Title

Statistical Challenges When Analyzing SARS-CoV-2 RNA Measurements Below the Assay Limit of Quantification in COVID-19 Clinical Trials

Journal

The Journal of Infectious Diseases, 228(Supplement_2)

ISSN

0022-1899

Authors

Moser, Carlee B
Chew, Kara W
Giganti, Mark J
et al.

Publication Date

2023-08-31

DOI

10.1093/infdis/jiad285

Peer reviewed

The Journal of Infectious Diseases
SUPPLEMENT ARTICLE

Statistical Challenges When Analyzing SARS-CoV-2 RNA Measurements Below the Assay Limit of Quantification in COVID-19 Clinical Trials

Carlee B. Moser,[1] Kara W. Chew,[2] Mark J. Giganti,[1] Jonathan Z. Li,[3] Evgenia Aga,[1] Justin Ritz,[1] Alexander L. Greninger,[4] Arzhang Cyrus Javan,[5] Rachel Bender Ignacio,[6] Eric S. Daar,[7] David A. Wohl,[8] Judith S. Currier,[2] Joseph J. Eron,[8] Davey M. Smith,[9] and Michael D. Hughes;[1,10] for the ACTIV-2/A5401 Study Team

[1]Center for Biostatistics in AIDS Research, Harvard T. H. Chan School of Public Health, Boston, Massachusetts, USA; [2]Department of Medicine, David Geffen School of Medicine, University of California, Los Angeles, Los Angeles, California, USA; [3]Department of Medicine, Brigham and Women's Hospital, Harvard Medical School, Cambridge, Massachusetts, USA; [4]Department of Laboratory Medicine and Pathology, University of Washington, Seattle, Washington, USA; [5]National Institutes of Health, Rockville, Maryland, USA; [6]Department of Medicine, University of Washington, Seattle, Washington, USA; [7]Lundquist Institute at Harbor-University of California, Los Angeles Medical Center, Torrance, California, USA; [8]Department of Medicine, Chapel Hill School of Medicine, University of North Carolina, Chapel Hill, North Carolina, USA; [9]Department of Medicine, University of California, San Diego, La Jolla, California, USA; and [10]Department of Biostatistics, Harvard T. H. Chan School of Public Health, Boston, Massachusetts, USA

Most clinical trials evaluating coronavirus disease 2019 (COVID-19) therapeutics include assessments of antiviral activity. In recently completed outpatient trials, changes in nasal severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) RNA levels from baseline were commonly assessed using analysis of covariance (ANCOVA) or mixed models for repeated measures (MMRM) with single imputation for results below assay lower limits of quantification (LLoQ). Analyzing changes in viral RNA levels with singly imputed values can lead to biased estimates of treatment effects. In this article, using an illustrative example from the ACTIV-2 trial, we highlight potential pitfalls of imputation when using ANCOVA or MMRM methods, and illustrate how these methods can be used when considering values <LLoQ as censored measurements. Best practices when analyzing quantitative viral RNA data should include details about the assay and its LLoQ, completeness summaries of viral RNA data, and outcomes among participants with baseline viral RNA ≥ LLoQ, as well as those with viral RNA < LLoQ.

Clinical Trials Registration. NCT04518410.

Keywords. COVID-19; SARS-CoV-2 RNA; linear regression for censored data; randomized trial.

Clinical trials designed to evaluate coronavirus disease 2019 (COVID-19) therapeutics should have clinically meaningful end points. Food and Drug Administration guidance states that clinical outcomes, such as the proportion of participants hospitalized or time to symptom recovery, are recommended as primary outcomes in phase 3 outpatient COVID-19 trials [1]. However, it also states that viral shedding should be measured to assess antiviral activity, primary virology outcomes are acceptable in phase 2, and quantitative and qualitative virological assessments are encouraged.

In typical COVID-19 randomized trials, samples such as nasopharyngeal swabs, anterior or midturbinate nasal swabs, oropharyngeal swabs, saliva, or plasma, are collected longitudinally for severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) RNA testing before and after intervention. Repeat sampling from early time points is common and in phase 3 trials typically includes 1 to 4 time points (Supplementary Table 1).

To evaluate virologic efficacy, SARS-CoV-2 RNA, henceforth called viral RNA (vRNA), is measured with quantitative reverse transcription polymerase chain reaction (RT-qPCR) assays. Like other nucleic acid assays, SARS-CoV-2 RNA assays have limits between which vRNA is accurately quantified, called the lower limit of quantification (LLoQ) and upper limit of quantification (ULoQ). For results > ULoQ, samples can be rerun with dilution to obtain quantifiable values. Assays may also indicate whether results <LLoQ are detectable or not.

Recent outpatient COVID-19 therapeutic trials considered various vRNA outcome measures and statistical methods. Most commonly, vRNA changes from baseline were analyzed using analysis of covariance (ANCOVA) at each time point or mixed models for repeated measures (MMRM). With these methods, single imputation was used to assign values for vRNA results <LLoQ (Supplementary Table 1) [2–18]. However, such imputation can introduce bias in estimating the magnitudes of treatment effects, as uncertainty for values <LLoQ is not captured [19].

Using an illustrative example from the Accelerating COVID-19 Therapeutic Interventions and Vaccines-2 (ACTIV-2) COVID-19 outpatient treatment trial, we describe bias that may arise when estimating treatment effects using

Correspondence: Carlee B. Moser, PhD, Center for Biostatistics in AIDS Research, Harvard T.H. Chan School of Public Health, 651 Huntington Avenue, FXB 513, Boston, MA 02115 (cmoser@sdac.harvard.edu).

The Journal of Infectious Diseases® 2023;228(S2):S101–10

https://doi.org/10.1093/infdis/jiad285

single imputation with ANCOVA and MMRM. This example was chosen to highlight statistical issues related to measurements below the assay LLoQ, including those that can arise due to large chance imbalances in levels between randomized arms at baseline. Drawing on the human immunodeficiency virus (HIV) literature [19], we describe and discuss alternative approaches for analyzing vRNA changes, that may be more appropriate by considering vRNA values <LLoQ as censored measurements. Finally, we provide recommendations for the analysis and presentation of results concerning vRNA changes in future trials.

METHODS

Description in this section is limited to aspects of trial design related to the ACTIV-2 illustrative example. As this manuscript aims to illustrate the impact of different statistical analysis approaches, further description of the analysis methods is integrated throughout the "Results" section.

