

Nowcasting epidemic trends using hospital- and community-based virologic test data

Corresponding Author: Dr Tse Yang Lim

This file contains all reviewer reports in order by version, followed by all author rebuttals in order by version.

Version 0:

Reviewer comments:

Reviewer #1

(Remarks to the Author)

In this study, Lim and colleagues evaluated the utility of virologic test data (i.e., cycle threshold values) in nowcasting epidemic trends in hospital and community settings. Simulations were conducted to evaluate the performance of using Ct values to estimate time-varying epidemic growth rates. The simulation setting also allowed the authors to explore the robustness of estimation to various sources of uncertainty/biases in real-world applications, such as viral kinetics, individual variation, and sampling strategies. The methods were then applied to real-world Ct data from three locations collected in both community and hospital settings. Nowcasting performance was assessed by altering several technical choices (sample size, accounting for population immunity or symptom status, trimming CT value outliers, etc.). Overall, the manuscript presents a comprehensive and technically sound analysis of nowcasting using Ct values. The thorough evaluation of the robustness of the proposed method to real-world data challenges is particularly useful. Here I have a few comments and questions on the study.

1. Using Ct values to estimate epidemic trends has been extensively discussed during the COVID-19 pandemic. Many existing studies have attempted to develop methods to perform nowcasting using Ct values, including works from the authors' team and other groups. For instance, in Ref. [30], Lin et al. applied a similar regression model approach to estimate COVID-19 transmission in Hong Kong. I am wondering if the authors could better highlight the novel insights from this study and put the findings in the context of existing literature.

2. The study was motivated by the advantage of using Ct values to track epidemic trends over more traditional data sources. However, there was no comparison with those approaches in the synthetic and real-world data analysis. Could the authors compare the performance of the Ct value-based method with existing models based on case numbers? I understand that there are many nowcasting methods and it is not possible to perform an exhaustive comparison. But presenting a comparison with at least one baseline method can better substantiate the advantage of using Ct values – whether it can detect the trend earlier or estimate the growth rate more accurately?

3. Throughout the analysis, the “ground truth” was set as the confirmed case number. Have the authors considered the impact of changing reporting rate over time? Maybe this variation is not important given that the goal here is to estimate instantaneous growth rates, but some analyses or discussions are warranted.

Some minor concerns:

1. Clarification of Modeling Approaches:

- The authors utilized two different models—Generalized Additive Models (GAMs) and logistic regression—but this distinction is not clearly articulated in the abstract and main text. In the GAMs, natural logarithms were applied, whereas in the logistic regression model, splines may have been introduced. Since GAMs are regression models and logistic regression is a classification model, it's important to clarify their differences and the specific purposes they serve in the study.
- The methods section lacks a clear explanation or formula for the logistic regression model used to predict directions, rendering the description of the method inadequate.

2. Residual Analysis and Model Testing:

While it's appreciated that the authors conducted sensitivity analyses and tested different model parameters, it's recommended to include residuals or model diagnostic tests for the GAMs. This would ensure there are no patterns in the residuals, confirming the model's adequacy.

3. Clarification of Train-Test Split:

- On page 6, line 113, and page 13, line 201, the term "train-test split date" is mentioned, but there's no corresponding information in Table S2.

- Please specify the exact periods used for training, validation, and testing, and detail the analyses performed during each phase.

4. Terminology Clarification:

The use of the term "rolling window" is misleading. Typically, a rolling window refers to a fixed-size window of data that moves forward over time. In the authors' approach, the training window sizes are not fixed but increase every two weeks, with only the prediction period being flexible. This method is usually called an "expanding window" with a fixed lead time (prediction period). The terminology should be corrected for accuracy.

5. Choice of Prediction Window Size:

The authors chose a two-week prediction window. Is there a specific reason for selecting this duration? Using shorter window sizes might improve the results, especially during periods of high growth rates. The rationale behind the chosen window size should be discussed.

6. Down-Sampling Analysis Reporting:

The down-sampling analysis results are reported as single point estimates. Given that 100 random samples were used, it would be more informative to report the results with error bars to reflect variability and uncertainty.

7. Figure Improvements:

In Figure 6 on page 18, the text "Alternate" is not fully displayed.

The second row of the plot lacks an explanation for the light grey bars. Please provide a legend or description to clarify their meaning.

(Remarks on code availability)

Reviewer #2

(Remarks to the Author)

I co-reviewed this manuscript with one of the reviewers who provided the listed reports. This is part of the Nature Communications initiative to facilitate training in peer review and to provide appropriate recognition for Early Career Researchers who co-review manuscripts.

(Remarks on code availability)

Reviewer #3

(Remarks to the Author)

This manuscript investigates whether viral load data, measured by Ct values, can be used to predict epidemic trends. The manuscript is clearly written. However, I have a few major comments:

1. While the idea of using community-based Ct values for outbreak surveillance is interesting, I have reservations about its practical utility. Unlike other population-level measures such as wastewater viral concentrations, this approach requires individual-level sampling and could therefore be heavily influenced by individual-level heterogeneity. Although the temporal dynamics of viral load may follow similar trends over the course of infection (e.g., declining after peak viral load), the absolute Ct values could vary significantly among patients at similar disease stages (such as during the peak of shedding). Capturing this variation may be challenging.

2. The authors present both RMSE (to measure the accuracy of predicting epidemic growth rates) and AUC (to measure the accuracy of predicting the epidemic direction). Interpreting the absolute value of the RMSE is difficult because it is generally used to compare the performance of competing models. For example, in line 115, the authors note that their in-sample RMSE is 0.00191, which is approximately 10% of the range in observed log incidence growth rates. It is unclear whether this 10% figure represents a good or poor performance without a clearer baseline for comparison.

3. I may have missed it, but I do not see a clear definition of "epidemic growth rate" in the manuscript.

4. The results suggest that the predictability of the epidemic growth rate is weak. For instance, the visual fit between observed and predicted rates appears poor in Figure 1, and most of the correlations between observed and predicted values are below 0.5.

5. The AUC for predicting epidemic direction seems reasonably large, although it is a less interesting target.

6. How does the performance of this approach compare to other methods, such as using wastewater data?

Overall, the manuscript is clearly written, but some key findings appear to be statistically insignificant, and the practical utility of the proposed approach might be limited.

(Remarks on code availability)

Reviewer #4

(Remarks to the Author)

In this paper, the authors conducted a systematic and comprehensive study to investigate the feasibility of using SARS-CoV-2 Ct values to nowcast epidemic trajectories over three years of the COVID-19 pandemic, using both synthetic data and real-world data from Mass General Brigham hospital system and Los Angeles County.

While I do think the study provides new methods in using viral load data in epidemic nowcasting, my main concern is that whether this manuscript has added sufficiently novel and important insights to warrant publication in Nature Communications, especially that the ideas to use Ct data in nowcasting have been published in several papers by the same group earlier. Please see my specific comments below.

1. Introduction: It might be a bit assertive to state that “estimating epidemic growth rate based on the distribution of measured viral loads does not depend on the number of positive tests”, because the estimation relies on a representative sample of the viral load measurements from the population which would be more likely to get when there were more positive tests.
2. The estimation is dependent on the assumption that a lower Ct value suggested that an individual was sampled earlier during their infection, which was true for SARS-CoV-2. For other pathogens, if the differences in viral load between early and late stage of infection were not as large as SARS-CoV-2, how would the performance of the nowcasting change? Does it fit to any synthetic scenarios (e.g., asymmetric?) described in Table S1?
3. Following the comment above, how would the model performance change with varying shedding profiles/distributions of other respiratory pathogens?
4. In practice, Ct data of earlier outbreaks were available for model training for later outbreaks. How would the performance of the MGB and LAC models change if Alpha and/or Delta data were used to train model for nowcasting of the Omicron outbreaks?
5. For the current model, only the viral load or Ct data were used in the prediction. Would the model performance be improved if the incidence data were also incorporated?
6. It would be useful to readers and users of the nowcasting methods if the growth rate estimates could be translated to real-time reproductive numbers with a given generation time distribution. The changes in growth rates were a bit difficult to interpret – for example, the predicted direction and the growth rate matter a lot when reproductive number is much higher than one, but it might be less worrying when the reproductive number is fluctuating around one.

(Remarks on code availability)

The methods are well documented, and the codes and synthetic data have been made available online.

Reviewer #5

(Remarks to the Author)

Lim et al. demonstrate a nowcasting method to predict SARS-CoV-2 epidemic growth rates based on cross-sectional Ct measurements. This follows up on the work by some of the same authors (Hay et al, Science, 2022), which first showed that Ct distributions contain relevant information to track epidemic dynamics. This new work extends the previous, through its development of a nowcasting model to predict epidemic growth rates from the CT-distribution, some sensitivity analyses and more extensive verification on two datasets from the COVID-19 pandemic.

This manuscript is well-written and clearly developed with scientific rigour. However, our primary reservations are around the relative advance posed by this paper, and the independence of this work with respect to the previous paper. These points may also affect the breadth of readers attracted by the manuscript. At present, we believe this manuscript may be most suited for a specialized audience of epidemiologists/biostatisticians. To aid such an audience in their use of the presented method, further sensitivity analyses and comparison to established techniques would be helpful.

1. (Line 77-94) The manuscript would benefit from a figure highlighting the intuition behind the relationship between CT value distribution and epidemic trajectory. While we realize that this has already been shown elsewhere, not including it here make the logic behind different modeling choices (e.g. a more or less symmetric distribution) harder to understand and possibly prohibitive for readers not familiar with the previous work.

