA-designated variants amongst sequenced cases is presented in Figure 6. UKHSA designated variants are those assigned for more comprehensive epidemiological studies and may incorporate multiple sub-lineages.

Of the sequenced cases from 2 October 2022 to 8 October 2022, 0.5% were BA.2 (VOC-22JAN-01), 0.6% BA.4 (VOC-22APR-03), 87.4% BA.5 (VOC-22APR-04), 5.5% BA.2.75 (V-22JUL-01), 4.5% BA.4.6 (V-22SEP-01) and 1.5% were classified as other.

The prevalence of lineages amongst sequences by Pangolin designation is presented in Figure 7. 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). Figure 8 shows the prevalence of lineages within the ONS sequence data only.

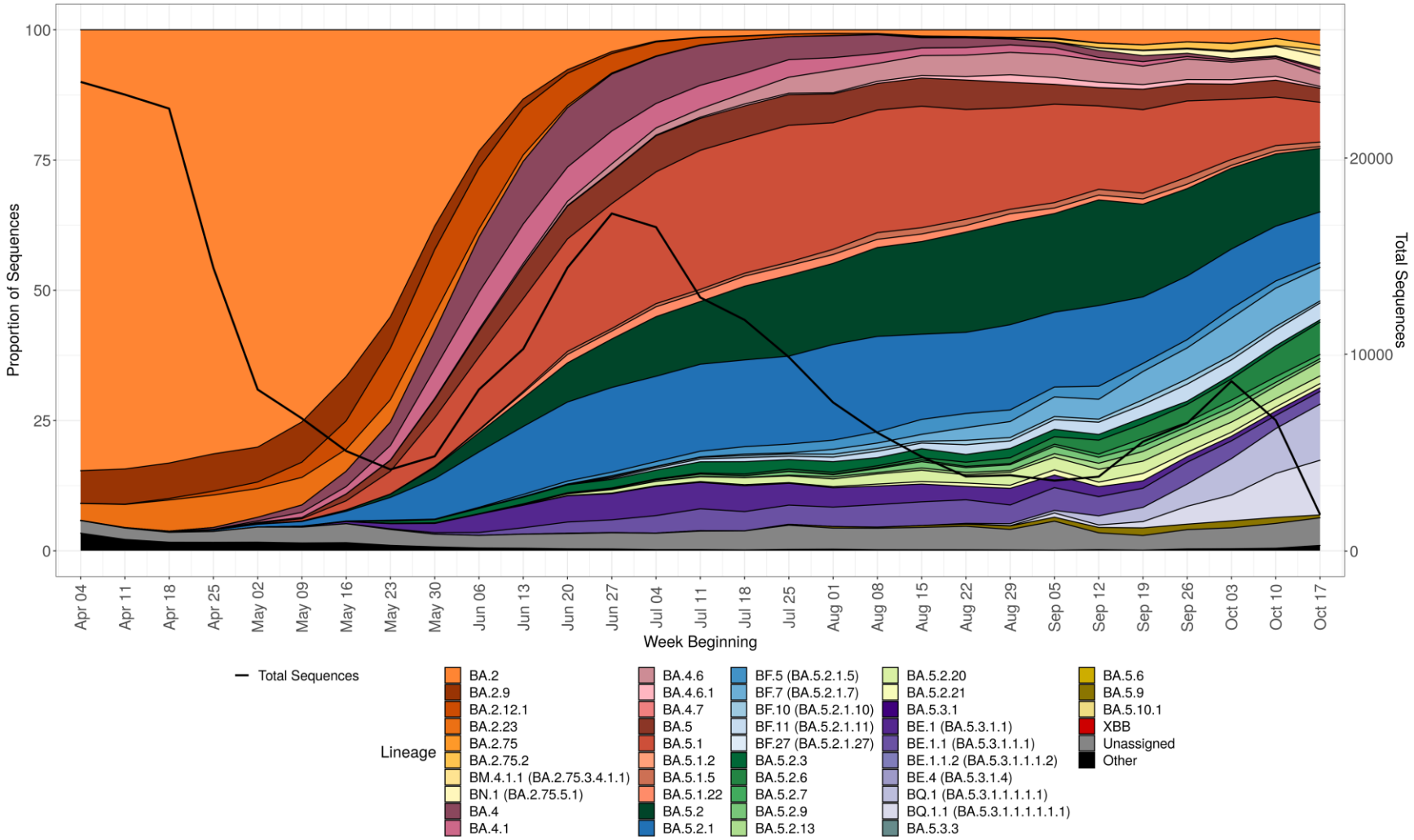
The lineages have been assigned using the accurate Ultrafast Sample placement on Existing tRee (UShER) mode and version 1.15 of the Pangolin data. The UShER mode identifies lineages based on their phylogenetic placement, rather than by specific mutation profiles. This allows sequences with reduced coverage to be assigned to lineages and easier separation of sub-lineages that are distinguished by a small number of mutations.