ACTIV-2 (NCT04518410) is an adaptive platform trial designed to evaluate potential outpatient therapeutics for COVID-19 [20]. Our illustrative example includes 114 participants randomized to receive tixagevimab/cilgavimab intravenously or placebo; the primary results were previously reported [21]. Nasopharyngeal swabs were collected before treatment at day 0 (baseline) and days 3, 7, and 14 for SARS-CoV-2 RNA quantitative testing using a RT-qPCR assay with LLoQ of 2 $\log_{10}$ copies/mL [22]. All results > ULoQ were rerun with dilution to obtain quantifiable results. ACTIV-2 was approved by a central institutional review board (IRB), Advarra (Pro00045266), with additional local IRB review and approval as required by participating sites. All participants provided written informed consent.

As this manuscript aims to illustrate and discuss different approaches to analyze vRNA changes, we provide an overview in Table 1, but integrate descriptions of each method in the "Results" section. For methods that use imputed values for results <LLoQ, 2 commonly used single-imputation strategies (Supplementary Table 1) were assessed: LLoQ-imputation with impute values <LLoQ as the LLoQ; and ½LLoQ-imputation with impute values <LLoQ as ½ the LLoQ.

See Supplementary Methods for additional details on model specifications and sample SAS software code.

RESULTS

Descriptive summaries of vRNA across time points for the 114 participants are shown in Table 2 and Figure 1*A* and 1*B*. At baseline, 15 participants (13%) had missing vRNA (Supplementary Figure 1). There was a chance imbalance in vRNA between the randomized arms, with median vRNA in the active arm 1.0 $\log_{10}$ copies/mL higher than the placebo arm, and a higher proportion of participants with vRNA ≥ LLoQ (72% vs 62%).

Following the recommendation of Marschner et al [19], we separately considered data for participants with vRNA < LLoQ from those ≥LLoQ at baseline. For those with vRNA < LLoQ at baseline (n = 33), vRNA remained <LLoQ at all follow-up time points in both arms, suggesting peak vRNA may have been achieved before enrollment. For the remaining analyses, we focus on the 66 participants with vRNA ≥ LLoQ at baseline. The proportion with vRNA < LLoQ increased over time: 27% (8 of 30) and 28% (7 of 25) at day 3, 62% (18 of 29) and 54% (14 of 26) at day 7, and 93% (27 of 29) and 89% (24 of 27) at day 14 for the active and placebo arms, respectively (Table 2 and Figure 1*C* and 1*D*).

Analyzing vRNA at a Single Time point

Using Imputed Values Leads to Biased Estimates

Fifty-five (83%) of the 66 participants had vRNA results at day 3 (Supplementary Figure 1). Among the 11 participants without results, 1 was due to hospitalization on day 3; however, 6 were due to laboratory/specimen issues and 4 were due to site error/visit scheduling issues, likely unrelated to RNA level. For these 55 participants, at baseline there was a modest difference (0.33 $\log_{10}$ copies/mL) in mean vRNA: 5.61 and 5.28 $\log_{10}$ copies/mL for the active and placebo arms, respectively.

Focusing first on changes within arm, using LLoQ imputation, the mean vRNA at day 3 was 3.43 and 3.97 $\log_{10}$ copies/mL for the active and placebo arms, respectively, with estimated mean changes from baseline of −2.18 and −1.30 $\log_{10}$ copies/mL. Within each arm, the estimated mean changes are conservative (biased towards zero) because for participants with vRNA < LLoQ at day 3, the true changes are at least as large in magnitude as the imputed changes. Using ½LLoQ imputation gives mean changes that are larger (more negative) compared to LLoQ imputation: −2.45 and −1.58 $\log_{10}$ copies/mL for the active and placebo arms, respectively. This imputation still results in biased estimates within each arm, but with an unknown direction (estimated mean changes may be larger or smaller than the truth). For both approaches, the larger mean change in the active arm could reflect higher average baseline values, and thus larger changes are observable. Because the estimated mean changes within each arm are biased, the estimated difference between arms will be biased, and further bias may be introduced with the baseline imbalances.

The estimated difference in mean change for the active versus placebo arms at day 3 was −0.87 $\log_{10}$ copies/mL using LLoQ imputation and −0.86 $\log_{10}$ copies/mL using ½LLoQ imputation (Table 3, A). Although these estimates are similar, this may not be the case in other datasets when using the 2 approaches. By day 14, when approximately 90% of participants had vRNA < LLoQ (and hence had imputed changes), the estimated difference in mean change between arms was approximately equal to the baseline mean difference for both

Table 1. Summary of Analytic Methods Considered in Our Illustrative Example for the Analysis of Changes From Baseline in SARS-CoV-2 RNA

Methods	No. of Time Points	Handing Values <LLoQ	Advantages	Caveats/Issues
ANCOVA/linear regression	1	Single imputation	Easy to implement in standard software With small proportion <LLoQ, impact of imputation is likely modest	Using imputation results in biased estimates of differences between randomized arms in mean change Normality assumption in model may be violated Those with RNA <LLoQ at both time points will have change imputed as zero, which could be problematic with larger proportion <LLoQ at baseline
Linear regression for censored data (tobit regression)	1	Not required	Easy to implement in standard software Analyses considering censored measurements avoids bias that may be created by using imputed values	Normality assumption in model cannot be confirmed when large proportion of data are censored Those with RNA <LLoQ at both time points will have change imputed as zero, which could be problematic with larger proportion <LLoQ at baseline
Median regression for censored data	1	Not required	Easy to implement in standard software Distribution-free model removes assumptions about distribution of the errors	Model cannot be fitted when large proportion of data are censored Those with RNA <LLoQ at both time points will have change imputed as zero, which could be problematic with larger proportion < LLoQ at baseline
MMRM	>1	Single imputation	Easy to implement in standard software Global test of no difference between randomized arms across time points can be easily generated	Using imputation results in biased estimates of the difference between randomized arms in mean change, with the bias at 1 time dependent on the proportion <LLoQ at other times (as information is shared among times through an assumed correlation structure) Multivariate normality assumption may be violated Those with RNA <LLoQ at both time points will have change imputed as zero, which could be problematic with larger proportion <LLoQ at baseline
MMRM for censored data (LMEC)	>1	Not required	Analyses considering censored measurements avoids bias that may be created by using imputed values Global test of no difference between randomized arms across time points can be easily generated Possible improved precision by sharing information over time points through an assumed model	Increase complexity in implementing model in standard software as the number of time points increases Multivariate normality assumption difficult to verify, particularly when large proportion of data are censored at 1 or more times Those with RNA <LLoQ at both time points will have change imputed as zero, which could be problematic with larger proportion <LLoQ at baseline
Binary regression	≥1	Not required	Easy to implement in standard software Includes all participants, regardless of baseline value Estimation of treatment effects not influenced by the proportion <LLoQ	Loss of statistical power when dichotomizing outcome from continuous variable to a binary variable

Abbreviations: ANCOVA, analysis of covariance; LLoQ, lower limit of quantification; LMEC, linear mixed effects models with censored response; MMRM, mixed model repeated measures.