2. The question that is set up to guide the advances made in this manuscript is: “it remains unclear under which conditions [population-level Ct values] are a practical source of epidemiological information. [...] Whether [a multitude of factors] are prohibitively confounding when using Ct value distributions for epidemic monitoring has yet to be explored.” (in lines 58-68) We are not sure this question is actually comprehensively covered by the four non-ideal modeling scenarios studied here.

a) “increased symmetry in viral kinetics” is an arbitrary choice. Why not study this for a number of different choices of “increased symmetry” (a continuous scale)?

b) Variation in the observed viral load given a time-since-infection is assumed to only change the variance, but what if there are two underlying populations (e.g. symptomatic/asymptomatic, or old/young) that have two separate distributions? What if those underlying classes are correlated with the delay between symptom and testing too, or only arise during a specific phase in the epidemic?

c) Why were the scenarios tested with only a single draw of a simulated epidemic scenario? What is the stochasticity in performance across different random seeds or slightly different initial values? Rather than simulating such a long dataset, we believe the focus should be on repeatability across the same scenario many times. One minimal way to approximate this

might be to split up the current simulated dataset into timeperiods that were similar (e.g. going from low to high incidence) and comparing performance per class of scenarios.

d) It is misleading that Fig 1 includes different colors for the different variants since no differences were assumed. Why didn't the authors simulate the impact of variant replacement dynamics and/or different Ct-time since infection relationships for the different variants?

3. In the two use cases on realistic data, how does the present nowcasting model (GAM with smoothing splines) differ from the SEIR model introduced in Hay et al 2022? The relevant comparison here seems to be against other Ct-based methods and/or against the RMSE of other transmission dynamic inference methods (e.g. Re estimation) on incidence data? In general it is extremely difficult to assign a sense of quality to the reported RMSE/AUC numbers. How well would a null model that just predicts the same average incidence growth rate as the past 2 weeks do?

4. How do the RMSE and AUC differ and depend on the phase of the epidemic?

— Minor Points —

Line 105: please define what RMSE and AUC stand for (abbreviation). Also specify which area under which curve you are measuring.

Line 168/169: what do the authors think could explain the difference in lagging/leading the incidence growth rates?

Fig 2B) It may be slightly misleading that the blue line here represents the average while in the first figure it was the simulated truth.

Fig 2C) what is the intuition behind the wildly different inferred mean/skewness relations for the different variants?

Line 217 What is "the directional discrimination test"?

Fig 6: the top left label is squished. Also the abbreviation "TFT" and the grey vertical bars are not mentioned in the caption.

Fig S13: visually the performance of the trained model for the mean of the Ct distribution seems poor on the test data. How do the authors explain that the model still performed well?

Supplement line 235 "using an extremely left-skewed viral kinetics curve" - Please include a reference to table S7 here. A figure showing the different distributions would also help a lot.

(Remarks on code availability)

We looked at the code very briefly. From a first glance it seems well organized and comprehensive. More careful review would be appropriate during a later round of revisions.

What we couldn't find but might be helpful is some pointers on how to re-plot the outcome of the synthetic data analysis without re-running the simulations.

Reviewer #6

(Remarks to the Author)

I co-reviewed this manuscript with one of the reviewers who provided the listed reports. This is part of the Nature Communications initiative to facilitate training in peer review and to provide appropriate recognition for Early Career Researchers who co-review manuscripts.

(Remarks on code availability)

Version 1:

Reviewer comments:

Reviewer #1

(Remarks to the Author)

The authors have revised the manuscript and addressed most of the concerns raised in the initial review. The updated version includes expanded explanations, additional analyses, and a clearer description of methods and results.

First, the revised manuscript articulates the novelty and connection to existing literature in the introduction.

Second, the revised version now includes a comparative analysis of three model types: Ct-only, case-only, and a combined

model (see Table 2).

Third, regarding the use of confirmed case numbers as ground truth—despite potential bias introduced by time-varying reporting rates—the authors acknowledge this limitation and argue that the short prediction horizon may reduce the impact of such biases. This explanation is reasonable, although a sensitivity analysis could further substantiate the claim.

In terms of methodological clarity, the revised methods section now clearly distinguishes between the generalized additive models (GAMs) used to predict continuous growth rates and logistic GAMs used to classify epidemic direction. Although complete model formulas are not provided, the descriptions of link functions and variables in the Supplementary Materials are adequate. Moreover, residual analysis is appropriately included and suggests the models are well-behaved (see Figure S2).

Finally, the authors have addressed other technical and formatting issues raised previously. Notably, they improved the terminology (e.g., clarifying “expanding” vs. “rolling” windows), reported variability in down-sampling analyses, and enhanced figure legibility. The expanded discussion of limitations—particularly the impact of assay heterogeneity and individual-level variability in Ct values—is thoughtful and well integrated.

(Remarks on code availability)

Reviewer #2

(Remarks to the Author)

I co-reviewed this manuscript with one of the reviewers who provided the listed reports. This is part of the Nature Communications initiative to facilitate training in peer review and to provide appropriate recognition for Early Career Researchers who co-review manuscripts.

(Remarks on code availability)

<https://github.com/gradlab/ct-nowcasting> page can not be opened.

Reviewer #3

(Remarks to the Author)

While I appreciate the authors' responses, I believe my earlier concerns have not yet been fully addressed.

(Remarks on code availability)

Reviewer #4

(Remarks to the Author)

Thank you for inviting me to review the revised manuscript. I believe the authors have adequately addressed my previous comments.

The revised manuscript provides a useful and informative comparison between nowcasting using case counts and nowcasting using Ct data, based on two real-world datasets (MGB and LCA). I concur with the earlier comments from Reviewer #1 and Reviewer #5. My primary concern remains regarding the relative novelty and the extent of advancement of the current findings in comparison to previous work by the co-authors (e.g., Hay et al.).

(Remarks on code availability)

I reviewed the codes in the first round of review.

Reviewer #5

(Remarks to the Author)

The authors have substantially revised the manuscript and succeeded in either addressing our concerns or arguing why this is not possible/necessary. We just have one follow-up question, which should not stand in the way of publication.

The new Fig. 1 is interesting and indeed helps with intuition. However, the dynamics in panel B, row 2, are not intuitive to us. Why do the fully symmetric viral kinetics lead to an inverse relationship between growth and Ct value? In this scenario the symmetry of the viral kinetics is not the only thing that changes, as also the date of peak Ct and the range of dates with Ct above 40 are affected. This will naturally change the timing between infection incidence and peak Ct, even without regard for the symmetry of the shedding. How can the authors be sure that the symmetry is the quantity that truly leads to the reported inverse relation between growth and Ct value?

Minor typos:

- The introduction was updated to read “virosolver R package, which models the convolution of viral kinetics and incidence

to predict population-level viral load distributions, to estimate infection incidence using Ct values from cross-sectional^{21,31,32,35,36,40}. “ but there is a word missing at the end of this sentence.

(Remarks on code availability)

The github folder was not accessible at the link provided. Instead the previous folder location <https://github.com/gradlab/ct-nowcasting-review> is still accessible. However, it was not updated for 10 months, so it is unclear whether this is the most current folder.

Reviewer #6

(Remarks to the Author)

I co-reviewed this manuscript with one of the reviewers who provided the listed reports. This is part of the Nature Communications initiative to facilitate training in peer review and to provide appropriate recognition for Early Career Researchers who co-review manuscripts.

(Remarks on code availability)

Open Access This Peer Review File is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

In cases where reviewers are anonymous, credit should be given to 'Anonymous Referee' and the source.

The images or other third party material in this Peer Review File are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.

To view a copy of this license, visit <https://creativecommons.org/licenses/by/4.0/>

Reviewers' comments:

Reviewer #1 (Remarks to the Author):

In this study, Lim and colleagues evaluated the utility of virologic test data (i.e., cycle threshold values) in nowcasting epidemic trends in hospital and community settings. Simulations were conducted to evaluate the performance of using Ct values to estimate time-varying epidemic growth rates. The simulation setting also allowed the authors to explore the robustness of estimation to various sources of uncertainty/biases in real-world applications, such as viral kinetics, individual variation, and sampling strategies. The methods were then applied to real-world Ct data from three locations collected in both community and hospital settings. Nowcasting performance was assessed by altering several technical choices (sample size, accounting for population immunity or symptom status, trimming CT value outliers, etc.). Overall, the manuscript presents a comprehensive and technically sound analysis of nowcasting using Ct values. The thorough evaluation of the robustness of the proposed method to real-world data challenges is particularly useful. Here I have a few comments and questions on the study.

1. Using Ct values to estimate epidemic trends has been extensively discussed during the COVID-19 pandemic. Many existing studies have attempted to develop methods to perform nowcasting using Ct values, including works from the authors' team and other groups. For instance, in Ref. [30], Lin et al. applied a similar regression model approach to estimate COVID-19 transmission in Hong Kong. I am wondering if the authors could better highlight the novel insights from this study and put the findings in the context of existing literature.

