Viral loads in symptomatic Covid-19 cases in England

Sebastian Funk, Sam Abbott + others tbc

March 22, 2022 - WORK IN PROGRESS DRAFT

For correspondence: sebastian.funk@lshtm.ac.uk

Introduction

Since the start of the SARS-CoV-2 pandemic in late 2019, hundreds of millions of people have been confirmed as infected with the virus. Confirmation of infection has most commonly been done by reverse transcriptase polymerase chain reaction (RT-PCR), with test results indicated as positive if the cycle threshold (Ct) is below a certain level, negative if it is not, and sometimes as ambiguous in an intermediate range. However, the level of the cycle threshold (the Ct value) of test positive cases contains additional information on the infection. Viral loads can be an indicator of severity and transmissibility (Fajnzylber et al., 2020; Knudtzen et al., 2021; Lee et al., 2021; Lyngse et al., 2021). They vary widely between individuals (Challenger et al., 2022). Part of this variation can be explained by changes in viral load over the course of infection in individuals (Kissler, Fauver, Mack, Olesen, et al., 2021). The shape of this so-called viral load curve has been found to vary by age, infecting variant and vaccination status (Hay et al., 2022; Jones et al., 2021; Kissler, Fauver, Mack, Tai, et al., 2021). At the population level this can translate to population-level changes over time (Kissler, Fauver, Mack, Olesen, et al., 2021; Walker et al., 2021).

Because of these properties, there is value in continuing to characterise viral loads in those infected with SARS-CoV-2 and monitor changes therein with changing immunity status due to vaccination or prior infection, as well as with new emerging variants. Here, we analyse viral loads in more than 3 million symptomatic COVID-19 cases in England since the start of the pandemic, as well as the degree to which individual viral loads can be predicted as a function of known individual characteristics such as age and vaccine status.

Methods

Data

We used individual N gene Ct values from symptomatic pillar 2 cases (i.e., mostly from self-initiated testing from individuals presenting with symptoms) in England with symptom onset during or after the week beginning 2020-08-24 alongside data on the week of symptom onset, sex, age, reinfection status, number of doses of vaccination received. We only included those with a positive test up to 6 days following symptom onset to avoid issues with recall. We assigned variant status based on either genetic confirmation or, where no genotyping or sequencing result was available, timing and S-gene target failure (SGTF) status (Table 1). Observations with missing data for any of the extracted variables were dropped from further analysis. Combinations of number of vaccine doses and variants with fewer than 1000 observations were then dropped from the extracted data, leaving 3632894 unique Ct values (Table 2).

Model

We modelled Ct values y for nucleocapsid protein (N) as target gene and assumed an additive relationship between covariates with gamma distributed errors. All covariates for mean Ct values (μ) were then modelled as fixed effects with the effect of variants and vaccines being modelled separately for each age group and

Table 1: Variant assignment for cases with S-gene result but without genetic confirmation.

Variant	Condition
Wildtype	Specimen date < 2021-01-01 AND S-gene positive
Alpha	Specimen date < 2021-04-01 AND S-gene negative
Delta	2021-05-01 <= Specimen date < 2022-01-01 AND S-gene positive
Omicron	Specimen date >= 2021-12-01 AND S-gene negative
Omicron BA.2	Specimen date >= 2022-02-02 AND S-gene positive

Table 2: Number of individual Ct values for each combination of variant and vaccine doses

Variant	Vaccine doses	n
Wildtype	0	504,069
Alpha	0	514,809
Alpha	1	22,096
Delta	0	681,686
Delta	1	192,136
Delta	2	561,437
Delta	3	17,831
Omicron BA.1	0	265,075
Omicron BA.1	1	68,599
Omicron BA.1	2	391,051
Omicron BA.1	3	351,754
Omicron BA.2	0	18,085
Omicron BA.2	1	3,099
Omicron BA.2	2	11,138
Omicron BA.2	3	30,029
Total		3,632,894

combined into a group level effect. Interactions between days from symptom onset to test and all other covariates were also included, again as fixed effects. Mathematically this can be represented as follows:

$$y_i \sim \text{Gamma}(\text{mean} = \mu_i, \text{shape} = k) \tag{1}$$
$$\mu_\text{i} = \beta_{vdr}\text{VDR}_i + \beta_s\text{Sex}_i + \beta_l\text{Lab}_i + f_{\text{Onset,Age}}(\text{Onset}_i, \text{Age}_i, \text{VDR}_i) \tag{2}$$

where VDR_i indicates the variant, number of vaccine doses received and reinfection status (known reinfection or no record of prior infection) of a given individual, s their sex, l the laboratory where the Ct values where measured, and $f_{\text{Onset,Age}}$ a 2-dimensional tensor product smooth varying as a function of days since symptom onset and age, estimated separately for each combination of variant, number of doses and reinfection status.