Table 2. Distribution of SARS-CoV-2 RNA by Study Visit in Each Treatment Arm in Overall Cohort and Among Those With vRNA ≥ LLoQ at Baseline/Day 0

	Active			Placebo		
Study Day	Median (Quartiles)	<LLoQ No. (%)	No. Missing	Median (Quartiles)	<LLoQ No. (%)	No. Missing
All participants in cohort (total n = 114, active n = 58, placebo n = 56)						
Baseline	4.0 (<LLoQ, 6.6)	14 (29)	9	3.0 (<LLoQ, 5.9)	19 (38)	6
Day 3	<LLoQ (<LLoQ, 3.9)	26 (52)	8	<LLoQ (<LLoQ, 3.9)	24 (52)	10
Day 7	<LLoQ (<LLoQ, 2.2)	37 (74)	8	<LLoQ (<LLoQ, 2.2)	35 (71)	7
Day 14	<LLoQ (<LLoQ, <LLoQ)	45 (96)	11	<LLoQ (<LLoQ, <LLoQ)	45 (97)	7
All participants with vRNA ≥LLoQ at baseline (total n = 66, active n = 35, placebo n = 31)						
Baseline	5.5 (3.7, 8.0)	0 (0)	0	5.0 (3.1, 6.7)	0 (0)	0
Day 3	3.0 (<LLoQ, 4.5)	8 (27)	5	3.4 (<LLoQ, 5.9)	7 (28)	6
Day 7	<LLoQ (<LLoQ, 2.5)	18 (62)	6	<LLoQ (<LLoQ, 3.3)	14 (54)	5
Day 14	<LLoQ (<LLoQ, <LLoQ)	27 (93)	6	<LLoQ (<LLoQ, <LLoQ)	24 (89)	4

Abbreviations: LLoQ, lower limit of quantification; vRNA, viral RNA.

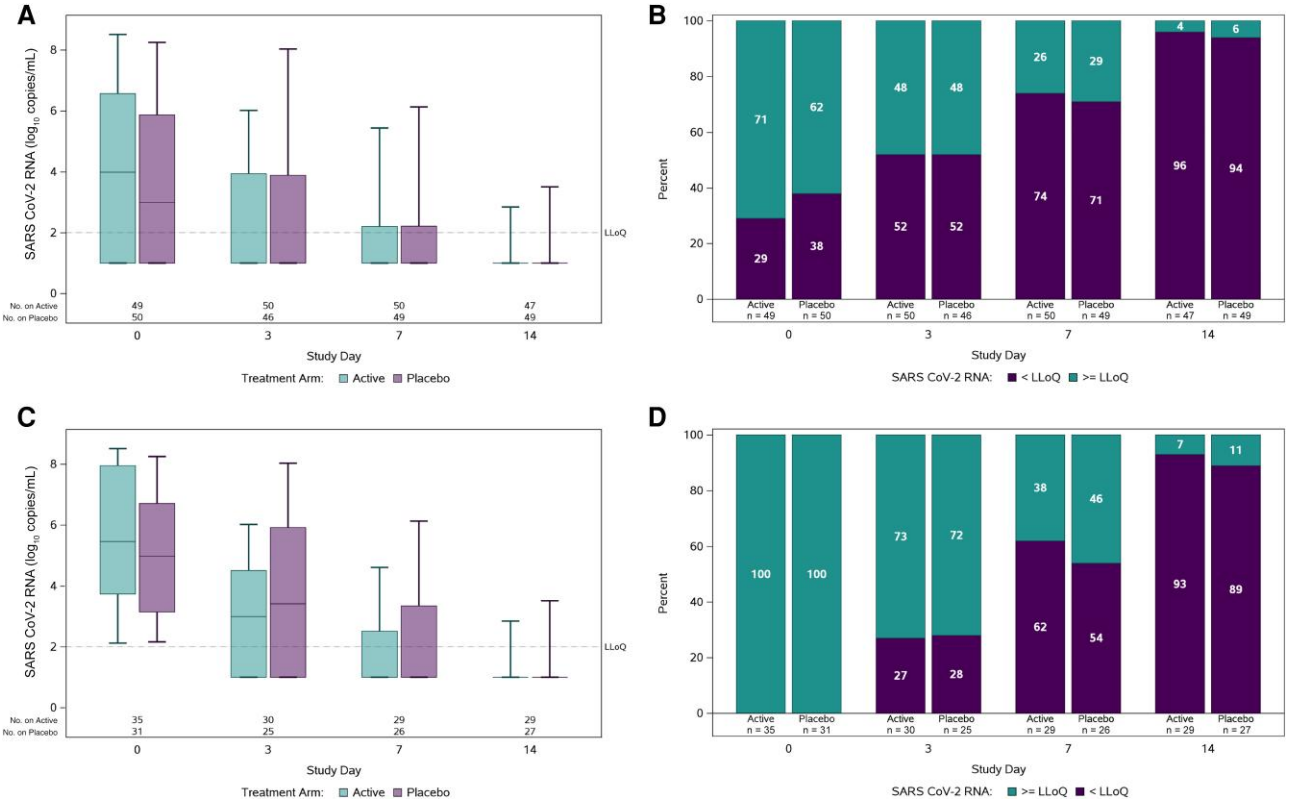

Figure 1. Distribution of SARS-CoV-2 RNA from nasopharyngeal swabs in active and placebo arms by study visit in overall cohort (*A* and *B*) and among those with vRNA ≥ LLoQ at baseline/day 0 (*C* and *D*). Levels of SARS-CoV-2 RNA (log_{10} copies/mL) with horizontal line = median, box = interquartile range, and whiskers = minimum/maximum (*A* and *C*); results below the LLoQ are plotted using an imputed value of 1 log_{10} copies/mL. Proportion with quantifiable vs unquantifiable SARS-CoV-2 RNA (*B* and *D*). Abbreviations: LLOQ, lower limit of quantification; SARS-CoV-2, severe acute respiratory syndrome coronavirus 2.

imputation approaches. If all participants had vRNA < LLoQ at day 14, the difference in mean change would equal the difference in mean vRNA at baseline, despite the choice of imputed value and underlying true difference. With larger proportions <LLoQ, differences between arms can reflect chance imbalances at baseline rather than true differences.