We believe that our study provides novel insights in four areas:

1. The relationship between Ct-value statistics and epidemic growth rates/ R_t is not necessarily linear, as has been assumed in the regression-based approaches that many have employed (e.g. Harrison et al. 2023, Lin et al. 2022, Khalil et al. 2022, Sharmin et al. 2023, Meng et al. 2025). We compare multiple, flexible GAMs to evaluate non-linearity.
2. The relationship between Ct-value statistics and growth rates depends on the populations sampled. By comparing two very different testing approaches (hospital-based and community-based) in a single framework, we show (1) the Ct-value to growth rate relationship is complex and varied (2) that a flexible regression framework can still lead to useful growth rate predictions.
3. Other than the referenced study by Lin et al., no study explores the use of Ct values in a true nowcasting analysis. Instead, past work typically trains models using data from one epidemic wave to perform predictions on a subsequent wave. In our analysis, we iteratively retrain our model over time. This better reflects a real-world use case, where new data will be continually collected and can be used to refine model predictions.
4. In addition, our study goes beyond Lin et al. and other examples of Ct value nowcasting by exploring in depth the impact of multiple factors that can affect the Ct value-growth rate relationship and subsequent predictive performance, including viral kinetics symmetry, variation in viral kinetics and Ct value measurements, sample size, and the delay between symptom onset and detection. On top of the empirical component, our study includes extensive theoretical and synthetic data analyses to this end.
5. We did not seek to model pathogens other than SARS-CoV-2, as this is being explored in a study by our collaborators at Hong Kong University (<https://doi.org/10.1101/2025.05.22.25328079>).

We reframed the introduction to better summarize what has been done in previous studies and what our study adds:

Multiple studies have reported on the feasibility of using population-level Ct values to track SARS-CoV-2 epidemic trends, the majority demonstrating a negative correlation between average Ct values and either the effective reproduction number (R_t), case counts, or hospitalizations^{28–38}. Some have gone further to predict R_t or epidemic growth rates from Ct value statistics using linear regression^{28,30,35,36,39}, autoregressive integrated moving average (ARIMA) models^{33,37}, or machine-learning algorithms⁴⁰. Other studies used the *virosolver* R package, which models the convolution of viral kinetics and incidence to predict population-level viral load distributions, to estimate infection incidence using Ct values from cross-sectional^{21,31,32,35,36,40}. To our knowledge only one study has explored the use of Ct-based statistics for nowcasting epidemics (i.e., regularly updating growth rate estimates on a rolling basis) as opposed to retrospectively predicting entire epidemic waves⁴³. In addition, these studies are generally limited to single settings and datasets, short time periods, and do not consider factors that may influence the accuracy of Ct-based inferences of epidemic growth rates, such as changes in viral kinetics and sampling schemes (e.g., symptom-based vs. random).

Here, we present an in-depth investigation of the real-world feasibility of using SARS-CoV-2 Ct values to nowcast epidemic trajectories, using a combination of theory, synthetic data, and multiple real-world datasets representing different sampling strategies and settings. We start with a theoretical exploration of how biological and logistical factors affect the non-linear relationship between population-level Ct value distributions and epidemic growth rates, using analytical convolutions of simple epidemic curves with viral load trajectories. We next establish how this relationship can be modeled to nowcast growth rates using reported Ct values. We then benchmarked nowcasting model performance using synthetic datasets, before applying the same models to three real SARS-CoV-2 RT-qPCR testing datasets, collected across multiple geographic areas in the United States and under different population sampling strategies, to assess and inform the use of this approach in real-time estimation of epidemic growth compared to using positive test counts alone. Our analyses demonstrate the possibilities and limitations of Ct-based epidemic growth rate surveillance.

2. The study was motivated by the advantage of using Ct values to track epidemic trends over more traditional data sources. However, there was no comparison with those approaches in the synthetic and real-world data analysis. Could the authors compare the performance of the Ct value-based method with existing models based on case numbers? I understand that there are many nowcasting methods and it is not possible to perform an exhaustive comparison. But presenting a comparison with at least one baseline method can better substantiate the advantage of using Ct values – whether it can detect the trend earlier or estimate the growth rate more accurately?

We thank the reviewer for this suggestion. Although our initial aim was just to demonstrate the predictive performance of Ct-based metrics alone rather than seeking to optimise predictions by including additional data, we agree that benchmarks for comparison would be useful. We updated the main model results to now compare three approaches:

1. Predicting growth rate of cases at the state or county level using the growth rate of new positive tests in the two datasets.
2. Predicting growth rate of cases at the state or county level using the Ct values from the two datasets.
3. Predicting growth rate of cases at the state or county level using both the growth rate of new positive tests and Ct values in the two datasets.

We compared the RMSE using these three sets of variables, establishing baseline predictive performance with (1) and better quantifying the additional information gained by including Ct values through comparing (1) with (3).

These new models showed contrasting results across locations. In Massachusetts, the growth rate of positive tests was a far better predictor of overall incidence growth rates than Ct values, and the inclusion of Ct values in the models marginally improved prediction accuracy. In LA County, Ct value based models performed better than models based on positive test numbers.

These results both provide a better quantitative sense of the accuracy of the Ct value models and reinforce what we propose is the use case for Ct values, namely providing an alternative means of tracking epidemic trajectories that is especially valuable when representative surveillance sampling or comprehensive testing are impractical. We also found in our sensitivity analyses of these new models that models using the growth rate of positive tests were far more sensitive to sample sizes than models based on Ct value distributions, further reinforcing this point. Note that our aim is not to claim Ct values universally provide more accurate tracking of epidemics than traditional data sources, but rather to demonstrate the possible uses and limitations of the method.

We also added a 'null model' as an additional performance benchmark, in which we used the mean growth rate over the preceding 2 weeks as the predicted growth rate for the following two weeks.

3. Throughout the analysis, the “ground truth” was set as the confirmed case number. Have the authors considered the impact of changing reporting rate over time? Maybe this variation is not important given that the goal here is to estimate instantaneous growth rates, but some analyses or discussions are warranted.

Choosing a “ground truth” for the real-world analyses is challenging, as the true infection incidence curve is not known for either LA County or Massachusetts without representative, random sampling studies (such as the ONS Covid-19 Infection Survey in the UK). As the reviewer points out, state- and county-wide case reports could be biased by changing reporting rates, testing algorithms, and test-seeking behaviour over time. We chose reported cases as the target variable for our predictions, but we note that it is likely a biased proxy for the true infection incidence. We have acknowledged this limitation in the Discussion and highlighted avenues for further work exploring the impact of alternative metrics:

Beyond confounding factors, it is plausible that the growth rate of reported COVID-19 cases may not be the most accurate benchmark against which to compare Ct value distributions. Reporting rates and delays may also change over time, causing deviations between reported case counts and true infection incidence. Alternative benchmarks, such as growth rate in hospitalizations, mortality, or wastewater viral loads, may therefore yield stronger relationships (possibly with some time-shifting); investigating these relationships would be a fruitful avenue for further research. In addition, geographically aggregated incidence may mask heterogeneous outbreak trajectories at finer scale, e.g., city or even neighborhood level. Such finer-scale incidence data may yield cleaner relationships with Ct value distributions, especially if matched to the catchment areas for the Ct value data collection process.

Some minor concerns:

1. Clarification of Modeling Approaches:

- The authors utilized two different models—Generalized Additive Models (GAMs) and logistic regression—but this distinction is not clearly articulated in the abstract and main text. In the GAMs, natural logarithms were applied, whereas in the logistic regression model, splines may have been

introduced. Since GAMs are regression models and logistic regression is a classification model, it's important to clarify their differences and the specific purposes they serve in the study.

We have updated the text throughout to clarify that these models predict different outputs but are otherwise very similar. They use the same covariates and splines but different link functions – a Gaussian link function for growth rate and a logit link function for epidemic direction. We made the following text changes:

Abstract: Observed Ct value distributions accurately predicted epidemic growth rates (growth rate RMSE ~ 0.039-0.052) and, using a separate model, direction (AUC ~ 0.72-0.78).

Results section: We also fit GAMs to predict the epidemic direction (i.e., whether incidence is growing or declining) using a logit link function and with the same explanatory variables and smoothing splines.

Methods: We also modeled epidemic direction using the same set of GAMs used for incidence growth rate, but with a logit link function to predict binary outcome data (growing or shrinking epidemic):

$$\ln\left(\frac{p_t}{1-p_t}\right) = \beta_0 + s_{\bar{x}}(\bar{x}_t) + s_g(g_t) + \beta_v v_t$$

Where p_t is the probability that the epidemic is growing at time t , and declining otherwise. This is implemented by setting `family=binomial()` in `mgcv:gam`.

- The methods section lacks a clear explanation or formula for the logistic regression model used to predict directions, rendering the description of the method inadequate.

We added this equation to the methods as described above.

2. Residual Analysis and Model Testing:

While it's appreciated that the authors conducted sensitivity analyses and tested different model parameters, it's recommended to include residuals or model diagnostic tests for the GAMs. This would ensure there are no patterns in the residuals, confirming the model's adequacy.

Thank you for the suggestion. We added residual plots for the simulation analyses and real data in the supplementary material.