Implementation

The model was implemented using the `mgcv` (Wood, 2011) package version 1.8-36 in `R` version 4.1.2 (R Core Team, 2019). All code and data required to reproduce this analysis is available from https://github.com/epiforecasts/covid19_ct_pillar2.

Results

Mean viral loads in the unvaccinated varied as a function of the number of days since symptom onset, with a maximum at one (Delta) or two (all other variants) days after infection (Figure 1A). Viral loads at peak were highest for Delta with, however, an earlier decline as a function of time since symptom onset compared to Alpha, the variant with the second highest viral load. Viral loads from Omicron infections declined slower as a function of time since symptom onset than from any of the other variants.

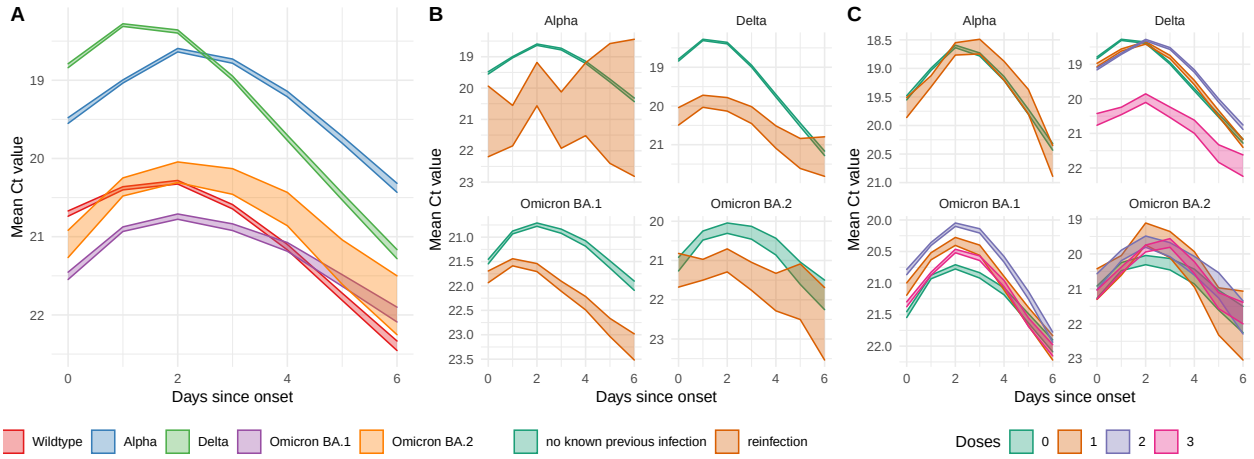

Figure 1: Estimated mean Ct values in (A) unvaccinated patients without any known prior infection, in unvaccinated patients as a function of whether there was knowledge of any prior infection (B), and in patients without any known prior infection as a function of the number of vaccine doses (C). Bands indicate estimated 90% confidence intervals of the means.

With all variants, evidence of a prior infection reduced the estimated mean viral loads in those that with symptomatic infection (Figure 1B). The number of vaccine doses, on the other hand, had less of a discernible effect, with the most discernible effect a reduction in viral loads for those with three doses of vaccination and infected with Delta (Figure 1C).

Viral loads changed over time, largely in response to a change in the dominant variant but also whilst one variant was dominating. Modelling viral loads as a function of time since symptom onset, age, evidence of

prior infection and number of vaccine doses yields poor predictions at the individual level (deviance explained: 9.6%) yet broadly recovers the time-trends in mean viral loads (Figure 2A-B). That said, it appears to do so mostly by capturing the changes with different variants, whilst being less able to recover trends in viral loads whilst a variant is dominating. Part of this may be explained by epidemic phase bias, as symptom onset follows infection and therefore, during periods of exponentially increasing infections, is biased towards more recent infections Hay et al. (2021). Indeed we observe higher viral loads during periods of increase than during periods of decrease for all variants (Figure 2C), and the trajectory of ct values compared to the case trajectories that this can explain differences between predicted and observed ct values.