Adjusting for Baseline Can Help Address Baseline Imbalances

Although adjusting for baseline does not remove the bias in estimating differences between arms using singly imputed values, it may help reduce the impact of baseline imbalances in mean vRNA when assessing treatment effects.

The estimated differences in mean changes between arms using standard linear regression are shown in Table 3 (A and B). In adjusted analyses, differences between arms have some attenuation at each time point compared with unadjusted analyses, reflecting the adjustment for higher baseline vRNA levels in the active arm.

Analysis Methods Considering vRNA < LLoQ as Censored

Statisticians refer to vRNA values <LLoQ as being left-censored because if the true vRNA could be measured, it would take a value between zero copies/mL and LLoQ (ie, a value to the left of LLoQ). This contrasts with right-censoring like in survival analysis where, for example, participants alive at the end of follow-up have time of death greater than (to the right of) the time at the end of follow-up. Statistical methods used for survival analysis can be used to analyze vRNA data, with the small adaptation that values are left-censored rather than right-censored. Change in vRNA is defined as the difference in vRNA at the follow-up time minus the baseline. However, for follow-up vRNA values that are <LLoQ or left-censored, the change in vRNA is calculated as the LLoQ minus baseline vRNA, and is also left-censored.

Linear regression using software designed to handle censored data (known as tobit regression) is a possible method. Using this approach, adjusting for baseline vRNA, the estimated difference between arms in mean change from baseline to day 3 was −0.97 log_{10} copies/mL (95% confidence interval [CI], −1.81 to −.13) favoring the active arm (Table 3, C), and is somewhat larger than the differences in mean change by either imputation approach (Table 3, B). At day 7, the difference in mean change from baseline was −1.36 log_{10} copies/mL, also favoring the active arm (95% CI, −2.31 to −.41), which is much

Table 3. Differences Between Treatment Arms in SARS-CoV-2 RNA (log_{10} Copies/mL) Change From Baseline

Imputation	Day 3	Day 7	Day 14
A. Linear regression model with imputation, separate model by day—unadjusted			
LLoQ imputation	−0.87 (−1.70 to −.06) *P* = .037	−0.82 (−1.79 to .15) *P* = .09	−0.25 (−1.30 to .81) *P* = .64
½LLoQ imputation	−0.86 (−1.69 to −.04) *P* = .041	−0.90 (−1.82 to .01) *P* = .053	−0.29 (−1.32 to .74) *P* = .58
B. Linear regression model with imputation, separate model by day—adjusted for baseline			
LLoQ imputation	−0.74 (−1.41 to −.06) *P* = .034	−0.56 (−1.01 to −.11) *P* = .015	−0.06 (−.18 to .07) *P* = .38
½LLoQ imputation	−0.77 (−1.53 to .002) *P* = .050	−0.69 (−1.29 to −.09) *P* = .024	−0.11 (−.37 to .16) *P* = .42
C. Linear regression model for censored data (tobit regression), separate model by day—adjusting for baseline			
NA	−0.97 (−1.81 to −.13) *P* = .023	−1.36 (−2.31 to −.41) *P* = .005	Not obtained[a]
D. Median regression model for censored data, separate model by day—adjusting for baseline			
NA	−1.17 (−2.42 to .07) *P* = .07	−0.96 (NE to NE) NE	NE
E. MMRM across all 3 days (day 3, 7, and 14) with imputation—adjusting for baseline			
LLoQ imputation	−0.39 (−1.23 to .45) *P* = .36	−0.49 (−.95 to −.04) *P* = .032	−0.07 (−.20 to .06) *P* = .27
½LLoQ imputation	−0.52 (−1.44 to .40) *P* = .26	−0.60 (−1.21 to .01) *P* = .052	−0.13 (−.40 to .14) *P* = .33
F. MMRM across days 3 and 7 with imputation—adjusting for baseline			
LLoQ imputation	−0.65 (−1.36 to .07) *P* = .08	−0.58 (−1.01 to −.15) *P* = .009	…
½LLoQ imputation	−0.72 (−1.50 to .06) *P* = .07	−0.71 (−1.29 to −.13) *P* = .018	…
G. MMRM for censored data across days 3 and 7—adjusting for baseline			
NA	−1.10 (−1.94 to −.26) *P* = .011	−1.33 (−2.23 to −.43) *P* = .004	…

Data are mean (A–C, E–G) and median (D), (95% confidence interval), and *P* value among those with quantifiable baseline vRNA.

Abbreviations: LLoQ, lower limit of quantification; MMRM, mixed model for repeated measures; NA, not applicable; NE, not estimable.

[a]Results are not shown at day 14 for the linear regression model for censored data because model assumptions cannot be reasonably verified due to the high level of censoring at day 14.

larger than differences observed by either imputation approach, illustrating the potential bias using those methods when the proportion with vRNA < LLoQ increases. We did not pursue an analysis of mean changes to day 14 using tobit regression because of the high level of censoring (approximately 90%) and hence the inability to check model assumptions.

As with standard linear regression, there is an assumption that the errors in the model are normally distributed. These errors are estimated by the residuals calculated as the observed vRNA value minus the predicted model value. The distributional assumption can be evaluated with quantile-quantile (Q-Q) plots, comparing the quantiles of the observed distribution of the residuals (calculated using Kaplan-Meier methods to account for censored residuals) against the corresponding quantiles of a standard normal distribution. If the assumption was satisfied, the plots would show linear associations. Figure 2 shows Q-Q plots for the distribution of standardized residuals from models for change from baseline, adjusting for baseline. For the models of change from baseline to days 3 and 7, the Q-Q plots appear reasonably linear, supporting normality assumptions. We note, however, the more restricted range of the Q-Q plot for changes to day 7, as shown by the lack of standardized residuals below −1. This reflects the higher proportion of censored values at day 7; thus, the normality assumption cannot be verified for the tail of the distribution, corresponding to large negative changes from baseline.

