

1 **TITLE PAGE**

2 **Title**

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

5 **Running Title**

6 Analyzing SARS-CoV-2 RNA in COVID Trials

7 **AUTHORS**

8 Carlee B Moser, PhD, Center for Biostatistics in AIDS Research, Harvard T.H. Chan School of
9 Public Health, Boston, 02115, USA

10 Kara W Chew, MD, MS, Department of Medicine, David Geffen School of Medicine at University
11 of California, Los Angeles, Los Angeles, 90024, USA

12 Mark J Giganti, PhD, Center for Biostatistics in AIDS Research, Harvard T.H. Chan School of
13 Public Health, Boston, 02115, USA

14 Jonathan Z Li, MD, MMSc, Department of Medicine, Brigham and Women's Hospital, Harvard
15 Medical School, Cambridge, 02139, USA

16 Evgenia Aga, MS, Center for Biostatistics in AIDS Research, Harvard T.H. Chan School of
17 Public Health, Boston, 02115, USA

18 Justin Ritz, MS, Center for Biostatistics in AIDS Research, Harvard T.H. Chan School of Public
19 Health, Boston, 02115, USA

20 Alexander L Greninger, MD, PhD, Department of Laboratory Medicine & Pathology, University
21 of Washington, Seattle, 98195, USA

22 Arzhang Cyrus Javan, MD, MPH, DTM&H, National Institutes of Health, Rockville, 20852, USA

23 Rachel Bender Ignacio, MD, MPH, Department of Medicine, University of Washington, Seattle,
24 98195, USA

25 Eric S Daar, MD, Lundquist Institute at Harbor-UCLA Medical Center, Torrance, 90502, USA

26 David A Wohl, MD, Department of Medicine, University of North Carolina at Chapel Hill School
27 of Medicine, Chapel Hill, 27599, USA

28 Judith S Currier, MD, MSc, Department of Medicine, David Geffen School of Medicine at
29 University of California, Los Angeles, Los Angeles, 90024, USA

30 Joseph J Eron, MD, Department of Medicine, University of North Carolina at Chapel Hill School
31 of Medicine, Chapel Hill, 27599, USA

32 Davey M Smith, MD, MAS, Department of Medicine, University of California, San Diego, La
33 Jolla, 92093, USA

34 Michael D Hughes, PhD, Department of Biostatistics and Center for Biostatistics in AIDS
35 Research, Harvard T.H. Chan School of Public Health, Boston, 02115, USA

36 For the ACTIV-2/A5401 Study Team

37

38 **Word Count:** 145 (abstract) / 3465 (text)

39

40 **Corresponding Author**

41 Carlee B. Moser, PhD, Center for Biostatistics in AIDS Research, Harvard T.H. Chan School of
42 Public Health, 651 Huntington Ave, FXB 513, Boston, MA, 02115, USA, phone: (617) 432-2526,
43 fax: (617) 432-3163, email: cmoser@sdac.harvard.edu

44 **FOOTNOTE PAGE**

45 **Disclosures**

46 CBM participated on a Data Safety Monitoring Board for the BONE STAR study. KWC has
47 received research funding to the institution from Merck Sharp & Dohme and is a consultant for
48 Pardes Biosciences. JZL has consulted for Abbvie and received research support from Merck
49 ALG reports contract testing from Abbott, Cepheid, Novavax, Pfizer, Janssen and Hologic and
50 research support from Gilead. RBI received consulting fees from SeaGen and Resverlogix
51 unrelated to this work and research funding through her institution from Novartis and Ascentage
52 Pharma Group. ESD receives consulting fees from Gilead Sciences, Merck, and GSK/ViiV and
53 research support through the institution from Gilead Sciences and GSK/ViiV. DAW has received
54 funding to the institution to support research and honoraria for advisory boards and consulting
55 from Gilead Sciences. JSC has consulted for Merck and Company. JJE is an ad hoc consultant
56 to GSK/VIR, data monitoring committee (DMC) chair for Adagio Phase III studies. DMS has
57 consulted for the following companies: Evidera, Fluxergy, Kiadis, Linear Therapies, Matrix
58 BioMed, Arena Pharmaceuticals, VxBiosciences, Model Medicines, Bayer Pharmaceuticals,
59 Signant Health and Brio Clinical. MJG, EA, JR, MN, ACJ and MDH report no competing
60 interests.