**Figure 6. Variant prevalence (UKHSA designated variant definitions only) of available sequenced cases for England from 1 February 2021 as of 24 October 2022**



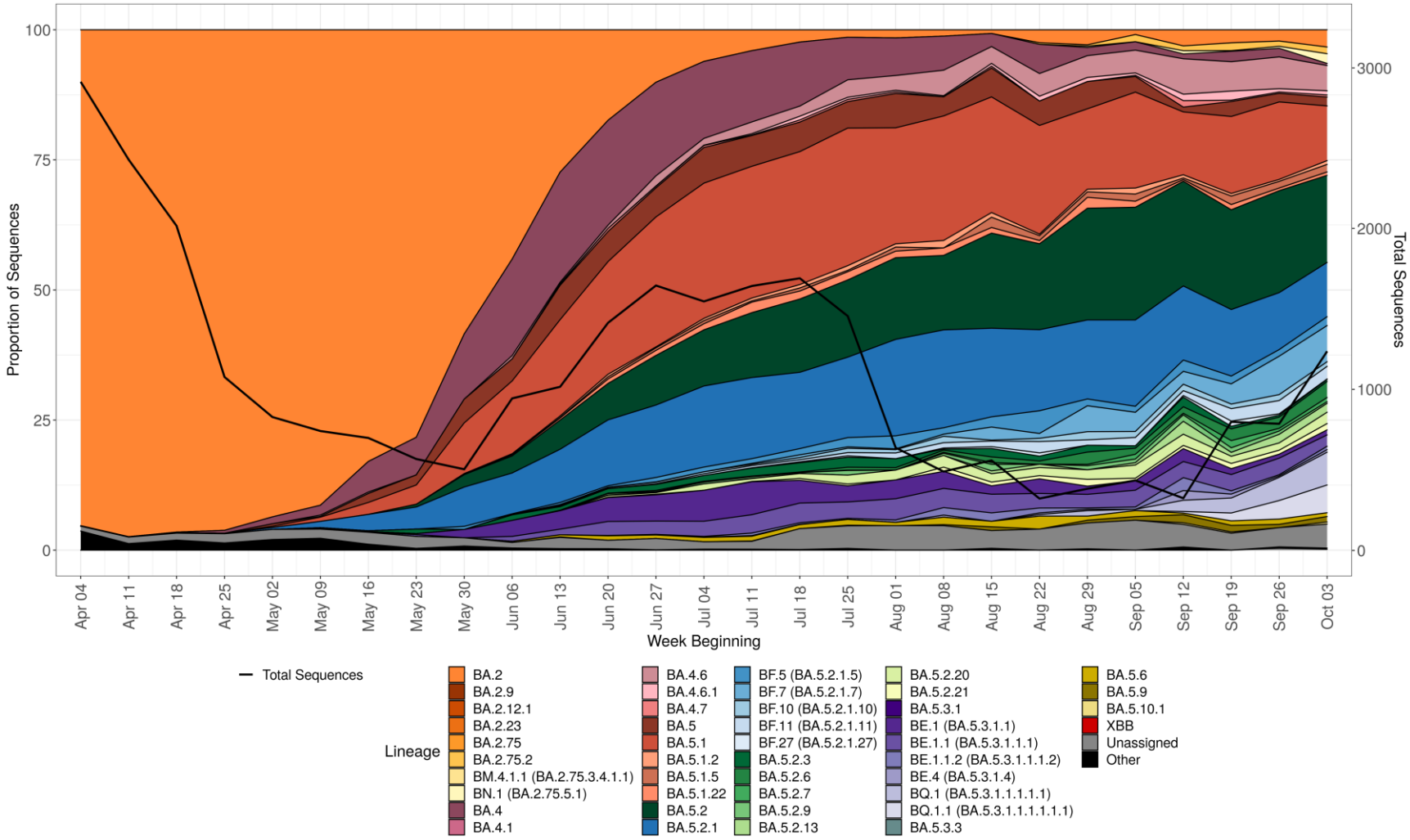

Find accessible data used in this graph in underlying data. Dashed lines indicate period incorporating issue at a sequencing site. Grey line indicates proportion of cases sequenced. The first red dashed line denotes the start of England's 'Living with COVID-19' Plan at the start of April 2022 and the second indicates the pause of asymptomatic testing for high-risk settings at the end of August 2022.