Quantile Regression as an Alternative Distribution-Free Method

An alternative to tobit regression is quantile regression applied to assay-censored data, for example to model median change in vRNA. With this approach, there are no assumptions concerning the distribution of the errors in the model. However, there is an assumption that the median change has linear associations with continuous covariates in the model, including baseline vRNA.

At day 3, the adjusted difference between arms in median change from baseline was −1.17 log_{10} copies/mL (95% CI, −2.42 to .07) favoring the active arm. This is reasonably similar to the adjusted difference in mean change of −0.97 log_{10} copies/mL obtained from tobit regression, although estimated without making the assumption of normally distributed errors. There is a somewhat narrower CI for the difference in means versus difference in medians, reflecting the gain in precision from assuming a normal distribution for the errors. At day 7, the adjusted

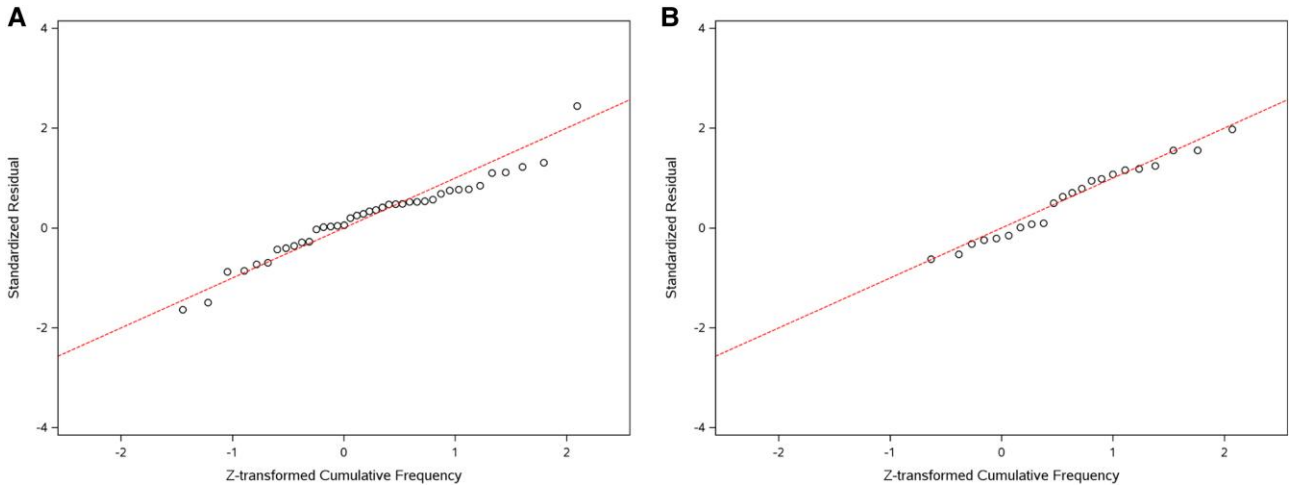

Figure 2. Quantile-quantile (Q-Q) plot for linear regression model for censored data for change in viral RNA from baseline to day 3 (*A*) and to day 7 (*B*); both models included an indicator variable for treatment versus placebo and adjusted for baseline viral RNA. Standardized residuals (for the noncensored observations) calculated by dividing the residuals by their standard deviation (estimated from the fitted model). Quantiles for a standard normal distribution plotted on the x-axis take account of censored residuals. Q-Q plots that show a linear association (data points falling along the diagonal line in a linear fashion) reflect that the normality assumption is reasonable. If the points depart markedly from the line, this implies the data are not normally distributed and may have outliers, are skewed (left or right), or are under- or overdispersed.

difference in median change was −0.96 $\log_{10}$ copies/mL, also favoring the active arm. However, it was not possible to obtain a CI from the numerical methods used to fit the model, due to the high proportion of participants with vRNA < LLoQ at day 7. At day 14, the higher proportion with vRNA < LLoQ meant the difference in median change between arms could not be estimated.

Analyzing Repeated vRNA Over Time

Imputed Values Can Affect Estimates From MMRM Due to Correlation Structure

Another strategy in several recent COVID-19 trials has been to use an MMRM with single imputation for vRNA values<LLoQ [2–14]. These models estimate the difference in mean vRNA change in each arm at each time point, in a similar manner to linear regression models fit separately by time point. However, MMRMs incorporate a stronger assumption about the distribution of errors across time points, specifically that they follow a multivariate normal distribution with a specified correlation structure. Using this assumption, a global test evaluating the null hypothesis of no difference between arms in vRNA change at any time point can be undertaken. The stronger assumption may provide improved precision in estimating the differences in mean change at each time point by borrowing information between time points. However, this assumption may not be appropriate when using singly imputed values for measurements <LLoQ as the correlation structure is affected by imputation. As an example, participants with vRNA < LLoQ at days 7 and 14 will have identical imputed changes at both time points leading to higher correlations of errors in the model than if the actual values < LLoQ were observed. To illustrate the impact of this, Table 3 (E) shows results from MMRMs for changes from baseline to days 3, 7, and 14. Compared with the estimates from models fitted separately at each time point (Table 3, B), the borrowing of information through the correlation structure leads to smaller estimated differences in mean change between arms, particularly at day 3 and to a lesser extent at day 7 for both imputation approaches. This attenuation is driven by including day 14, where approximately 90% of participants had vRNA < LLoQ; removal of this time point from the MMRM reduces the magnitude of the attenuation (Table 3, F). The estimates remain biased, however, for the same reasons as those obtained from separate regression models at each time point.

Extensions to MMRM that account for censored data exist (also known as linear mixed effects models for censored responses [LMEC]), but still require the multivariate normality assumption [23, 24]. A caveat with these models is that they can be difficult to implement in standard statistical software, especially as the number of time points increases. Estimated differences between arms in mean change from baseline to days 3 and 7 from LMEC are shown in Table 3 (G). The estimates are similar to those from the tobit regression models fitted separately at days 3 and 7 (Table 3, C). The stronger multivariate normal assumption leads to small gains in precision at day 7 as seen by the narrower CI, although the gain at day 3, where there is less censoring, is negligible. As with the separate regression models, we did not pursue LMEC over the 3 days, as the high level of censoring at day 14 meant that a normality assumption could not be reasonably verified.