61 **Funding Statement**

62 This work was supported by the National Institute of Allergy and Infectious Diseases of the
63 National Institutes of Health under Award Number UM1 AI068634 (MDH), UM1 AI068636 (JSC)
64 and UM1 AI106701 (GA). The content is solely the responsibility of the authors and does not
65 necessarily represent the official views of the National Institutes of Health.

66 **Meeting Information**

67 N/A

68 **Corresponding Author**

69 Carlee B. Moser, PhD, Center for Biostatistics in AIDS Research, Harvard T.H. Chan School of
70 Public Health, 651 Huntington Ave, FXB 513, Boston, MA, 02115, USA, phone: (617) 432-2526,
71 fax: (617) 432-3163, email: cmoser@sdac.harvard.edu

72 **ABSTRACT**

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

85

86 **Key Words:** SARS-CoV-2 RNA, COVID-19, linear regression for censored data, randomized
87 trial

88 **Trial Registration:** ClinicalTrials.gov Identifier: NCT04518410

89 **BACKGROUND**

90 Clinical trials designed to evaluate COVID-19 therapeutics should have clinically meaningful
91 endpoints. FDA guidance states that clinical outcomes, such as the proportion of participants
92 hospitalized or time to symptom recovery, are recommended as primary outcomes in phase III
93 outpatient COVID-19 trials [1]. However, it also states that viral shedding should be measured to
94 assess antiviral activity, primary virology outcomes are acceptable in phase II, and quantitative
95 and qualitative virological assessments are encouraged.

96
97 In typical COVID-19 randomized trials, samples such as nasopharyngeal swabs, anterior or
98 mid-turbinate nasal swabs, oropharyngeal swabs, saliva, or plasma, are collected longitudinally
99 for SARS-CoV-2 RNA testing before and after intervention. Repeat sampling from early
100 timepoints is common and in phase III typically includes one to four timepoints (Supplemental
101 Table 1).

102
103 To evaluate virologic efficacy, SARS-CoV-2 RNA, henceforth called viral RNA (vRNA), is
104 measured with quantitative reverse transcription polymerase chain reaction (RT-qPCR) assays.
105 Like other nucleic acid assays, SARS-CoV-2 RNA assays have limits between which vRNA is
106 accurately quantified, called the lower limit of quantification (LLoQ) and upper limit of
107 quantification (ULoQ). For results >ULoQ, samples can be rerun with dilution to obtain
108 quantifiable values. Assays may also indicate whether results <LLoQ are detectable or not.

109
110 Recent outpatient COVID-19 therapeutic trials considered various vRNA outcome measures
111 and statistical methods. Most commonly, vRNA changes from baseline were analyzed using

112 analysis of covariance (ANCOVA) at each timepoint or mixed models for repeated measures
113 (MMRM). With these methods, single-imputation was used to assign values for vRNA results
114 <LLoQ (Supplemental Table 1) [2–18]. However, such imputation can introduce bias in
115 estimating the magnitudes of treatment effects, as uncertainty for values <LLoQ isn't captured
116 [19].

117
118 Using an illustrative example from the ACTIV-2 COVID-19 outpatient treatment trial, we
119 describe bias that may arise when estimating treatment effects using single-imputation with
120 ANCOVA and MMRM. Drawing on the HIV literature [19], we describe and discuss alternative
121 approaches for analyzing vRNA changes, that may be more appropriate by considering vRNA
122 values <LLoQ as censored measurements. Finally, we provide recommendations for the
123 analysis and presentation of results concerning vRNA changes in future trials.