**Figure 7. Prevalence of Pangolin lineages in the UK with sequence data from 4 April 2022 to 23 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.

**Figure 8. Prevalence of Pangolin lineages in the UK ONS sequence data from 4 April 2022 to 9 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.

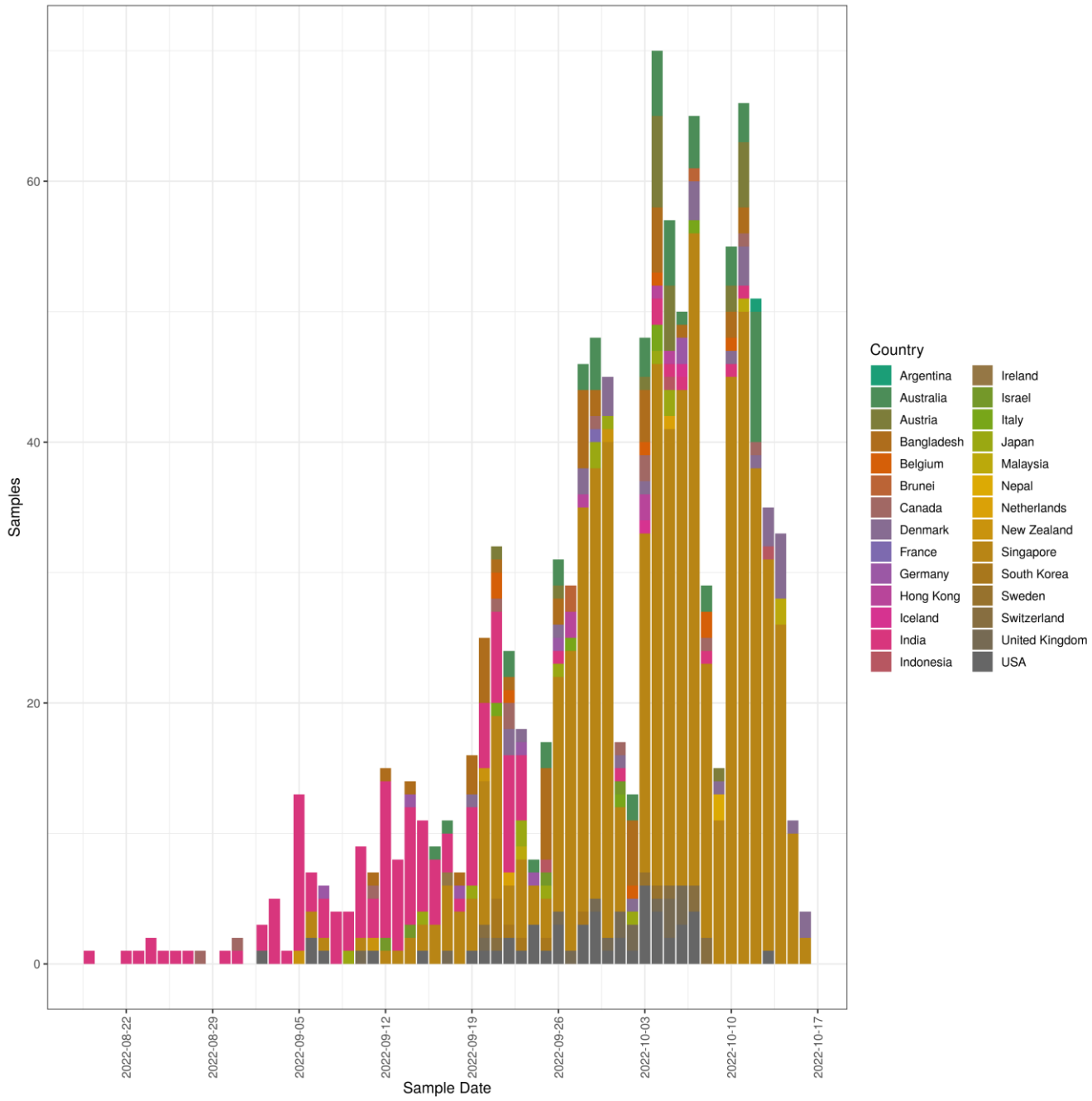
# Part 2. Horizon scanning: newly designated variants

## 2.1 Newly escalated variant: V-22OCT-02 (XBB recombinant)

XBB was first raised as a signal on 11 October 2022 as part of horizon scanning. This recombinant lineage is composed of 2 BA.2 parent lineages BJ.1 and BM.1.1.1, with an approximate break point between spike mutations G446S and N460K. This recombinant is characterised by the acquisition of E: T11A, Spike: V83A, H146Q, Q183E, F486S, F490S. Spike mutations inherited from BJ.1 are G339H, R346T, V445P, G446S and from BM.1.1.1 are N460K, F486V, F490S, and R493Q. XBB contains more RBD mutations at antigenic sites than any other widespread circulating variant and, according to Cao and others, is less well neutralised than either BQ.1.1 or BA.2.75.2.