Analyzing Proportion of Participants With vRNA < LLoQ Over Time

Strategies That Do Not Rely on Quantitative Values May Be Preferred With Large Percent <LLoQ

When there is a high proportion of participants with vRNA < LLoQ at 1 or more time points, it may be more appropriate to focus on how this proportion changes with time. This could be analyzed over time using log-binomial models fit using generalized estimating equations (GEE). However, due to problems with numerical algorithms, in ACTIV-2 we used Poisson regression models modified for binary outcomes [25] fit using GEEs with independence working correlation structure and robust standard errors, adjusting for baseline vRNA. When implementing this model across the 3 days, the proportion with vRNA < LLoQ did not differ between arms (Supplementary Table 2). When excluding the day 14 measurements, where approximately 90% of participants had vRNA < LLoQ, the results for days 3 and 7 were almost identical, confirming this method is not sensitive to including time points with high proportions <LLoQ. This strategy can lead to loss in statistical power compared to analyses of quantitative vRNA, so is best reserved for when high proportions of participants are expected to have vRNA < LLoQ at 1 or more time points. However, there is also no need to restrict the analysis population to participants with vRNA ≥ LLoQ, potentially providing more comprehensive analyses of qualitative vRNA in the overall study population.

DISCUSSION

In this article we summarize methods commonly used in outpatient COVID-19 therapeutic trials for analyzing quantitative changes in SARS-CoV-2 RNA over time, and through an illustrative example from the ACTIV-2 study, highlight potential pitfalls. In ACTIV-2, our primary virology analyses focused on comparing the proportion of participants with vRNA < LLoQ over time, and examined vRNA levels rather than changes. As the pandemic has evolved and we have learned more about viral trajectories and variability, so has our thinking about the best analytic strategy. Since designing ACTIV-2, we have implemented exploratory analyses examining treatment effects on changes in vRNA over time using tobit regression models with adjustment for baseline RNA, restricted to participants with baseline vRNA ≥ LLoQ, a method we advocate for in this article [19, 26, 27].

In our illustrative example, the primary focus was on the population with quantifiable vRNA at baseline, which has been a focus in recent COVID-19 studies. This was reasonable in our analysis as none who were<LLoQ at baseline had quantifiable vRNA at later time points. Including these individuals in analyses using imputed values would have led to imputed changes of zero and likely attenuation of the estimated mean changes. Regression analyses for censored data are more complex if such individuals are included, requiring strong, unverifiable assumptions about the distribution of vRNA changes over time among those with baseline vRNA < LLoQ. Looking more broadly across the study population in phase 2 placebo-controlled evaluations in ACTIV-2 (n = 1565 enrolled with a median of 6 days from symptom onset), we observed that only 14% (of 287) of those with vRNA < LLoQ at baseline later had quantifiable vRNA. As new studies are developed, potentially with enrollment closer to onset of symptoms, the decision to exclude those <LLoQ at baseline should be carefully scrutinized, as doing so could remove individuals on an upward viral load trajectory and we lack understanding of these trajectories in the setting of vaccination, reinfection, and emergent variants. At a minimum, documenting viral shedding changes among participants with baseline vRNA < LLoQ is important, and analyses stratified by level (<LLoQ and ≥LLoQ at baseline) might be pursued. Our illustrative example also highlights some of the additional complexities that arise when there is a substantial chance imbalance between randomized arms in mean vRNA levels at baseline. Such imbalances are not uncommon because of the variability in sampling when taking nasal swabs, which in smaller trials can make treatment effects more difficult to detect. Trials should be designed with sufficient sample size to reduce the impact of such imbalances on power and precision in estimating treatment effects of the anticipated proportion of vRNA values below the LLoQ [28]. Timing of measurements during follow-up is also important in considering power and precision, but should take account of population characteristics such as symptom duration at enrollment and anticipated speed of an antiviral affect based on pharmacokinetic properties of an agent and its mode of administration.

The methods considered in this article are not exhaustive of imputation or modeling strategies, but were chosen to align with methods from recent publications of COVID-19 trials. We focus on single imputation, and do not evaluate the performance of multiple-imputation strategies, which are more complicated and rely on distributional assumptions to support the imputation, but may reduce potential biases with imputation highlighted in this article [29, 30]. We also have not evaluated the statistical performance of these methods through formal simulation studies, which may add further insights to benefits or downsides of the analytic strategies, particularly when high proportions of participants have vRNA < LLoQ during follow-up, where verification of model assumptions becomes more difficult. We also have not considered potential biases due to missing data, for example, missingness arising due to hospitalization (which only affected 1 participant at each of days 3 and 7 in our example), if hospitalized participants have higher vRNA levels. Finally, analysis of vRNA changes among participants with baseline levels above a threshold (eg, the LLoQ) leads to estimated mean changes within each arm that are affected by

regression to the mean, although inference on the differences between randomized arms are not anticipated to be affected provided that the analysis adjusts for baseline level. Despite these limitations, our article highlights key issues and considerations when analyzing SARS-CoV-2 RNA data from outpatient treatment trials. These methods are not only applicable in the COVID-19 setting, but should be considered when analyzing any biomarker that is measured with an assay with an LLoQ.

Recommendations

The best practices in analyzing SARS-CoV-2 RNA from outpatient trials depend on the number of time points and proportion of results <LLoQ. Regardless of the planned analysis, some key details should be reported to facilitate interpretation.

1. Provide sufficient details of the RT-qPCR assay, including the LLoQ.
2. Explain who is included in the analysis, such as via a CONSORT-type diagram (see Supplementary Figure 1), including an accounting of missing data and the reasons for missing (eg, death, hospitalization, loss to follow-up, sample not obtained, sample processing/shipping issue).
3. If restricting the analysis population to those with quantifiable baseline vRNA, describe outcomes among those with vRNA < LLoQ.
4. Although we do not recommend the use of single imputation, if used, the choice of imputed values should be provided, and implications for interpretation of results discussed.
5. Include descriptive summaries of vRNA by treatment arm and time point. We suggest including 2 figures (see Figure 1): distributions of quantitative levels (eg, box and whisker plots) and distribution of vRNA categories (eg, <LLoQ vs ≥LLoQ).