124

125 **METHODS**

126 ACTIV-2 (NCT04518410) is an adaptive platform trial designed to evaluate potential outpatient
127 therapeutics for COVID-19[20]. Our illustrative example includes 114 participants randomized to
128 receive tixagevimab/cilgavimab intravenously or placebo; the primary results previously reported
129 [21]. Nasopharyngeal swabs were collected before treatment at Day 0 (baseline) and Days 3, 7
130 and 14 for SARS-CoV-2 RNA quantitative testing using a RT-qPCR assay with LLoQ of $2 \log_{10}$
131 copies/ml [22]. All results >ULoQ were rerun with dilution to obtain quantifiable results. ACTIV-2
132 was approved by a central institutional review board (IRB), Advarra (Pro00045266), with
133 additional local IRB review and approval as required by participating sites. All participants
134 provided written informed consent.

135

136 As this manuscript aims to illustrate and discuss different approaches to analyze vRNA
137 changes, we provide an overview in Table 1, but integrate descriptions of each method in the
138 Results. For methods that use imputed values for results <LLoQ, two commonly-used single-
139 imputation strategies (Supplemental Table 1) were assessed:

140 (1) "*LLoQ-imputation*": impute values <LLoQ as the LLoQ,

141 (2) "*½LLoQ-imputation*": impute values <LLoQ as ½ the LLoQ.

142 See Supplemental Methods for additional details on model specifications and sample SAS
143 software code.

144

145 RESULTS

146 Descriptive summaries of vRNA across timepoints for the 114 participants are shown in Table
147 2A and Figure 1A and 1B. At baseline, 15 participants (13%) had missing vRNA (Supplemental
148 Figure 1). There was a chance imbalance in vRNA between the randomized arms, with median
149 vRNA in the active arm 1.0 log₁₀ copies/ml higher than the placebo arm, and a higher proportion
150 of participants with vRNA ≥LLoQ (72% versus 62%).

151

152 Following the recommendation of Marschner et al. [19], we separately considered data for
153 participants with vRNA <LLoQ from those ≥LLoQ at baseline. For those with vRNA <LLoQ at
154 baseline (N=33), vRNA remained <LLoQ at all follow-up timepoints in both arms, suggesting
155 peak vRNA may have been achieved before enrollment. For the remaining analyses, we focus
156 on the 66 participants with vRNA ≥LLoQ at baseline. The proportion with vRNA <LLoQ
157 increased over time: 27% and 28% at Day 3, 62% and 54% at Day 7, and 93% and 89% at Day
158 14 for the active and placebo arms, respectively (Table 2B and Figure 1C and 1D).

159

160 Analyzing vRNA at a Single Timepoint

161 1. *Using imputed values leads to biased estimates*

162 Fifty-five (83%) of the 66 participants had vRNA results at Day 3 (Supplemental Figure 1). For
163 these 55 participants, at baseline there was a modest difference (0.33 log₁₀ copies/ml) in mean
164 vRNA: 5.61 and 5.28 log₁₀ copies/ml for the active and placebo arms, respectively.

165

166 Using *LLoQ-imputation*, the mean vRNA at Day 3 was 3.43 and 3.97 log₁₀ copies/ml for the
167 active and placebo arms, respectively, with estimated mean changes from baseline of -2.18 and
168 -1.30 log₁₀ copies/ml. Within each arm, the estimated mean changes are conservative and
169 biased because for participants with vRNA <LLoQ at Day 3, the true changes are at least as
170 large in magnitude as the imputed changes. Using *½LLoQ-imputation* gives mean changes that
171 are larger (more negative) compared to *LLoQ-imputation*: -2.45 and -1.58 log₁₀ copies/ml for the
172 active and placebo arms, respectively. This imputation still results in biased estimates, but with
173 an unknown direction (estimated changes may be larger or smaller than the truth). For both
174 approaches, the larger mean change in the active arm could reflect higher average baseline
175 values, and thus larger changes are observable. Since the estimated mean changes within each
176 arm are biased, the estimated difference between arms will be biased, and further bias may be
177 introduced with the baseline imbalances.