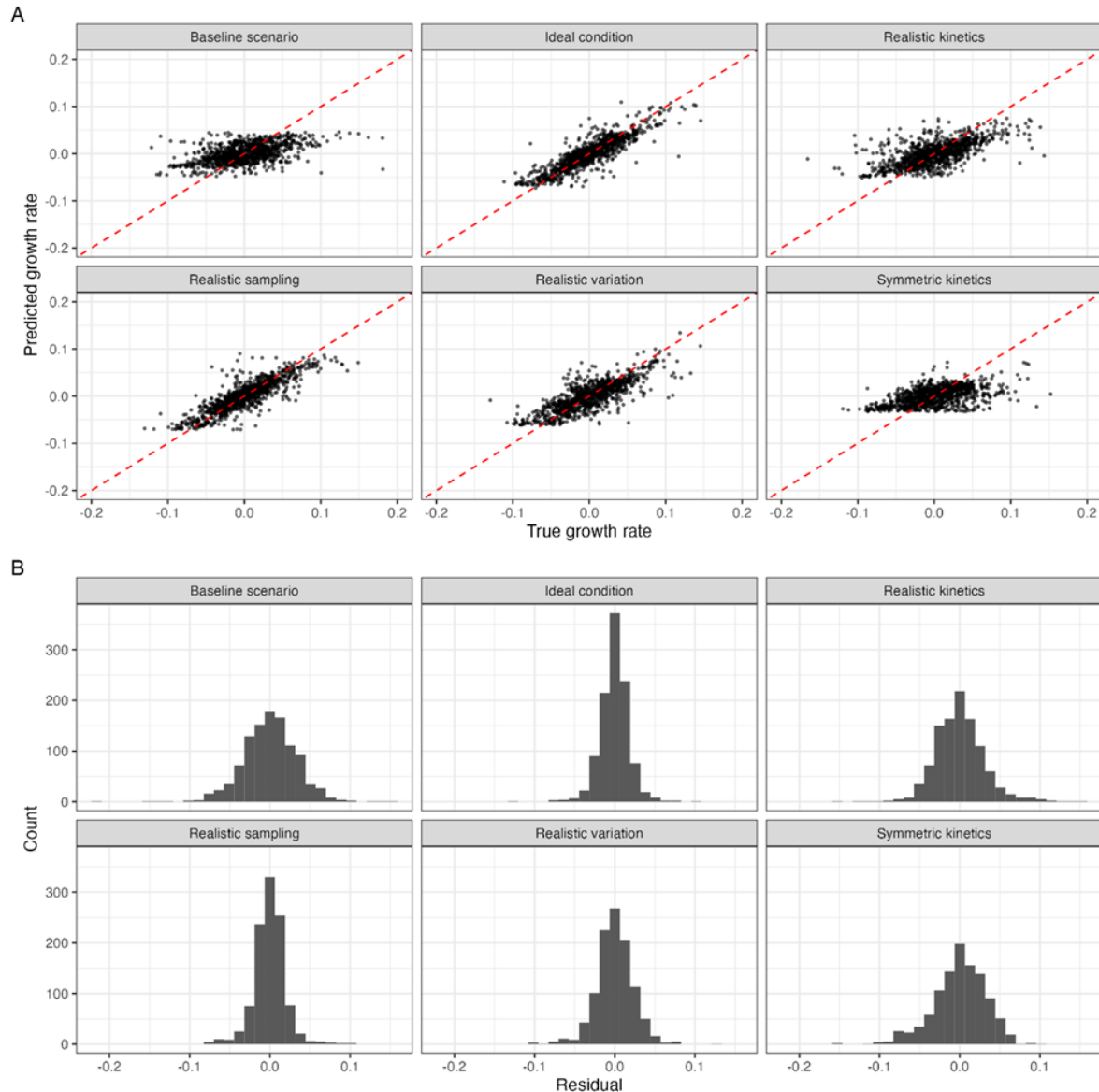


Figure S2. (A) Comparison of model-predicted incidence growth rates with true growth rates from the synthetic dataset analyses. Estimates shown are from in-sample fits for each scenario using the full time series of data to train the GAM. (B) Distribution of residuals between model predictions and true growth rates.

3. Clarification of Train-Test Split:

- On page 6, line 113, and page 13, line 201, the term "train-test split date" is mentioned, but there's no corresponding information in Table S2.

We clarified that the split date is 31 Dec 2021 in these locations and also made this clearer in the methods section, under "Evaluating model performance".

- Please specify the exact periods used for training, validation, and testing, and detail the analyses performed during each phase.

We clarified the methods section, "Evaluating model performance". For the fixed test/train analysis, the data was split on or before and after 31 Dec 2021. For the rolling nowcast test, we initially trained the model using the first 16 weeks of available data (date varies by dataset) and then sequentially retrained the model in expanding 2-week windows.

4. Terminology Clarification:

The use of the term "rolling window" is misleading. Typically, a rolling window refers to a fixed-size window of data that moves forward over time. In the authors' approach, the training window sizes are not fixed but increase every two weeks, with only the prediction period being flexible. This method is usually called an "expanding window" with a fixed lead time (prediction period). The terminology should be corrected for accuracy.

Thank you for highlighting this. We changed the wording to "expanding window" throughout. We kept the term "rolling nowcast", as we believe this is still an accurate description of our analysis. We nowcast epidemic growth rates in 2-week windows; the nowcast is on a rolling basis, but we agree that the training windows are expanding.

5. Choice of Prediction Window Size:

The authors chose a two-week prediction window. Is there a specific reason for selecting this duration? Using shorter window sizes might improve the results, especially during periods of high growth rates. The rationale behind the chosen window size should be discussed.

We thank the reviewer for this important point. We chose the window size arbitrarily to demonstrate how the prediction models can be retrained on a rolling basis but without re-fitting the model each day, which would be computationally demanding. We agree that in practice a public health authority might retrain the model more frequently, perhaps even daily. But the choice of updating frequency would depend on many practical factors such as the resources available, the amount and timeliness of new data, and the desired frequency for updating nowcasts or predictions. For example, large testing volume might permit daily updates, whereas sparse testing might require updates bi-weekly or less.

We have added the following text to the methods section in "*Evaluating model performance*":

A 2-week window was chosen to demonstrate how models could be updated on a rolling basis. In practice, the appropriate window would depend on multiple factors such as computational capacity and frequency of data updates. Shorter windows could improve predictive performance, by allowing faster detection of changes in the Ct-growth rate relationship.

6. Down-Sampling Analysis Reporting:

The down-sampling analysis results are reported as single point estimates. Given that 100 random samples were used, it would be more informative to report the results with error bars to reflect variability and uncertainty.

Thank you for this suggestion. We modified the results reported in the figure and table to show ± 1 s.d. error bars, reflecting the variation between instantiations of the randomly downsampled data. Where applicable, we included error bars on the comparison results as well.

7. Figure Improvements:

In Figure 6 on page 18, the text "Alternate" is not fully displayed.

We corrected this in the figure.

The second row of the plot lacks an explanation for the light grey bars. Please provide a legend or description to clarify their meaning.

We added the explanation to the figure caption:

Grey bars indicate the number of days of data included for the various downsampling analyses; the random downsampling can reduce some days below the 10-sample inclusion threshold, hence the variation

Reviewer #2 (Remarks to the Author):

I co-reviewed this manuscript with one of the reviewers who provided the listed reports. This is part of the Nature Communications initiative to facilitate training in peer review and to provide appropriate recognition for Early Career Researchers who co-review manuscripts.

Reviewer #3 (Remarks to the Author):

This manuscript investigates whether viral load data, measured by Ct values, can be used to predict epidemic trends. The manuscript is clearly written. However, I have a few major comments:

1. While the idea of using community-based Ct values for outbreak surveillance is interesting, I have reservations about its practical utility. Unlike other population-level measures such as wastewater viral concentrations, this approach requires individual-level sampling and could therefore be heavily influenced by individual-level heterogeneity. Although the temporal dynamics of viral load may follow similar trends over the course of infection (e.g., declining after peak viral load), the absolute Ct values could vary significantly among patients at similar disease stages (such as during the peak of shedding). Capturing this variation may be challenging.

These are valid concerns, and we do not claim that community-based Ct surveillance will necessarily outperform other population-level metrics including wastewater surveillance. Our study demonstrates that community-based Ct values hold useful but imperfect information on epidemic growth rates, as is the case for any other data stream, including case counts, wastewater, internet search trends, etc. As with all these systems, there are settings where some will be suitable but not others. For example, wastewater testing can be sensitive to environmental fluctuations and the flow of a particular system, and wastewater is not available in settings that solely use septic tanks. Infectious disease surveillance systems benefit from multiple, orthogonal data streams, as different data streams might fail or succeed at different times. Our argument is that Ct-based metrics can be an informative, additional source of epidemiological information. We have reframed the Introduction and Discussion, as well as reorganized the analyses presented, to better reflect that argument.

Regarding heterogeneities in viral load, these are good points, and we updated our discussion accordingly. The specific point about capturing variation across patients at similar disease stages is explored by one of our synthetic scenarios ("increased variation"), where we find that increasing the variation in Ct values for a given time-since-infection only slightly decreases predictive performance. The extent to which this variation is problematic will depend on the sample size used for training the model and performing predictions, demonstrated by our new sample size analyses. However, this heterogeneity becomes a problem if there are differences in viral kinetics between subgroups and the weighting of these subgroups in the data changes in an unknown way. For example, if peak viral load is higher in older age groups, and there is a shift in incidence towards older ages, then an untrained model might erroneously attribute higher population viral loads to a higher growth rate.

However, we note that such heterogeneity would also affect population-level metrics such as wastewater. Using the same example, consider that older individuals exhibit increased viral shedding in urine/faeces. If infections started increasing disproportionately in older adults, that might be reflected in an increase in wastewater viral load, in addition to the increase caused by higher prevalence. If the wastewater/Rt prediction model was trained over time periods with a different age composition of infections, then model predictions would likely be inaccurate.

In fact, data streams using individual-level data are arguably better equipped to detect these biases, as the data can be stratified by relevant covariates (e.g., constructing Ct-based metrics stratified by age group). We did not consider this here, as we had access to limited metadata for each sample due

to data privacy restrictions. But a real-life implementation of this system might be able to include covariates in the modelling analysis and thus account for the changing mixture of patients.