From 16 October 2022 to 24 October 2022, there have been 1,104 XBB samples uploaded to GISAID from 28 countries, across 5 continents (Figure 9). Currently, the majority of samples uploaded from GISAID are from Singapore (639, 58%). Eighteen English samples have been uploaded to GISAID.

**Figure 9. Epicurve of international XBB lineage sequences by collection date**



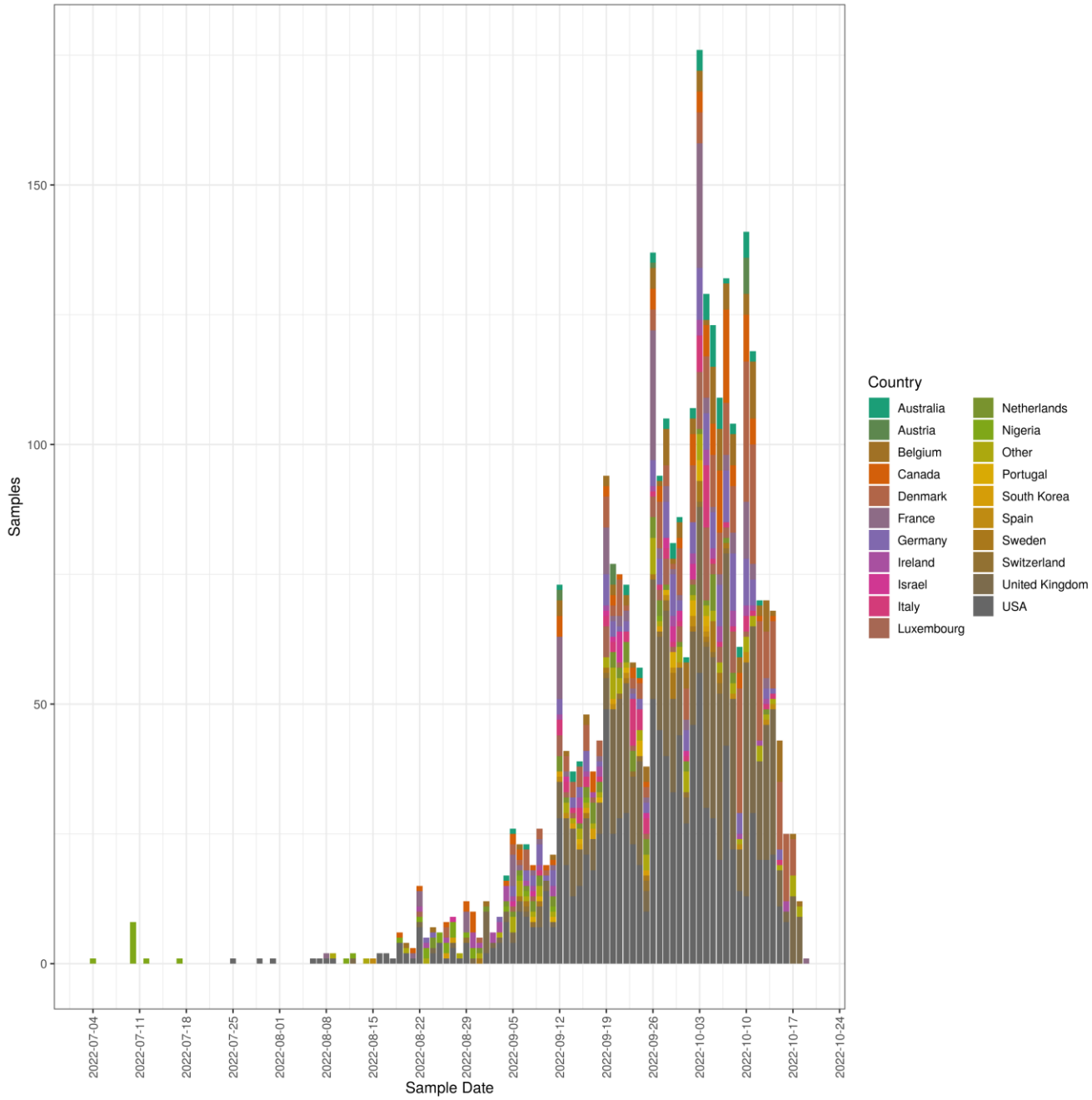

Supplementary data is not available for this figure.