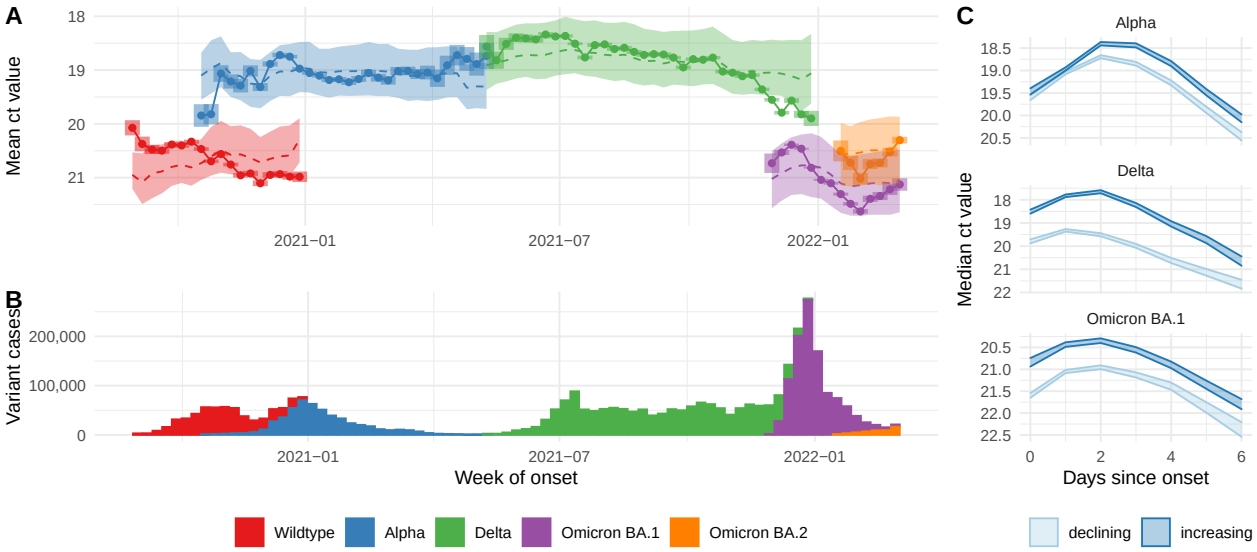

Figure 2: (A) Mean ct values (points and solid lines) estimated independently for each week and variant, with associated 90% confidence intervals as bars; means of modelled ct values (dashed lines) and their interquartile ranges (ribbons). (B) Case numbers identified as being due to a give variant.

Discussion

Analysing a large data set of Ct values in symptomatic COVID-19 cases in the UK, we found variation in viral load by time since symptom onset, vaccine and reinfection status, age and variant. We found highest viral loads in those tested 1-3 days after symptom onset for each of the variants, including when considering variation by vaccination or reinfection status. We further found that the Alpha and Delta variants produced higher viral loads than wildtype and Omicron variants, and that both vaccination and prior infection appeared to reduce viral loads.

In comparing these results to observations made with other data sets, it is important to bear in mind the specific biases that exist in the data we analysed. We only considered positive tests in symptomatic individuals as part of the so-called Pillar 2 testing in England, the vast majority of which would have been based on self-initiated testing at a community. Because of this, our data may preferentially reflect cases with symptoms that are both noticable enough to prompt individuals to get tested whilst not severe enough that the first test would happen in hospital, with age and other variables probably affecting the probability of getting tested as a function of viral load. Moreover, we did not have any information on why individuals would seek out a test only on day 5 or 6 of symptoms and whether this reflects a particular symptom trajectory such as from very mild to more severe, potentially with corresponding behaviour in the viral load trajectory. All that said, whilst these biases and the selection bias of Ct values less than 30 means our observed Ct values cannot be interpreted as samples of individual-level viral load trajectories as in previous work, their smooth behaviour which is broadly in line with that observed previously (Jones et al., 2021; Kissler, Fauver, Mack, Olesen, et al., 2021), indicates that conclusions on broad trends may be valid.

Whilst it has previously been suggested that viral loads may be related to both severity and transmissibility, this does not appear translate into a relationship that would make it possible to infer these from observed viral loads. If anything, the observed viral loads in our data set may relate to severity, previously observed to have been higher in Alpha and Delta than the previous or subsequent variants to date (Davies et al., 2021; Twohig et al., 2022), and lower than with Omicron (Nyberg et al., 2022). Delta and Alpha are also the variant with the highest viral loads we observed. The Omicron variant, on the other hand, which has shown a transmission advantage but decreased severity compared to other lineages (Nyberg et al., 2022; Pearson et al., 2021), shows significantly reduced viral loads.