Added discussion text:

While the relationship between sampled viral loads, viral kinetics, and epidemic dynamics can be described mathematically under ideal conditions, in practice several issues complicate its application as an epidemic monitoring tool. Measured Ct values are determined by a combination of biological factors, such as individual-level heterogeneity, immunological history and infecting variant⁴⁵, and practical factors such as whether individuals are tested at a random point in their infection or around the time of peak viral load (prompted by symptom onset)⁴⁶, demography of the tested population⁴⁷, sample type^{30,33,34}, and RT-qPCR platform^{22,31,36,48}.

2. The authors present both RMSE (to measure the accuracy of predicting epidemic growth rates) and AUC (to measure the accuracy of predicting the epidemic direction). Interpreting the absolute value of the RMSE is difficult because it is generally used to compare the performance of competing models. For example, in line 115, the authors note that their in-sample RMSE is 0.00191, which is approximately 10% of the range in observed log incidence growth rates. It is unclear whether this 10% figure represents a good or poor performance without a clearer baseline for comparison.

Thank you for this suggestion. We have included two additional baseline estimates for comparison.

1. Using the growth rate of new positive tests in the LA County and hospital testing datasets to predict growth rate at the state level as a new baseline measure, described in the new results section, *Nowcasting state- or county-level epidemic growth rates combining Ct values and number of new positive tests*.
2. Using the mean growth rate over the preceding 2 weeks as the predicted growth rate for the following two weeks. New results text:

In comparison, a null model using the mean growth rate over the preceding two weeks as the prediction for the following two-week window gave an RMSE of 0.0336 and AUC of 0.770.

3. I may have missed it, but I do not see a clear definition of “epidemic growth rate” in the manuscript. We defined epidemic growth rate in the “Statistical methods” section of the Methods:

We calculated daily incidence-based epidemic growth rates as the natural log-transformed ratio of 7-day moving average new reported cases for each day to the 7-day moving average for the preceding day:

$$y_t = \ln \frac{\sum_{k=0}^6 f_{(t-k)}}{\sum_{k=1}^7 f_{(t-k)}}$$

where y_t is incidence growth rate and f_t is daily incidence at time t . We defined epidemic direction as growing when $y_t > 0$ and declining when $y_t \leq 0$.

4. The results suggest that the predictability of the epidemic growth rate is weak. For instance, the visual fit between observed and predicted rates appears poor in Figure 1, and most of the correlations between observed and predicted values are below 0.5.

We agree that the correlation between predicted and observed growth rates is poor in some places, and thus our Ct-based growth rate estimates are not always accurate. However, we argue that the model clearly has some predictive power despite the low correlation coefficients and provides additional information for situational awareness. Our additional analyses comparing the Ct-based model predictions against those based on growth rate of within-dataset positive tests demonstrates

this point, with Ct-based estimates outperforming within-dataset incidence-based estimates under certain circumstances. We do not claim that Ct-based growth rate estimation is the silver bullet of infectious disease surveillance but instead offer a rigorous evaluation of the contexts in which it can be helpful.

5. The AUC for predicting epidemic direction seems reasonably large, although it is a less interesting target.

We agree that an ideal surveillance system would predict infection growth rates accurately, rather than simply whether an epidemic is growing or declining. However, we argue that increased information on the direction of an epidemic is still valuable for decision-makers. Take, for example, when during the pandemic there was uncertainty as to whether a new wave was starting, scepticism as to whether infections were truly increasing or whether a “testing epidemic” was taking place, and detecting when a wave had reached its peak. Additional data to confirm true changes in growth rate would have been invaluable.

6. How does the performance of this approach compare to other methods, such as using wastewater data?

We feel that this is an apples-to-oranges comparison, as it will depend on the setting and data availability. Some states or settings may not have reliable wastewater testing available to them or only have a sewer system serving a subset of the population (e.g., Alaska). We hope that future studies that tackle comparisons to other methods will benefit from the comprehensive approach and results we present in our manuscript.

Overall, the manuscript is clearly written, but some key findings appear to be statistically insignificant, and the practical utility of the proposed approach might be limited.

Reviewer #4 (Remarks to the Author):

In this paper, the authors conducted a systematic and comprehensive study to investigate the feasibility of using SARS-CoV-2 Ct values to nowcast epidemic trajectories over three years of the COVID-19 pandemic, using both synthetic data and real-world data from Mass General Brigham hospital system and Los Angeles County.

While I do think the study provides new methods in using viral load data in epidemic nowcasting, my main concern is that whether this manuscript has added sufficiently novel and important insights to warrant publication in Nature Communications, especially that the ideas to use Ct data in nowcasting have been published in several papers by the same group earlier. Please see my specific comments below.

1. Introduction: It might be a bit assertive to state that “estimating epidemic growth rate based on the distribution of measured viral loads does not depend on the number of positive tests”, because the estimation relies on a representative sample of the viral load measurements from the population which would be more likely to get when there were more positive tests.

This is a fair point, and we updated the introduction accordingly:

Whereas count-based surveillance assumes that the number of reported positive tests tracks the true incidence of new infections, estimating epidemic growth rates using viral loads uses

information in the distribution of Ct values from a representative sample of infected individuals.

2. The estimation is dependent on the assumption that a lower Ct value suggested that an individual was sampled earlier during their infection, which was true for SARS-CoV-2. For other pathogens, if the differences in viral load between early and late stage of infection were not as large as SARS-CoV-2, how would the performance of the nowcasting change? Does it fit to any synthetic scenarios (e.g., asymmetric?) described in Table S1?

We highlight two changes to the manuscript to address the relationship between viral kinetics symmetry and nowcasting performance:

1. We added a synthetic scenario to the supplementary analyses assuming a fully symmetric viral kinetics curve with equal growth and clearance phase durations (5 days in each). These additional results show that the relationship between mean/skew of Ct values and growth rates is not sufficiently clear to allow good predictive performance.
2. To generalise the findings from our synthetic data analyses, we added a new section and analysis (new Figure 1) demonstrating the analytical (as opposed to simulated) relationship between mean population Ct values and the epidemic growth rate under different viral kinetics models.

More broadly, while the shape of the viral kinetics curve contributes towards the relationship between Ct values and epidemic growth rate, it does not entirely determine whether the method will work. The other component to the relationship is the time-since-infection distribution, which is essentially the incidence curve in the weeks prior to the observation. The method works because the average time-since-infection increases over the course of the outbreak (because the growth rate declines), which leads to a change in the average viral load. As long as this change in the average time-since-infection covers a period of monotonically declining or increasing viral load, then there will be information to detect a changing growth rate.

3. Following the comment above, how would the model performance change with varying shedding profiles/distributions of other respiratory pathogens?

This is an interesting research area. We did not simulate specific pathogens here, as this has been explored by colleagues at Hong Kong University (<https://doi.org/10.1101/2025.05.22.25328079>).

4. In practice, Ct data of earlier outbreaks were available for model training for later outbreaks. How would the performance of the MGB and LAC models change if Alpha and/or Delta data were used to train model for nowcasting of the Omicron outbreaks?

This is an important point. To the reviewer's specific example, our sensitivity analyses using a fixed train-test split date of 31 Dec 2021 are effectively testing the scenario described - the vast majority of data in the model training set then are from Alpha and Delta (and wild type) waves, with only the very beginnings of the first Omicron outbreak included in the training set, so the trained model is largely shaped by the earlier variants. The resulting nowcasts are slightly biased because, as would be expected in this situation, the effects of Omicron are incompletely accounted for.

We added a sensitivity analysis to explore the impact of variant-specific viral kinetics on our method. We updated our simulation to incorporate three distinct variants with differing viral kinetics, described in Supplementary Text 1 – synthetic datasets section 4. We then repeated our in-sample, fixed train/test split, and nowcasting analyses with these synthetic datasets, finding that predictive performance is very poor for the Omicron period (new Fig S13). This analysis also revealed that the relationship between Ct values and growth rates are very different between variants (new Fig S7), with different slopes and intercepts for each variant; standardizing the Ct values by viral variant (using the mean and variance of all Ct values from that variant) reduces these differences considerably,

resulting in more similar Ct-growth rate relationships across variants (new Fig S9). However, we note that standardizing Ct values by variant would not be possible in real time because a) not all samples are sequenced and thus cannot be categorised and b) we could only standardize our synthetic data using the entire time series of data.

Our models do incorporate information from earlier outbreaks to inform nowcasting of later ones, even across different variants. The inclusion of an additive term representing 'variant era' (i.e., approximate time period when a particular variant is most prevalent) in the GAMs means that information from earlier variants is still used, just with potentially different intercept terms representing the effects of a variant on the relationship between incidence growth rate and Ct value distribution. Estimation of that term for any given variant becomes more accurate as more data becomes available. In cases where a new variant has rapidly emerged but no data are yet available for model training (in the first few weeks of a new variant era), the model makes the conservative assumption that the preceding variant is still dominant.

5. For the current model, only the viral load or Ct data were used in the prediction. Would the model performance be improved if the incidence data were also incorporated?

Thank you for this suggestion, which we incorporated into the updated manuscript. Please see our response to Reviewer 1, Comment 2 for details.

6. It would be useful to readers and users of the nowcasting methods if the growth rate estimates could be translated to real-time reproductive numbers with a given generation time distribution. The changes in growth rates were a bit difficult to interpret – for example, the predicted direction and the growth rate matter a lot when reproductive number is much higher than one, but it might be less worrying when the reproductive number is fluctuating around one.