# 2.2 Newly escalated variant: V-22OCT-01 (BQ.1)

BQ.1 (B.1.1.529.5.3.1.1.1.1.1) was first raised as a signal in monitoring on 12 September 2022 as part of horizon scanning. This BA.5 sub-lineage has acquired spike mutations L452R, N460K, and K444T. The affinity impact of N460K is reported to be significant (Saito and others 2022, Cao and others 2022). Additionally, the BQ.1.1 sub-lineage has acquired the spike mutation R346T, a site which has been associated with a recent influx of convergent mutations and appears to infer a notable growth advantage. This sub-lineage is captured as part of UKHSA V-22OCT-01 lineage definition.

From 4 July 2022 to 24 October 2022 there have been a total of 3,207 BQ.1 samples uploaded to GISAID from 48 distinct countries, across 6 continents (Figure 10). Since it was first allocated a Pangolin lineage, BQ.1 GISAID uploads have increased by more than 75% each week. BQ.1 is most commonly uploaded from the United States of America (USA) (1,060) and the UK (717).

**Figure 10. Epicurve of International BQ.1 lineage sequences by collection date – countries with fewer than 10 sequences uploaded have been grouped as "Other"**




Supplementary data is not available for this figure.

A further 3,004 BQ.1.1 samples have been uploaded to GISAID since 21 June 2022, spanning 35 countries and 6 continents. Since BQ.1.1 was first raised as a distinct signal on 20 September 2022, the total number of GISAID uploads have increased by more than 90% each week.

Additionally, 317 BQ.1.2 (Spike I666V) samples have been uploaded to GISAID with collection dates starting from 18 July 2022; 143 BQ.1.3 (Spike E619Q) samples have been uploaded with collection dates starting from 22 August 2022; and 34 BQ.1.4 (Spike R190T) samples have been uploaded to GISAID with collection dates starting from 11 July 2022.

# Part 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 logistic regression and generalised additive models.

Variant growth rates were estimated using 3 models in comparison to different background reference data sets (Table 3). The multinomial model (MM) is fitted with the UShER assigned sequences described in Section 1.2. This differs from previous versions of this report, where the MM was fitted to sequenced cases from the Sanger data set where pangoLEARN is used to assigned lineages. Growth rates are estimated with respect to a given reference lineage. 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 English sequenced cases from 4 April 2022 to 23 October 2022, to model the relative fitness advantages of Omicron lineages. The model is fit at the regional level to account for geographic heterogeneity in variant dynamics. The data set used is the same as described in Section 1.2, with a mixture of Pillar 1 and 2 sequences.

The modelled percentage representation is shown in Figure 11, with relative growth rates compared to BA.5.2 lineages Figure 12. Note that the multinomial model includes several emerging, competitive lineages. This means that a large relative growth rates relative to the presently dominant BA.5.2 must be considered in the context of other competing variants which may also be increasing in representation.

## Logistic regression and generalised additive models

The growth rate is estimated by logistic regression of a variant or lineage of each sample unit on time of sample selection, relative to all other variants. 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 was 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