It has previously been suggested that viral loads may be a useful quantity to monitor as part of real-time surveillance of infectious disease dynamics (Hay et al., 2021). Our results show evidence of "epidemic phase bias", whereby a model taking into account relevant determinants of viral loads such as variant, age and variant and vaccination status underestimates viral loads at times of increasing case numbers and overestimates them at times of decreasing viral loads. This suggest future avenues of research whereby integrating Ct values into analysis of real-time dynamics, in particular at a time when community may have been scaled back and overall case numbers are sampled from a smaller proportion of the population, could improve the accuracy and timeliness of estimates.

References

Challenger, J. D., Foo, C. Y., Wu, Y., Yan, A. W. C., Marjaneh, M. M., Liew, F., Thwaites, R. S., Okell, L. C., & Cunnington, A. J. (2022). Modelling upper respiratory viral load dynamics of SARS-CoV-2. *BMC Medicine*, *20*(1). https://doi.org/10.1186/s12916-021-02220-0

Davies, N. G., Jarvis, C. I., CMMID COVID-19 Working Group, Edmunds, W. J., Jewell, N. P., Diaz-Ordaz, K., & Keogh, R. H. (2021). Increased mortality in community-tested cases of SARS-CoV-2 lineage b.1.1.7. *Nature*, *593*(7858), 270–274. https://doi.org/10.1038/s41586-021-03426-1

Fajnzylber, J., Regan, J., Coxen, K., Corry, H., Wong, C., Rosenthal, A., Worrall, D., Giguel, F., Piechocka-Trocha, A., Atyeo, C., Fischinger, S., Chan, A., Flaherty, K. T., Hall, K., Dougan, M., Ryan, E. T., Gillespie, E., Chishti, R., Li, Y., … Li, J. Z. (2020). SARS-CoV-2 viral load is associated with increased disease severity and mortality. *Nature Communications*, *11*(1). https://doi.org/10.1038/s41467-020-19057-5

Hay, J. A., Kennedy-Shaffer, L., Kanjilal, S., Lennon, N. J., Gabriel, S. B., Lipsitch, M., & Mina, M. J. (2021). Estimating epidemiologic dynamics from cross-sectional viral load distributions. *Science*, *373*(6552). https://doi.org/10.1126/science.abh0635

Hay, J. A., Kissler, S. M., Fauver, J. R., Mack, C., Tai, C. G., Samant, R. M., Connelly, S., Anderson, D. J., Khullar, G., MacKay, M., Patel, M., Kelly, S., Manhertz, A., Eiter, I., Salgado, D., Baker, T., Howard, B., Dudley, J. T., Mason, C. E., … Grad, Y. H. (2022). *Viral dynamics and duration of PCR positivity of the SARS-CoV-2 omicron variant*. https://doi.org/10.1101/2022.01.13.22269257

Jones, T. C., Biele, G., Mühlemann, B., Veith, T., Schneider, J., Beheim-Schwarzbach, J., Bleicker, T., Tesch, J., Schmidt, M. L., Sander, L. E., Kurth, F., Menzel, P., Schwarzer, R., Zuchowski, M., Hofmann, J., Krumbholz, A., Stein, A., Edelmann, A., Corman, V. M., & Drosten, C. (2021). Estimating infectiousness throughout SARS-CoV-2 infection course. *Science*, *373*(6551). https://doi.org/10.1126/science.abi5273

Kissler, S. M., Fauver, J. R., Mack, C., Olesen, S. W., Tai, C., Shiue, K. Y., Kalinich, C. C., Jednak, S., Ott, I. M., Vogels, C. B. F., Wohlgemuth, J., Weisberger, J., DiFiori, J., Anderson, D. J., Mancell, J., Ho, D. D., Grubaugh, N. D., & Grad, Y. H. (2021). Viral dynamics of acute SARS-CoV-2 infection and applications to diagnostic and public health strategies. *PLOS Biology*, *19*(7), e3001333. https://doi.org/10.1371/journal.pbio.3001333

Kissler, S. M., Fauver, J. R., Mack, C., Tai, C. G., Breban, M. I., Watkins, A. E., Samant, R. M., Anderson, D. J., Metti, J., Khullar, G., Baits, R., MacKay, M., Salgado, D., Baker, T., Dudley, J. T., Mason, C. E., Ho, D. D., Grubaugh, N. D., & Grad, Y. H. (2021). Viral dynamics of SARS-CoV-2 variants in vaccinated and unvaccinated persons. *New England Journal of Medicine*, *385*(26), 2489–2491. https://doi.org/10.1056/nejmc2102507

Knudtzen, F. C., Jensen, T. G., Lindvig, S. O., Rasmussen, L. D., Madsen, L. W., Hoegh, S. V., Bek-Thomsen, M., Laursen, C. B., Nielsen, S. L., & Johansen, I. S. (2021). SARS-CoV-2 viral load as a predictor for disease severity in outpatients and hospitalised patients with COVID-19: A prospective cohort study.