178

179 The estimated difference in mean change for the active versus placebo arms at Day 3 was -0.87
180 log₁₀ copies/ml using *LLoQ-imputation* and -0.86 log₁₀ copies/ml using *½LLoQ-imputation* (Table
181 3A). Although these estimates are similar, this may not be the case in other datasets when

182 using the two approaches. By Day 14, when ~90% of participants had vRNA <LLOQ (and hence
183 had imputed changes), the estimated difference in mean change between arms was
184 approximately equal to the baseline mean difference for both imputation approaches. If all
185 participants had vRNA <LLOQ at Day 14, the difference in mean change would equal the
186 difference in mean vRNA at baseline, despite the choice of imputed value and underlying true
187 difference. With larger proportions <LLOQ, differences between arms can reflect chance
188 imbalances at baseline rather than true differences.

189

190 *2. Adjusting for baseline can help address baseline imbalances*

191 Although adjusting for baseline doesn't remove the bias in estimating differences between arms
192 using singly-imputed values, it may help reduce the impact of baseline imbalances in mean
193 vRNA when assessing treatment effects.

194

195 The estimated differences in mean changes between arms using standard linear regression are
196 shown in Table 3A-B. In adjusted analyses, differences between arms have some attenuation at
197 each timepoint compared with unadjusted analyses, reflecting the adjustment for higher
198 baseline vRNA levels in the active arm.

199

200 *3. Analysis methods considering vRNA <LLOQ as censored*

201 Statisticians refer to vRNA values <LLOQ as being left-censored because the if the true vRNA
202 could be measure it would be a value between zero and LLOQ (i.e., a value to the left of LLOQ).
203 This contrasts with right-censoring like in survival analysis where, for example, participants alive
204 at the end of follow-up have time of death greater than (to the right of) the time at the end of

205 follow-up. Statistical methods used for survival analysis can be used to analyze vRNA data, with
206 the small adaptation that values are left-censored rather than right-censored. Change in vRNA
207 is defined as the difference in vRNA at the follow-up time minus the baseline. However, for
208 follow-up vRNA values that are <LLOQ or left-censored, the change in vRNA is calculated as the
209 LLOQ minus baseline vRNA, and is also left-censored.

210

211 Linear regression using software designed to handle censored data (known as tobit regression)
212 is a possible method. Using this approach, adjusting for baseline vRNA, the estimated
213 difference between arms in mean change from baseline to Day 3 was $-0.97 \log_{10}$ copies/ml
214 (95% confidence interval [CI]: $-1.81, -0.13$) favoring the active arm (Table 3C), and is somewhat
215 larger than the differences in mean change by either imputation approach (Table 3B). At Day 7,
216 the difference in mean change from baseline was $-1.36 \log_{10}$ copies/ml, also favoring the active
217 arm (95% CI: $-2.31, -0.41$), which is much larger than differences observed by either imputation
218 approach, illustrating the potential bias using those methods when the proportion with vRNA
219 <LLOQ increases. We didn't pursue an analysis of mean changes to Day 14 using tobit
220 regression because of the high level of censoring (~90%) and hence the inability to check model
221 assumptions.

222

223 As with standard linear regression, there is an assumption that the errors in the model are
224 normally distributed. These errors are estimated by the residuals calculated as the observed
225 vRNA value minus the predicted model value. The distributional assumption can be evaluated
226 with quantile-quantile (Q-Q) plots, comparing the quantiles of the observed distribution of the
227 residuals (calculated using Kaplan-Meier methods to account for censored residuals) against
228 the corresponding quantiles of a standard normal distribution. If the assumption was satisfied,

229 the plots would show linear associations. Figure 2 shows Q-Q plots for the distribution of
230 standardized residuals from models for change from baseline, adjusting for baseline. For the
231 models of change from baseline to Days 3 and 7, the Q-Q plots appear reasonably linear,
232 supporting normality assumptions. We note, however, the more restricted range of the Q-Q plot
233 for changes to Day 7, as shown by the lack of standardized residuals below -1. This reflects the
234 higher proportion of censored values at Day 7; thus, the normality assumption cannot be verified
235 for the tail of the distribution, corresponding to large negative changes from baseline.