Analytic strategies to estimate differences between arms we recommend are:

1. Methods that address censoring without imputation, such as tobit or median regression, or LMEC [23, 24] should be prioritized. But with increased censoring:
   (a) Normality assumptions underlying regression analysis for censored data cannot be evaluated over the full range of the distribution, and dropping time points with high levels of censoring from analysis may be appropriate.
   (b) Differences in medians (and their CIs) between arms might not be estimable from quantile regression.
2. Alternatively, consider nonparametric tests to analyze quantitative vRNA, such as the censored version of the Wilcoxon test (Gehan-Wilcoxon), which is implementable in standard software as a stratified test to account for baseline vRNA.
3. Comparing the proportion of participants with vRNA < LLoQ between arms over time may be preferred if there are high amounts of censoring.
4. With early, frequent measurements (eg, daily), more complex extensions of LMEC that evaluate viral dynamics (eg, estimating initial increases and subsequent vRNA decay) [20, 31–34], or time-to-viral clearance via methods for time-to-event data [4–7, 10, 35, 36] might be used.

Supplementary Data

Supplementary materials are available at *The Journal of Infectious Diseases* online. Consisting of data provided by the authors to benefit the reader, the posted materials are not copyedited and are the sole responsibility of the authors, so questions or comments should be addressed to the corresponding author.

Notes

Acknowledgments. We thank the study participants, site staff, site investigators, and the entire ACTIV-2/A5401 study team; the AIDS Clinical Trials Group, including Lara Hosey, Jhoanna Roa, and Nilam Patel; the ACTG Laboratory Center, including Grace Aldrovandi and William Murtaugh; Frontier Science, including Marlene Cooper, Howard Gutzman, and Kevin Knowles; the Harvard Center for Biostatistics in AIDS Research and ACTG Statistical and Data Analysis Center including Ronald Bosch and Linda Harrison; the National Institute of Allergy and Infectious Diseases/Division of AIDS; Bill Erhardt; the Foundation for the National Institutes of Health and the ACTIV partnership, including Stacey Adams; and the PPD clinical research business of Thermo Fisher Scientific. We also thank AstraZeneca for partnership in the ACTIV-2 study and for donation of study drug for the cohort used as the illustrative example in this analysis.

Disclaimer. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

Financial support. This work was supported by the National Institute of Allergy and Infectious Diseases of the National Institutes of Health (grant numbers UM1 AI068634 to M. D. H., UM1 AI068636 to J. S. C., and UM1 AI106701 to G. A.).

Supplement sponsorship. This article appears as part of the supplement "Findings From the ACTIV-2/AIDS Clinical Trials Group A5401 Adaptive Platform Trial of Investigational Therapies for Mild-to-Moderate COVID-19," sponsored by the National Institutes of Health through a grant to the University of California, Los Angeles.

Potential conflicts of interest. C. B. M. participated on a data safety monitoring board for the BONE STAR study. K. W. C. has received research funding to the institution from Merck Sharp & Dohme; and is a consultant for Pardes Biosciences. J. Z. L. has consulted for Abbvie; and received research support from Merck. A. L. G. reports contract testing

from Abbott, Cepheid, Novavax, Pfizer, Janssen, and Hologic; and research support from Gilead. R. B. I. received consulting fees from SeaGen and Resverlogix unrelated to this work; and research funding through her institution from Novartis and Ascentage Pharma Group. E. S. D. receives consulting fees from Gilead Sciences, Merck, and GSK/ViiV; and research support through the institution from Gilead Sciences and GSK/ViiV. D. A. W. has received funding to the institution to support research; and honoraria for advisory boards and consulting from Gilead Sciences. J. S. C. has consulted for Merck and Company. J. J. E. is an ad hoc consultant to GSK/VIR; and data monitoring committee chair for Adagio phase 3 studies. D. M. S. has consulted for Evidera, Fluxergy, Kiadis, Linear Therapies, Matrix BioMed, Arena Pharmaceuticals, VxBiosciences, Model Medicines, Bayer Pharmaceuticals, Signant Health, and Brio Clinical. All other authors report no potential conflicts.

All authors have submitted the ICMJE Form for Disclosure of Potential Conflicts of Interest. Conflicts that the editors consider relevant to the content of the manuscript have been disclosed.