PLOS ONE, *16*(10), e0258421. https://doi.org/10.1371/journal.pone.0258421

Lee, L. Y., Rozmanowski, S., Pang, M., Charlett, A., Anderson, C., Hughes, G. J., Barnard, M., Peto, L., Vipond, R., Sienkiewicz, A., Hopkins, S., Bell, J., Crook, D. W., Gent, N., Walker, A. S., Peto, T. E., & Eyre, D. W. (2021). *SARS-CoV-2 infectivity by viral load, s gene variants and demographic factors and the utility of lateral flow devices to prevent transmission*. https://doi.org/10.1101/2021.03.31.21254687

Lyngse, F. P., Mølbak, K., Træholt Franck, K., Nielsen, C., Skov, R. L., Voldstedlund, M., Cohen, A. S., & Kirkeby, C. T. (2021). *Association between SARS-CoV-2 transmissibility, viral load, and age in households*. https://doi.org/10.1101/2021.02.28.21252608

Nyberg, T., Ferguson, N. M., Nash, S. G., Webster, H. H., Flaxman, S., Andrews, N., Hinsley, W., Bernal, J. L., Kall, M., Bhatt, S., Blomquist, P., Zaidi, A., Volz, E., Aziz, N. A., Harman, K., Funk, S., Abbott, S., Hope, R., Charlett, A., ... Thelwall, S. (2022). Comparative analysis of the risks of hospitalisation and death associated with SARS-CoV-2 omicron (b.1.1.529) and delta (b.1.617.2) variants in england: A cohort study. *The Lancet*. https://doi.org/10.1016/s0140-6736(22)00462-7

Pearson, C. A. B., Silal, S. P., Li, M. W. Z., Dushoff, J., Bolker, B. M., Abbott, S., Schalkwyk, C. van, Davies, N. G., Barnard, R. C., Edmunds, W. J., Bingham, J., Meyer-Rath, G., Jamieson, L., Glass, A., Wolter, N., Govender, N., Stevens, W. S., Scott, L., Mlisana, K., ... Pulliam, J. R. C. (2021). *Bounding the levels of transmissibility & immune evasion of the omicron variant in south africa*. https://doi.org/10.1101/2021.12.19.21268038

R Core Team. (2019). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. https://www.R-project.org/

Seaman, S. R., Nyberg, T., Overton, C. E., Pascall, D., Presanis, A. M., & De Angelis, D. (2021). *Adjusting for time of infection or positive test when estimating the risk of a post-infection outcome in an epidemic*. https://doi.org/10.1101/2021.08.13.21262014

Twohig, K. A., Nyberg, T., Zaidi, A., Thelwall, S., Sinnathamby, M. A., Aliabadi, S., Seaman, S. R., Harris, R. J., Hope, R., Lopez-Bernal, J., Gallagher, E., Charlett, A., De Angelis, D., Presanis, A. M., Dabrera, G., & the COVID-19 Genomics UK (COG-UK) consortium. (2022). Hospital admission and emergency care attendance risk for SARS-CoV-2 delta (b.1.617.2) compared with alpha (b.1.1.7) variants of concern: A cohort study. *The Lancet Infectious Diseases*, *22*(1), 35–42. https://doi.org/10.1016/s1473-3099(21)00475-8

Walker, A. S., Pritchard, E., House, T., Robotham, J. V., Birrell, P. J., Bell, I., Bell, J. I., Newton, J. N., Farrar, J., Diamond, I., Studley, R., Hay, J., Vihta, K.-D., Peto, T. E., Stoesser, N., Matthews, P. C., Eyre, D. W., & Pouwels, K. B. (2021). Ct threshold values, a proxy for viral load in community SARS-CoV-2 cases, demonstrate wide variation across populations and over time. *eLife*, *10*. https://doi.org/10.7554/elife.64683

Wood, S. N. (2011). Fast stable restricted maximum likelihood and marginal likelihood estimation of semiparametric generalized linear models. *Journal of the Royal Statistical Society (B)*, *73*(1), 3–36.