236

237 *4. Quantile regression as an alternative distribution-free method*

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

243

244 At Day 3, the adjusted difference between arms in median change from baseline was $-1.17 \log_{10}$
245 copies/ml (95% CI: -2.42, 0.07) favoring the active arm. This is reasonably similar to the
246 adjusted difference in mean change of $-0.97 \log_{10}$ copies/ml obtained from tobit regression,
247 though estimated without making the assumption of normally distributed errors. There is a
248 somewhat narrower CI for the difference in means, versus difference in medians, reflecting the
249 gain in precision from assuming a normal distribution for the errors. At Day 7, the adjusted
250 difference in median change was $-0.96 \log_{10}$ copies/ml, also favoring the active arm. However, it
251 wasn't possible to obtain a CI from the numerical methods used to fit the model, due to the high

252 proportion of participants with vRNA <LLOQ at Day 7. At Day 14, the higher proportion with
253 vRNA <LLOQ meant the difference in median change between arms couldn't be estimated.

254

255 Analyzing Repeated vRNA Over Time

256 *5. Imputed values can affect estimates from MMRM due to correlation structure*

257 Another strategy in several recent COVID-19 trials has been to use an MMRM with single-
258 imputation for vRNA values <LLOQ [2–14]. These models estimate the difference in mean vRNA
259 change in each arm at each timepoint, in a similar manner as linear regression models fit
260 separately by timepoint. However, MMRMs incorporate a stronger assumption about the
261 distribution of errors across timepoints, specifically that they follow a multivariate normal
262 distribution with a specified correlation structure. Using this assumption, a global test evaluating
263 the null hypothesis of no difference between arms in vRNA change at any timepoint can be
264 undertaken. The stronger assumption may provide improved precision in estimating the
265 differences in mean change at each timepoint by borrowing information between timepoints.
266 However, this assumption may not be appropriate when using singly-imputed values for
267 measurements <LLOQ as the correlation structure is affected by imputation. As an example,
268 participants with vRNA <LLOQ at Days 7 and 14 will have identical imputed changes at both
269 timepoints leading to higher correlations of errors in the model, than if the actual values <LLOQ
270 were observed. To illustrate the impact of this, Table 3E shows results from MMRMs for
271 changes from baseline to Days 3, 7, and 14. Compared with the estimates from models fitted
272 separately at each timepoint (Table 3B), the borrowing of information through the correlation
273 structure leads to smaller estimated differences in mean change between arms, particularly at
274 Day 3 and to a lesser extent at Day 7 for both imputation approaches. This attenuation is driven
275 by including Day 14, where ~90% of participants had vRNA <LLOQ; removal of this timepoint

276 from the MMRM reduces the magnitude of the attenuation (Table 3F). The estimates remain
277 biased, however, for the same reasons as those obtained from separate regression models at
278 each timepoint.

279

280 Extensions to MMRM that account for censored data exist (also known as linear mixed effects
281 models for censored responses[LMEC]), but still require the multivariate normality
282 assumption[23,24]. A caveat with these models is that they can be difficult to implement in
283 standard statistical software, especially as the number of timepoints increases. Estimated
284 differences between arms in mean change from baseline to Days 3 and 7 from LMEC are
285 shown in Table 3G. The estimates are similar to those from the tobit regression models fitted
286 separately at Days 3 and 7 (Table 3C). The stronger multivariate normal assumption leads to
287 small gains in precision at Day 7 as seen by the narrower CI, though the gain at Day 3, where
288 there's less censoring, is negligible. As with the separate regression models, we didn't pursue
289 LMEC over the three days, as the high level of censoring at Day 14 meant that a normality
290 assumption couldn't be reasonably verified.

291

292 Analyzing Proportion of Participants with vRNA <LLOQ Over Time

293 6. *Strategies that don't rely on quantitative values may be preferred with large % <LLOQ*

294 When there is a high proportion of participants with vRNA <LLOQ at one or more timepoints, it
295 may be more appropriate to focus on how this proportion changes with time. This could be
296 analyzed over time using log-binomial models fit using generalized estimating equations (GEE).
297 However, due to problems with numerical algorithms, in ACTIV-2 we used Poisson regression
298 models modified for binary outcomes [25] fit using GEEs with independence working correlation
299 structure and robust standard errors, adjusting for baseline vRNA. When implementing this