References

1. U.S. Food and Drug Administration. COVID-19: developing drugs and biological products for treatment or prevention [Internet]. **2021**. https://www.fda.gov/regulatory-information/search-fda-guidance-documents/covid-19-developing-drugs-and-biological-products-treatment-or-prevention. Accessed 30 October 2022.
2. Weinreich DM, Sivapalasingam S, Norton T, et al. REGEN-COV antibody combination and outcomes in outpatients with Covid-19. N Engl J Med **2021**; 385:e81.
3. Weinreich DM, Sivapalasingam S, Norton T, et al. REGN-COV2, a neutralizing antibody cocktail, in outpatients with Covid-19. N Engl J Med **2021**; 384:238–51.
4. Dougan M, Azizad M, Mocherla B, et al. A randomized, placebo-controlled clinical trial of bamlanivimab and etesevimab together in high-risk ambulatory patients with COVID-19 and validation of the prognostic value of persistently high viral load. Clin Infect Dis **2021**; 75: e440–9.
5. Dougan M, Nirula A, Azizad M, et al. Bamlanivimab plus etesevimab in mild or moderate Covid-19. N Engl J Med **2021**; 385:1382–92.
6. Dougan M, Azizad M, Chen P, et al. Bebtelovimab, alone or together with bamlanivimab and etesevimab, as a broadly neutralizing monoclonal antibody treatment for mild to moderate, ambulatory COVID-19. medRxiv, doi: 10.1101/2022.03.10.22272100v1, 12 March **2022**, preprint: not peer reviewed.
7. Fischer WA, Eron JJ, Holman W, et al. A phase 2a clinical trial of molnupiravir in patients with COVID-19 shows accelerated SARS-CoV-2 RNA clearance and elimination of infectious virus. Sci Transl Med **2021**; 14:eabl7430.
8. Montgomery H, Hobbs FDR, Padilla F, et al. Efficacy and safety of intramuscular administration of tixagevimab-cilgavimab for early outpatient treatment of COVID-19 (TACKLE): a phase 3, randomised, double-blind, placebo-controlled trial. Lancet Respir Med **2022**; 10:985–96.
9. Gupta A, Gonzalez-Rojas Y, Juarez E, et al. Effect of sotrovimab on hospitalization or death among high-risk patients with mild to moderate COVID-19: a randomized clinical trial. JAMA **2022**; 327:1236–46.
10. Jagannathan P, Andrews JR, Bonilla H, et al. Peginterferon lambda-1a for treatment of outpatients with uncomplicated COVID-19: a randomized placebo-controlled trial. Nat Commun **2021**; 12:1967.
11. Alemany A, Millat-Martinez P, Corbacho-Monné M, et al. High-titre methylene blue-treated convalescent plasma as an early treatment for outpatients with COVID-19: a randomised, placebo-controlled trial. Lancet Respir Med **2022**; 10:278–88.
12. Reis G, Moreira Silva EAS, Medeiros Silva DC, et al. Early treatment with pegylated interferon lambda for COVID-19. N Engl J Med **2023**; 388:518–28.
13. Ramakrishnan S, Nicolau DV, Langford B, et al. Inhaled budesonide in the treatment of early COVID-19 (STOIC): a phase 2, open-label, randomised controlled trial. Lancet Respir Med **2021**; 9:763–72.
14. Mitjà O, Corbacho-Monné M, Ubals M, et al. Hydroxychloroquine for early treatment of adults with mild Covid-19: a randomized-controlled trial. Clin Infect Dis **2020**; 73:e4073–81.
15. Hammond J, Leister-Tebbe H, Gardner A, et al. Oral nirmatrelvir for high-risk, nonhospitalized adults with Covid-19. N Engl J Med **2022**; 386:1397–408.
16. Bernal A J, da Silva MM G, Musungaie DB, et al. Molnupiravir for oral treatment of Covid-19 in nonhospitalized patients. N Engl J Med **2022**; 386:509–20.
17. Gottlieb RL, Vaca CE, Paredes R, et al. Early remdesivir to prevent progression to severe Covid-19 in outpatients. N Engl J Med **2022**; 386:305–15.
18. Mukae H, Yotsuyanagi H, Ohmagari N, et al. Efficacy and safety of ensitrelvir in patients with mild-to-moderate coronavirus disease 2019 (COVID-19): the phase 2b part of a randomized, placebo-controlled, phase 2/3 study. Clin Infect Dis **2022**; 76:1403–11.
19. Marschner IC, Betensky RA, DeGruttola V, Hammer SM, Kuritzkes DR. Clinical trials using HIV-1 RNA-based primary endpoints: statistical analysis and potential biases. J Acquir Immune Defic Syndr Hum Retrovirol **1999**; 20: 220–7.
20. Chew KW, Moser C, Daar ES, et al. Antiviral and clinical activity of bamlanivimab in a randomized trial of non-

hospitalized adults with COVID-19. Nat Commun **2022**; 13:4931.
21. Ignacio RA B, Chew KW, Moser C, et al. Tixagevimab/cilgavimab IM and IV for COVID-19: a randomized controlled ACTIV-2 trial. JAMA Netw Open **2023**; 6: e2310039.
22. Berg MG, Zhen W, Lucic D, et al. Development of the real-time SARS-CoV-2 quantitative laboratory developed test and correlation with viral culture as a measure of infectivity. J Clin Virol **2021**; 143:104945.
23. Hughes JP. Mixed effects models with censored data with application to HIV RNA levels. Biometrics **1999**; 55:625–9.
24. Vaida F, Liu L. Fast implementation for normal mixed effects models with censored response. J Comput Graph Stat **2009**; 18:797–817.
25. Zou G. A modified Poisson regression approach to prospective studies with binary data. Am J Epidemiol **2004**; 159:702–6.
26. Taiwo BO, Chew KW, Moser C, et al. Phase 2 safety and antiviral activity of SAB-185, a novel polyclonal antibody therapy for non-hospitalized adults with COVID-19. J Infect Dis **2023**; 228:133–42.
27. Moser C, Li JZ, Eron JJ, et al. Predictors of SARS-CoV-2 RNA from nasopharyngeal swabs and concordance with other compartments in nonhospitalized adults with mild to moderate COVID-19. Open Forum Infect Dis **2022**; 9: ofac618.
28. Hughes MD. Analysis and design issues for studies using censored biomarker measurements with an example of viral load measurements in HIV clinical trials. Stat Med **2000**; 19:3171–91.
29. Uh H-W, Hartgers FC, Yazdanbakhsh M, Houwing-Duistermaat JJ. Evaluation of regression methods when immunological measurements are constrained by detection limits. BMC Immunol **2008**; 9:59.
30. Lubin JH, Colt JS, Camann D, et al. Epidemiologic evaluation of measurement data in the presence of detection limits. Environ Health Perspect **2004**; 112: 1691–6.
31. Watson JA, Kissler SM, Day NPJ, Grad YH, White NJ. Characterizing SARS-CoV-2 viral clearance kinetics to improve the design of antiviral pharmacometric studies. Antimicrob Agents Chemother **2022**; 66: e00192-22.
32. Perelson AS, Ke R. Mechanistic modeling of SARS-CoV-2 and other infectious diseases and the effects of therapeutics. Clin Pharmacol Ther **2021**; 109:829–40.
33. Wang S, Pan Y, Wang Q, Miao H, Brown AN, Rong L. Modeling the viral dynamics of SARS-CoV-2 infection. Math Biosci **2020**; 328:108438.
34. Ke R, Zitzmann C, Ho DD, Ribeiro RM, Perelson AS. In vivo kinetics of SARS-CoV-2 infection and its relationship with a person's Infectiousness. Proc Natl Acad Sci U S A **2021**; 118:e2111477118.
35. Hoffmann-La Roche. Study to evaluate the effects of AT-527 in non-hospitalized adult patients with mild or moderate COVID-19. 2022. https://clinicaltrials.gov/ct2/show/results/NCT04709835. Accessed 30 October 2022.
36. Golan Y, Campos JAS, Woolson R, et al. Favipiravir in patients with early mild-to-moderate coronavirus disease 2019 (COVID-19): a randomized controlled trial. Clin Infect Dis **2023**; 76:e10–7.