We note that it is not universally agreed whether the reproductive number or growth rate is the more useful metric. There is a linear relationship between the growth rate and R_t , provided a model is assumed for the generation time distribution (Parag et al. 2022, <https://doi.org/10.1111/rssa.12867>). We prefer to present the growth rate here, as the generation time distribution is likely to have changed repeatedly over the 3-year time period of our datasets. Thus, we feel that this would introduce additional assumptions, complexity, and potential for error if the generation time distribution is misspecified, which would complicate the analyses further.

Reviewer #4 (Remarks on code availability):

The methods are well documented, and the codes and synthetic data have been made available online.

Reviewer #5 (Remarks to the Author):

Lim et al. demonstrate a nowcasting method to predict SARS-CoV-2 epidemic growth rates based on cross-sectional Ct measurements. This follows up on the work by some of the same authors (Hay et al, Science, 2022), which first showed that Ct distributions contain relevant information to track epidemic dynamics.

This new work extends the previous, through its development of a nowcasting model to predict epidemic growth rates from the CT-distribution, some sensitivity analyses and more extensive verification on two datasets from the COVID-19 pandemic.

This manuscript is well-written and clearly developed with scientific rigour. However, our primary reservations are around the relative advance posed by this paper, and the independence of this work

with respect to the previous paper. These points may also affect the breadth of readers attracted by the manuscript. At present, we believe this manuscript may be most suited for a specialized audience of epidemiologists/biostatisticians. To aid such an audience in their use of the presented method, further sensitivity analyses and comparison to established techniques would be helpful.

Thank you for your positive appraisal of our manuscript. A similar question about the advances in this manuscript was made by other reviewers; please see our response to Reviewer 1, Comment 1.

1. (Line 77-94) The manuscript would benefit from a figure highlighting the intuition behind the relationship between CT value distribution and epidemic trajectory. While we realize that this has already been shown elsewhere, not including it here make the logic behind different modeling choices (e.g. a more or less symmetric distribution) harder to understand and possibly prohibitive for readers not familiar with the previous work.

As the reviewer mentions, the intuition behind the fundamental relationship between Ct values and epidemic trajectory is described elsewhere. To build on this intuition, we added a detailed exploration of factors that can influence the Ct value-epidemic trajectory relationship, as summarized in the new Fig. 1. We hope this helps develop a reader's intuition while also advancing the field.

2. The question that is set up to guide the advances made in this manuscript is: "it remains unclear under which conditions [population-level Ct values] are a practical source of epidemiological information. [...] Whether [a multitude of factors] are prohibitively confounding when using Ct value distributions for epidemic monitoring has yet to be explored. " (in lines 58-68)

We are not sure this question is actually comprehensively covered by the four non-ideal modeling scenarios studied here.

Thank you for pointing this out. We substantially modified our synthetic data analyses (and reframed the introduction accordingly) to better address these questions. We now first present analytical convolutions highlighting how the Ct-growth rate relationship—which underlies this nowcasting method—changes based on several of the factors mentioned. We then present simulation-based nowcasting results quantifying the impact of such factors on nowcasting accuracy. We believe this two-stage approach has allowed us to more thoroughly address the question of how various biological and logistical factors could impact the feasibility of Ct value nowcasting.

a) "increased symmetry in viral kinetics" is an arbitrary choice. Why not study this for a number of different choices of "increased symmetry" (a continuous scale)?

We agree that the distinction between symmetric and asymmetric is arbitrary. However, exploring the impact of symmetry in viral kinetics over a continuous scale of symmetry would be computationally intensive while providing limited additional insight, in our opinion. We added a viral kinetics symmetry scenario (symmetric viral kinetics) such that we now have three levels of viral kinetics symmetry that we explore analytically. We believe the comparison of these three scenarios will help develop readers' intuition and understanding for how these factors impact the Ct-growth rate relationship.

b) Variation in the observed viral load given a time-since-infection is assumed to only change the variance, but what if there are two underlying populations (e.g. symptomatic/asymptomatic, or old/young) that have two separate distributions? What if those underlying classes are correlated with the delay between symptom and testing too, or only arise during a specific phase in the epidemic?

This is an important limitation impacting the representativeness of any surveillance metric from routine testing, including Ct values. There is some evidence that symptomatic infections have slightly different viral kinetics than asymptomatic infections for SARS-CoV-2, but the effect is very small. For other pathogens where this difference may be bigger, stratifying the data by symptom status, as done here in some of the analyses, would be recommended.

Regarding the correlation between symptom onset and testing, that is already accounted for in the model, as we simulated symptom-based testing by i) simulating an incubation period for each individual and then ii) simulating a delay between symptom onset and testing for each individual. More complex correlations will indeed affect the Ct/growth rate relationship further. If this relationship changes over time, then the trained prediction model will become mis-specified and lead to errors. This is the case for all the factors we explored; whether or not this is a problem in practice depends on the degree of misspecification.

We did attempt to update our simulation framework to incorporate a correlation between peak viral load and symptom onset and to change the mixture of symptomatic/asymptomatic infections over time. However, this became an extremely complex analysis for two reasons:

1. We do not simulate individual-level viral kinetics trajectories, but draw Ct values from a population-level model describing the overall mean and distribution of Ct value over time-since-infection. This is appropriate for our analyses as we simulate at most one observation per individual, and are not interested in longitudinal trajectories. Thus we could not introduce correlation of individual-level parameters without rewriting the entire framework.
2. There are some data for the correlation between viral kinetics parameters, such as timing of peak viral load relative to symptom onset (see Russell et al. 2024 <https://doi.org/10.1371/journal.pbio.3002463>). However, any modelling decisions on other important factors such as the relative proportion of symptomatic/asymptomatic cases, or correlation between symptom onset and testing behaviour, would be arbitrary. We could test reasonable assumptions that would drastically affect the Ct/growth rate relationship, and other assumptions that would not.

We therefore decided not to include additional analyses. We believe that the overall idea is covered in the discussion: not accounting for all factors affecting the observed Ct/growth rate relationship can reduce model accuracy, depending on the magnitude of those effects. We have added text to the discussion that addresses the possibility of model misspecification that may arise if the relationship between Ct values and growth rate diverges from the model assumptions:

Ultimately, predictions may become biased when the relationship between observed Ct values and the epidemic growth rate differ between training and test data. For example, if there are strong, time-varying correlations between symptom status, symptom onset, testing delay, and a time-varying mixture of symptomatic vs. asymptomatic samples, then the prediction model may become mis-specified over time if not updated.

c) Why were the scenarios tested with only a single draw of a simulated epidemic scenario? What is the stochasticity in performance across different random seeds or slightly different initial values? Rather than simulating such a long dataset, we believe the focus should be on repeatability across the same scenario many times. One minimal way to approximate this might be to split up the current simulated dataset into timeperiods that were similar (e.g. going from low to high incidence) and comparing performance per class of scenarios.

We agree that presenting simulations from the full, multiwave epidemic make it difficult to draw clear conclusions. One reason the full incidence curve is useful to show is that it incorporates a wide range of growth rates and epidemic dynamics: focusing on a single epidemic wave would mask a lot of the complexity arising from waves continuously waning and increasing (e.g., sampled Ct values in a given week might be from infections straddling both an epidemic decrease and increase phase). Indeed, in our exploratory work (not shown), the nowcasting analyses from simpler simulations with two epidemic waves performed extremely well.

We felt that repeating the same scenario many times with different random seeds was not necessary, as the simulated datasets are so large that they already incorporate substantial stochasticity. Instead,

we added an analytical model (new Figure 1) using a simple SIR model to illustrate the theoretical relationship between Ct values and growth rates over a single epidemic wave. Rather than use repeated stochastic draws to incorporate uncertainty, we instead illustrated how different sample sizes lead to different levels of confidence in the population-level mean Ct value, and how this maps onto uncertainty in growth rate estimates. We felt that this was a comprehensive solution to exploring this point.

d) It is misleading that Fig 1 includes different colors for the different variants since no differences were assumed. Why didn't the authors simulate the impact of variant replacement dynamics and/or different Ct-time since infection relationships for the different variants?

We included the shading for better consistency with the subsequent real-world figures but agree that this was confusing. We removed the shading from the equivalent figure (now Figure 2). As suggested, we also added a sensitivity analysis to explore the impact of variant-specific viral kinetics on our method. Please see our response to Reviewer 4, Comment 4 for details.

3. In the two use cases on realistic data, how does the present nowcasting model (GAM with smoothing splines) differ from the SEIR model introduced in Hay et al 2022? The relevant comparison here seems to be against other Ct-based methods and/or against the RMSE of other transmission dynamic inference methods (e.g. Re estimation) on incidence data?

The Hay et al. `virosolver` method requires assumptions about an underlying viral kinetics model. That model explicitly captures the convolution between the incidence curve and the viral kinetics curve. In contrast, the nowcasting models presented here are purely statistical - they rely on the relationship between Ct distribution and growth rates that results from those viral kinetics, but no explicit representation of the kinetics is required. Thus, the two approaches are very different. We updated the discussion (paragraph beginning "Another challenge for modeling Ct value dynamics is the choice of mathematical model to capture the relationship between observed Ct values and underlying epidemic growth rates.") to place these two approaches in context.

Furthermore, to make it easier to compare model performance to more relevant comparators, we added several comparison models, incorporating the within-dataset growth rate of positive tests as a predictor, as well as the null model proposed below. These comparison models should allow readers a better sense of how the Ct-based models are performing.

In general it is extremely difficult to assign a sense of quality to the reported RMSE/AUC numbers. How well would a null model that just predicts the same average incidence growth rate as the past 2 weeks do?