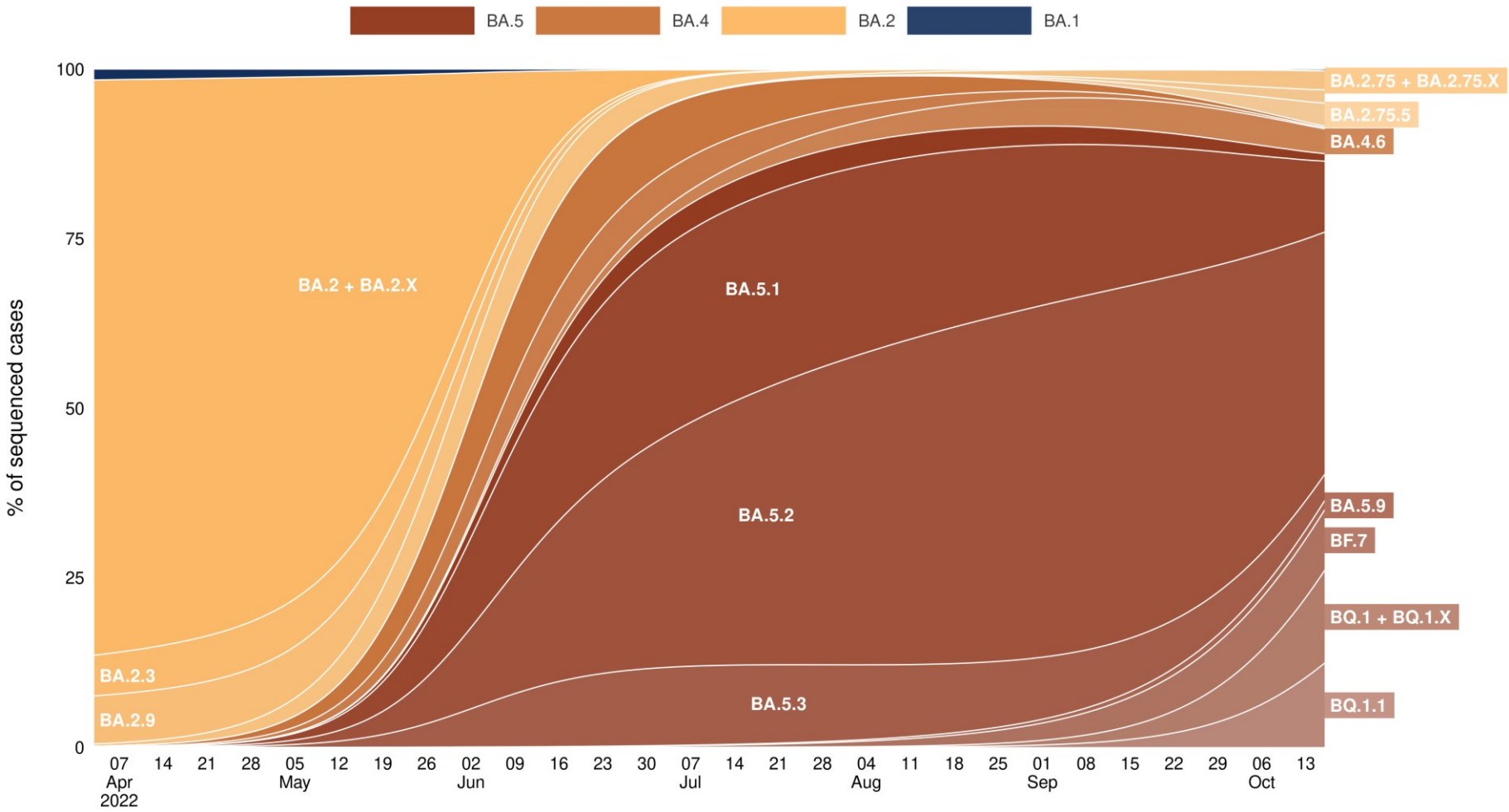
to a geographically matched sample of genomes. A logistic growth rate of zero would indicate no difference in growth rates between a given lineage and other variants.

**Table 3. Growth rate (GR) of variants and signals under monitoring as of 20 October 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) |
|---|---|---|---|---|---|---|
| BQ.1.1 | 1,170 | 12.41% (95% CrI: 10.18 to 15.47) | 63.14% (95% CrI: 58.74 to 67.67) | 1,272 | 54.9% | 23.9% |
| BQ.1 + BQ.1.X* | 1,473 | 13.66% (95% CrI: 11.52 to 16.15) | 53.26% (95% CrI: 49.94 to 56.92) | 1,610 | 43.4% | 25.6% |
| BF.7 | 1,952 | 8.92% (95% CrI: 7.75 to 10.19) | 25.93% (95% CrI: 24.57 to 27.4) | 2,644 | 45.3% | 57.0% |
| BA.5.9 | 407 | 1.36% (95% CrI: 1.02 to 1.79) | 20.02% (95% CrI: 17.73 to 22.34) | 551 | 6.8% | 45.4% |
| BA.4.6 | 3,477 | 3.67% (95% CrI: 3.04 to 4.54) | 4.27% (95% CrI: 3.61 to 4.92) | 2,181 | -1.8% | -8.2% |
| BA.2.75.5 | 401 | 3.4% (95% CrI: 2.36 to 4.92) | 45.06% (95% CrI: 40.31 to 50.07) | 131 | 30% | 19.8% |
| BA.2.75.2 | 325 | 1.96% (95% CrI: 1.47 to 2.56) | 37.32% (95% CrI: 33.22 to 41.6) | 394 | 21.7% | -14.3% |
| BA.2.75 + BA.2.75.X | 657 | 2.81% (95% CrI: 2.23 to 3.4) | 24.18% (95% CrI: 22.11 to 26.24) | 1,130 | 25.6% | 0.7% |