As suggested, we added a null model using the mean growth rate over the preceding 2 weeks as the predicted growth rate for the following two weeks. New results text:

In comparison, a null model using the mean growth rate over the preceding two weeks as the prediction for the following two-week window gave an RMSE of 0.0336 and AUC of 0.770.

4. How do the RMSE and AUC differ and depend on the phase of the epidemic?

In response to this question, we did preliminary analysis (on the synthetic data nowcast) comparing RMSE and AUC in different epidemic phases: rapidly declining (true epidemic growth rate ≤ -0.025), declining ($-0.025 < GR \leq 0$), growing ($0 < GR \leq 0.025$), and rapidly growing ($0.025 < GR$). We found no systematic differences between the phases, other than larger RMSE for larger absolute growth rates, which is to be expected; see table below. As such we elected not to include further analysis of performance by epidemic phase in the results.

Model	RMSE, in- sample	RMSE, nowcas t	RMSE, inflectio n	AUC, in- sample	AUC, nowca st	AUC, inflecti on	RMSE, gr < - 0.025	AUC, gr < - 0.025	RMSE, - 0.025<=g r<0	AUC, - 0.025<= gr<0	RMSE, 0<=gr<0. 025	AUC, 0<=gr<0 .025	RMSE, gr>=0.0 25	AUC, gr>=0.0 25
Ideal condition	0.0184	0.0188	0.0234	0.936	0.926	0.95	0.0214	0.805	0.0162	0.766	0.0156	0.808	0.0216	0.851
Realistic kinetics	0.0319	0.0319	0.0378	0.788	0.757	0.842	0.0398	0.557	0.0243	0.522	0.0248	0.579	0.0375	0.793
Realistic variation	0.0238	0.0246	0.0303	0.862	0.839	0.87	0.0256	0.655	0.0224	0.586	0.0213	0.658	0.0289	0.793
Realistic sampling	0.0201	0.0209	0.0282	0.916	0.9	0.892	0.0214	0.803	0.0184	0.706	0.0177	0.732	0.0256	0.72
Baseline scenario	0.0352	0.0352	0.0452	0.729	0.682	0.745	0.0444	0.557	0.0244	0.511	0.0221	0.565	0.0457	0.662
High variation	0.039	0.0395	0.0483	0.645	0.559	0.548	0.0556	0.678	0.0263	0.592	0.0223	0.569	0.048	0.668
Symmetric kinetics	0.0337	0.0346	0.0415	0.718	0.664	0.735	0.0437	0.575	0.0251	0.521	0.0238	0.613	0.0432	0.749
Baseline, asymptomatic only	0.0351	0.036	0.0374	0.609	0.498	0.509	0.0519	0.515	0.0217	0.5	0.0191	0.523	0.0449	0.818
Realistic sampling, asymptomatic	0.0214	0.0226	0.0237	0.898	0.892	0.873	0.021	0.818	0.0198	0.675	0.0198	0.735	0.0293	0.821
Ideal condition	0.0181	0.0192	0.0231	0.936	0.926	0.95	0.0215	0.776	0.0167	0.766	0.0156	0.81	0.0225	0.875
Realistic kinetics	0.0309	0.0333	0.0398	0.798	0.752	0.836	0.04	0.639	0.0268	0.543	0.0266	0.575	0.0388	0.824
Realistic variation	0.0231	0.0251	0.0297	0.863	0.832	0.837	0.0253	0.654	0.0235	0.587	0.0228	0.645	0.0286	0.717
Realistic sampling	0.0195	0.022	0.03	0.92	0.893	0.863	0.0234	0.784	0.0186	0.713	0.0193	0.725	0.0263	0.686
Baseline scenario	0.0339	0.0348	0.0444	0.728	0.683	0.733	0.043	0.619	0.0262	0.506	0.0233	0.572	0.044	0.645
High variation	0.0368	0.0386	0.0471	0.678	0.618	0.652	0.0522	0.502	0.0273	0.503	0.0243	0.565	0.0468	0.76
Symmetric kinetics	0.0326	0.0355	0.0438	0.742	0.682	0.658	0.0422	0.546	0.028	0.519	0.0266	0.612	0.0434	0.649
Baseline, asymptomatic only	0.0345	0.0366	0.0367	0.649	0.558	0.505	0.054	0.453	0.023	0.468	0.0218	0.549	0.0424	0.565
Realistic sampling, asymptomatic	0.021	0.0234	0.0242	0.901	0.893	0.886	0.0209	0.82	0.0217	0.683	0.0219	0.735	0.0287	0.805

— Minor Points —

Line 105: please define what RMSE and AUC stand for (abbreviation). Also specify which area under which curve you are measuring.

We added the definitions in the introduction.

Line 168/169: what do the authors think could explain the difference in lagging/leading the incidence growth rates?

We are not sure. We are cautious and seek to avoid overinterpreting these leads/lags, as the results shown are from calculating the cross-correlation across the entire time series rather than investigating lags and leads for specific time periods.

For the MGB data, the Ct value trends appear to lag the state-level case curve substantially. This may be due to delays in the MGB testing and reporting delays relative to the cases being tested as part of the state-wide reported incidence curve; if infections being detected as part of the MGB dataset arise later than a randomly selected infection in the community, we might see this lag.

For the LAC data, the Ct trends appear to lead the count-level case curve, though the correlations are very small. As these are all municipal tests and reflect 10% of all municipal testing done in LAC during the sample period, it may be that these are a good and early representative of the overall incidence curve, though we emphasise that the correlation coefficients are very small (peaks of around -0.06).

Fig 2B) It may be slightly misleading that the blue line here represents the average while in the first figure it was the simulated truth.

Thank you for highlighting this. The simulation figures use a lighter blue which should hopefully make this distinction clear enough. Note the blue line in the simulation also represents the smoothed growth rate.

Fig 2C) what is the intuition behind the wildly different inferred mean/skewness relations for the different variants?

These are modelled relationships using quite flexible smoothing splines. Some of the differences may be due to identifiability issues and stochasticity, others may be because there are other factors affecting the Ct/growth rate relationship independent of different variant viral kinetics. The new simulation analyses comparing different variant waves (new Fig S7 and S9) should provide more intuition on the impact of the variant itself on these relationships.

Line 217 What is “the directional discrimination test”?

This refers to the same epidemic direction (logistic) prediction model; we rephrased to avoid confusion.

Fig 6: the top left label is squished. Also the abbreviation “TFT” and the grey vertical bars are not mentioned in the caption.

We corrected the label display and added explanations to the caption.

Fig S13: visually the performance of the trained model for the mean of the Ct distribution seems poor on the test data. How do the authors explain that the model still performed well?

We sought to clarify our comparisons of model performance, providing more points of comparison and a more nuanced assessment. We also clarified that our overall aim was not to demonstrate superiority of Ct-based inference, but rather to demonstrate its possible uses and limitations where reliable, representative case reporting numbers are not available.

Supplement line 235 “using an extremely left-skewed viral kinetics curve” - Please include a reference to table S7 here. A figure showing the different distributions would also help a lot.

[This figure is included as Fig S6.](#)

Reviewer #5 (Remarks on code availability):

We looked at the code very briefly. From a first glance it seems well organized and comprehensive. More careful review would be appropriate during a later round of revisions.

What we couldn't find but might be helpful is some pointers on how to re-plot the outcome of the synthetic data analysis without re-running the simulations.

[Unfortunately, it is not possible to upload the synthetic data outputs as the files exceed the GitHub maximum file size. Instead, we have uploaded smaller synthetic datasets which should allow the remainder of the analyses to be run without re-running the simulations. We have included instructions in the simulation README file on how to restore the simulation sizes to the values used in the manuscript.](#)

[In addition, the code is designed to be modular such that plotting outcomes can be done separately from generating analysis results. This can be done by altering the values in the control file ``/code/0_code_settings.R``, changing the various ‘update’ settings \(e.g. ``update_data``, ``update_nowcast``, etc.\) to FALSE.](#)

Reviewer #6 (Remarks to the Author):

I co-reviewed this manuscript with one of the reviewers who provided the listed reports. This is part of the Nature Communications initiative to facilitate training in peer review and to provide appropriate recognition for Early Career Researchers who co-review manuscripts.

REVIEWERS' COMMENTS

Reviewer #1 (Remarks to the Author):

The authors have revised the manuscript and addressed most of the concerns raised in the initial review. The updated version includes expanded explanations, additional analyses, and a clearer description of methods and results.

Thank you for the supportive remarks and constructive comments, which helped greatly improve the manuscript.

First, the revised manuscript articulates the novelty and connection to existing literature in the introduction.

Second, the revised version now includes a comparative analysis of three model types: Ct-only, case-only, and a combined model (see Table 2).

Third, regarding the use of confirmed case numbers as ground truth—despite potential bias introduced by time-varying reporting rates—the authors acknowledge this limitation and argue that the short prediction horizon may reduce the impact of such biases. This explanation is reasonable, although a sensitivity analysis could further substantiate the claim.

In terms of methodological clarity, the revised methods section now clearly distinguishes between the generalized additive models (GAMs) used to predict continuous growth rates and logistic GAMs used to classify epidemic direction. Although complete model formulas are not provided, the descriptions of link functions and variables in the Supplementary Materials are adequate. Moreover, residual analysis is appropriately included and suggests the models are well-behaved (see Figure S2).

Finally, the authors have addressed other technical and formatting issues raised previously. Notably, they improved the terminology (e.g., clarifying “expanding” vs. “rolling” windows), reported variability in down-sampling analyses, and enhanced figure legibility. The expanded discussion of limitations—particularly the impact of assay heterogeneity and individual-level variability in Ct values—is thoughtful and well integrated.

Reviewer #2 (Remarks to the Author):

I co-reviewed this manuscript with one of the reviewers who provided the listed reports. This is part of the Nature Communications initiative to facilitate training in peer review and to provide appropriate recognition for Early Career Researchers who co-review manuscripts.

Reviewer #2 (Remarks on code availability):

<https://github.com/gradlab/ct-nowcasting> page can not be opened.

We apologize for the confusion. We are updating the repository link for the final publication version.

Reviewer #3 (Remarks to the Author):

While I appreciate the authors' responses, I believe my earlier concerns have not yet been fully addressed.

We appreciate the reviewer's previous comments that Ct value nowcasting faces inherent limitations which affect its practical utility; the revisions and additions to the manuscript have sought to elucidate and quantify some of these limitations. Our contention is not that Ct value methods provide a universally superior way to track epidemic dynamics – indeed we show that their accuracy is often quite limited. We argue that the challenges of epidemic monitoring are such that having a greater range of potential monitoring tools is valuable, and as one such potential tool, Ct-based methods are worth understanding further. We believe our revised analyses have been appropriately framed with this context, with a fair assessment of their potential strengths, limitations, and use cases. The challenges of data sparsity and variation are addressable in theory, but we acknowledge that they may be challenging to solve in practice. Nonetheless, we feel that our work provides a comprehensive exploration of potential issues and scenarios for researchers developing biomarker-based models for tracking epidemics..

Reviewer #4 (Remarks to the Author):

Thank you for inviting me to review the revised manuscript. I believe the authors have adequately addressed my previous comments.

Thank you for your previous comments, which were helpful for improving the manuscript.

The revised manuscript provides a useful and informative comparison between nowcasting using case counts and nowcasting using Ct data, based on two real-world datasets (MGB and LCA). I concur with the earlier comments from Reviewer #1 and Reviewer #5. My primary concern remains regarding the relative novelty and the extent of advancement of the current findings in comparison to previous work by the co-authors (e.g., Hay et al.).

We believe that the in-depth exploration of factors affecting the relationship between Ct values and epidemic dynamics, and therefore the usability of the method under real-world conditions, provides a substantial advance over previous work. Our initial publication described the theory behind why there is a predictable relationship between population viral loads and epidemic growth rates (i.e., illustrated the idea), but it did not explore real-world limitations under different scenarios and datasets (i.e., implement the idea into a practical tool). We hope that where previous work initially proposed use of Ct values in this manner, this present manuscript can serve as more of a reference guide on the method, its strengths and limitations.

Reviewer #4 (Remarks on code availability):

I reviewed the codes in the first round of review.

Reviewer #5 (Remarks to the Author):

The authors have substantially revised the manuscript and succeeded in either addressing our concerns or arguing why this is not possible/necessary. We just have one follow-up question, which should not stand in the way of publication.

The new Fig. 1 is interesting and indeed helps with intuition. However, the dynamics in panel B, row 2, are not intuitive to us. Why do the fully symmetric viral kinetics lead to an inverse relationship between growth and Ct value? In this scenario the symmetry of the viral kinetics is not the only thing that changes, as also the date of peak Ct and the range of dates with Ct above 40 are affected. This will naturally change the timing between infection incidence and peak Ct, even without regard for the symmetry of the shedding. How can the authors be sure that the symmetry is the quantity that truly leads to the reported inverse relation between growth and Ct value?

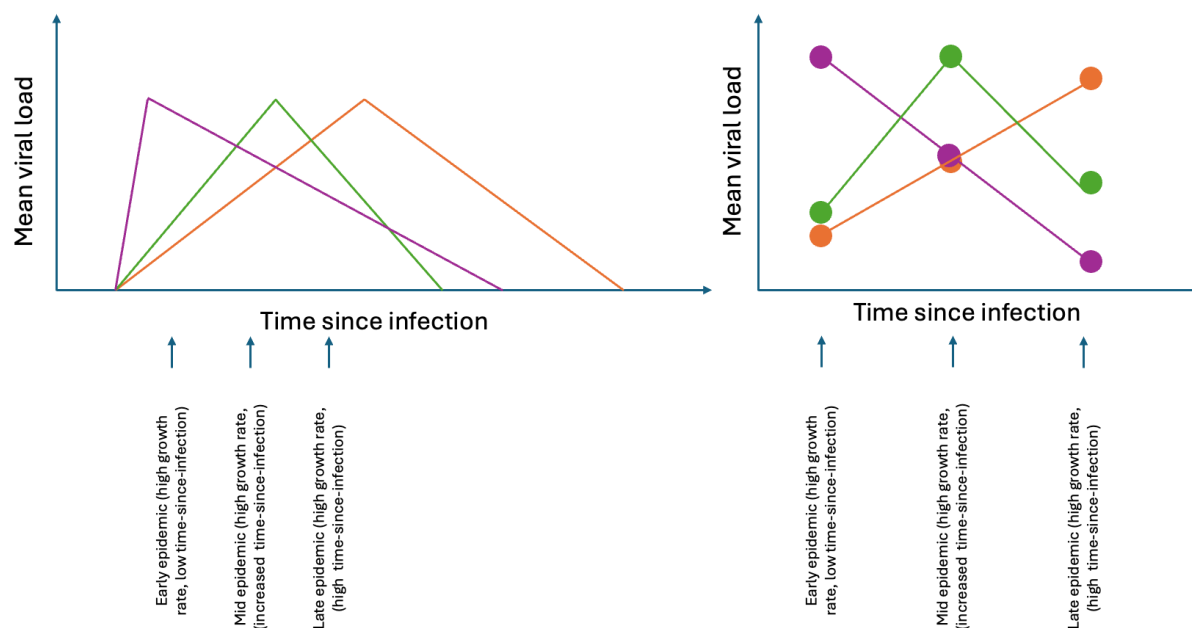
Thank you for this insightful observation. The key idea is that the epidemic growth rate shapes the distribution of time-since-infection in a cross-sectional sample. Since viral loads (Ct values) follow a predictable trajectory over infection, their distribution in a cross-sectional sample reflects this time-since-infection pattern of those individuals and thus the growth rate.

As you note, it is not symmetry itself that determines the Ct value/growth rate relationship, but whether the sampled time-since-infection range covers the viral growth phase, clearance phase, or sits across both. For example, if the time-since-infection range is entirely within the viral growth phase, then a decrease in growth rate leading to an increase in time-since-infection leads to monotonically increasing viral load. If the time-since-infection range sits entirely within the viral clearance phase (as with SARS-CoV-2), then the relationship flips as the increasing time-since-infection corresponds to monotonically decreasing viral load. However, if the time-since-infection range straddles the peak, as is the case in the highlighted scenario, then the relationship becomes non-monotonic, first increasing with viral growth and then decreasing with viral clearance.

We agree that the intuition is based around where the time-since-infection distribution sits along the viral kinetics curve, rather than whether the kinetics curve is symmetric. We have updated the text and figure to refer to “viral growth phase duration” rather than symmetry. We have updated the text to explain the point that what matters is where the time-since-infection range over the epidemic sits on the viral kinetics curve.

Changed text:

Figure 1 demonstrates the impact of viral kinetics properties (peak viral load, time from infection to peak viral load, individual-level variance) and sampling approaches (distribution of delays between infection and sampling) on the predictability of growth rates using mean Ct values. These results demonstrate several key ideas. First, in most scenarios, there is a monotonic relationship between mean Ct value and growth rate, providing a basis for predicting the latter from the former. This monotonic relationship arises because the epidemic growth rate determines the time-since-infection distribution in cross-sectional samples, which in turn determines the Ct value distribution. For instance, over the course of an outbreak with $R_0=1.5$, the mean time-since-infection ranges from 8 days to 25 days (assuming a maximum time-since-infection of 35 days), which sits entirely within the viral decline phase; as such the population-mean Ct value monotonically increases with decreasing growth rate. Conversely, if the time-since-infection range sat entirely within the viral growth phase (which could arise either from a much higher epidemic growth rate or a longer viral growth phase), the relationship would be reversed. Second, the precision of growth rates estimated from Ct values depends on the shape and gradient of this relationship. Because there is uncertainty in the true population mean Ct value from a finite number of observations, a steeper gradient will correspond to a wider range of compatible growth rates for a given set of measured Ct values. Third, biological and logistical factors influence the shape of the relationship and hence prediction accuracy, but in different ways. For example, increasing the duration of the viral growth phase decreases the precision of the estimated growth rate; a sufficiently long growth phase (producing a fully symmetric viral kinetics curve) results in a non-monotonic relationship and thus bi-modal estimates (Figure 1, second row, high scenario).



Minor typos:

- The introduction was updated to read “virosolver R package, which models the convolution of viral kinetics and incidence to predict population-level viral load distributions, to estimate infection incidence using Ct values from cross-sectional^{21,31,32,35,36,40}. “ but there is a word missing at the end of this sentence.

Thank you for catching this; the sentence should end with “cross-sectional samples”.

Reviewer #5 (Remarks on code availability):

The github folder was not accessible at the link provided. Instead the previous folder location <https://github.com/gradlab/ct-nowcasting-review> is still accessible. However, it was not updated for 10 months, so it is unclear whether this is the most current folder.

We apologize for the confusion. We are updating the repository link for the final publication version.

Reviewer #6 (Remarks to the Author):

I co-reviewed this manuscript with one of the reviewers who provided the listed reports. This is part of the Nature Communications initiative to facilitate training in peer review and to provide appropriate recognition for Early Career Researchers who co-review manuscripts.