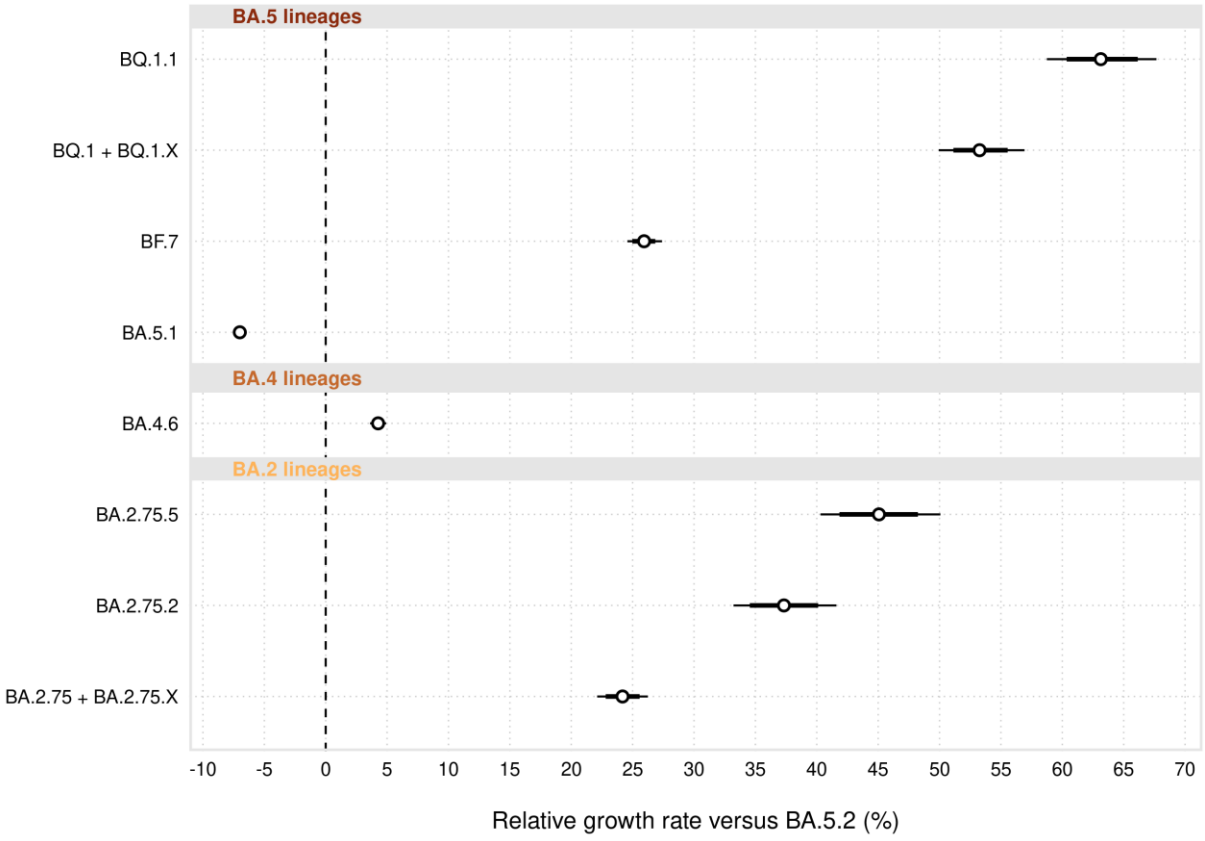
*BQ.1 + BQ.1.X excludes BQ.1.1 which was modelled separately

**Figure 11. Area plot showing the predicted representation of each lineage of the multinomial model of sequenced Pillar 1 and 2 sequenced samples**




This figure shows the predicted representation of different lineages from the multinomial model. Supplementary data is not available for this figure.

**Figure 12. Comparison of the estimated relative growth rates for emerging BA.5, BA.4 lineages versus that for specifically BA.5.2 lineages**




The relative growth rates are taken from a multinomial model of sequenced Pillar 1 and 2 cases in England, described above. Supplementary data is not available for this figure.

# Growth of convergent RBD variants

Because many lineages have independently acquired similar changes to the RBD of the spike protein, growth among sets of sequences defined by the number of RBD mutations at sites of known antigenic significance were separately characterised. Table 4 shows the mutations used to categorise samples.

The growth of sequences with one, 2 or 3 antigenically-significant RBD mutations was assessed against those with no RBD mutations. Logistic growth rates of RBD categories were estimated using a geographically stratified generalised additive model with a smooth effect of time and a binomial response. Four thousand one hundred and eighty-four samples were collected through ONS testing from 1 August 2022 to 7 October 2022.

Figure 13 shows growth of RBD categories in log odds terms and estimated sample frequency. In the most recent week, logistic growth of variants with one, 2 or 3 convergent and antigenically significant RBD mutations was respectively 23%, 47%, and 66% per week. The category with 3

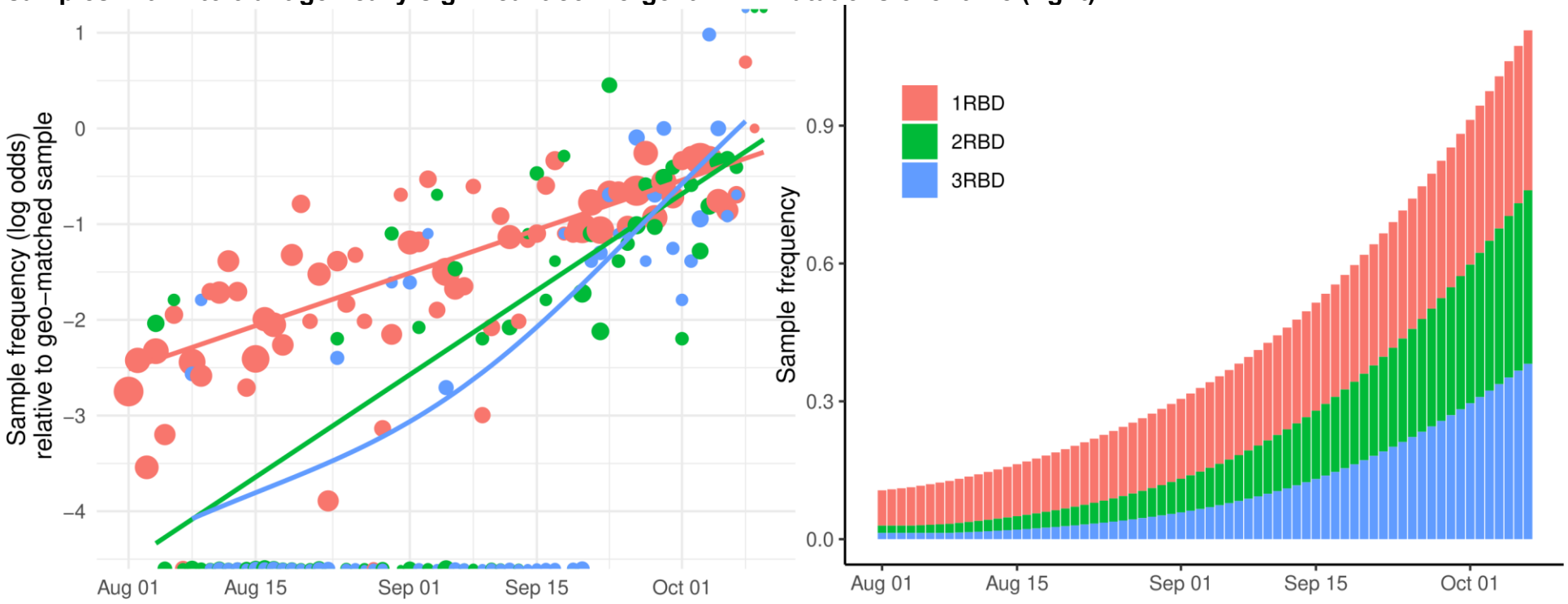
RBD mutations consisted largely of BQ.1.1 (59%) with the remainder consisting primarily of a mixture of BA.2.75 sub-lineages (29%).

**Table 4. RBD mutations* used for categorising sequences for growth modelling – data is drawn from ONS testing 1 August 2022 to 7 October 2022**

| Convergent RBD mutation | n | Proportion |
|---|---|---|
| 1 S:R346T | 804 | 49% |
| 2 S:N460K | 279 | 17% |
| 3 S:K444T | 167 | 10% |
| 4 S:G446S | 128 | 8% |
| 5 S:F486S | 59 | 4% |
| 6 S:R346I | 49 | 3% |
| 7 S:K444M | 36 | 2% |
| 8 S:N450D | 34 | 2% |
| 9 S:V445A | 31 | 2% |
| 10 S:K444R | 30 | 2% |
| 11 S:F490S | 19 | 1% |
| 12 S:F486P | 2 | 0% |

*List of mutations derived from Cao et al., 2022

**Figure 13. Growth of RBD mutation categories relative to a geographically matched sample (left) and estimated proportion of samples with 1 to 3 antigenically-significant convergent RBD mutations over time (right)**




Supplementary data is not available for this figure.

# 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 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 |
| **Virology and Immunology** | |
| Bassam Hallis | UKHSA |

| Person | Institution |
|---|---|
| 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** | |
| Chris Williams | Public Health Wales |
| Daniela de Angelis | University of Cambridge |
| Derek Smith | University of Cambridge |
| Erik Volz | UKHSA / Imperial College London |
| Fergus Cumming | UKHSA |
| 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 |
| Sarah Walker | University of Oxford |
| 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 health secure.

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


Version 1.0

Published: October 2022
Publishing reference: GOV-13512