300 model across the three days, the proportion with vRNA <LLoQ didn't differ between arms
301 (Supplemental Table 2). When excluding the Day 14 measurements, where ~90% of
302 participants had vRNA <LLoQ, the results for Days 3 and 7 were almost identical, confirming
303 this method isn't sensitive to including timepoints with high proportions <LLoQ. This strategy
304 can lead to loss in statistical power compared to analyses of quantitative vRNA, so is best
305 reserved for when high proportions of participants are expected to have vRNA <LLoQ at one or
306 more timepoints. However, there is also no need to restrict the analysis population to
307 participants with vRNA \geq LLoQ, potentially providing more comprehensive analyses of qualitative
308 vRNA in the overall study population.

309

310 **DISCUSSION**

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

321

322 In our illustrative example, the primary focus was on the population with quantifiable vRNA at
323 baseline, which has been a focus in recent COVID-19 studies. This was reasonable in our

324 analysis as none who were <LLoQ at baseline had quantifiable vRNA at later timepoints.
325 Including these individuals in analyses using imputed values would have led to imputed
326 changes of zero and likely attenuation of the estimated mean changes. Regression analyses for
327 censored data are more complex if such individuals are included, requiring strong, unverifiable
328 assumptions about the distribution of vRNA changes over time among those with baseline
329 vRNA <LLoQ. Looking more broadly across the study population in phase II placebo-controlled
330 evaluations in ACTIV-2 (N=1565 enrolled with a median of 6 days from symptom onset), we
331 observed that only 14% (of 287) of those with vRNA <LLoQ at baseline later had quantifiable
332 vRNA. As new studies are developed, potentially with enrollment closer to onset of symptoms,
333 the decision to exclude those <LLoQ at baseline should be carefully scrutinized, as doing so
334 could remove individuals on an upward viral load trajectory and we lack understanding of these
335 trajectories in the setting of vaccination, reinfection, and emergent variants. At a minimum,
336 documenting viral shedding changes among participants with baseline vRNA <LLoQ is
337 important, and analyses stratified by level (<LLoQ and ≥LLoQ at baseline) might be pursued.

338
339 The methods considered in this paper aren't exhaustive of imputation or modeling strategies,
340 but were chosen to align with methods from recent publications of COVID-19 trials. We focus on
341 single-imputation, and don't evaluate the performance of multiple-imputation strategies, which
342 are more complicated and rely on distributional assumptions to support the imputation, but may
343 reduce potential biases with imputation highlighted in this paper [28,29]. We also haven't
344 evaluated the statistical performance of these methods through formal simulation studies, which
345 may add further insights to benefits or downsides of the analytic strategies, particularly when
346 high proportions of participants have vRNA <LLoQ during follow-up, where verification of model
347 assumptions becomes more difficult. We also haven't considered potential biases due to
348 missing data, for example, missingness arising due to hospitalization, if hospitalized participants

349 have higher vRNA levels. In designing studies, the impact on power and precision in estimating
350 treatment effects needs consideration [30]. Finally, analysis of vRNA changes among
351 participants with baseline levels above a threshold (e.g., the LLoQ) leads to estimated mean
352 changes within each arm that are affected by regression to the mean, though estimated
353 differences in mean changes between randomized arms are not. Despite these limitations, our
354 paper highlights key issues and considerations when analyzing SARS-CoV-2 RNA data from
355 outpatient treatment trials. These methods aren't only applicable in the COVID-19 setting, but
356 should be considered when analyzing any biomarker that is measured with an assay with an
357 LLoQ.

358

359 **Recommendations**

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

- 363 1. Provide sufficient details of the RT-qPCR assay, including the LLoQ.
- 364 2. Explain who is included in the analysis, such as via a CONSORT-type diagram (see
365 Supplemental Figure 1), including an accounting of missing data and the reasons for
366 missing (e.g., death, hospitalization, loss to follow-up, sample not obtained, sample
367 processing/shipping issue).
- 368 3. If restricting the analysis population to those with quantifiable baseline vRNA, describe
369 outcomes among those with vRNA <LLoQ.
- 370 4. Although we don't recommend the use of single-imputation, if used, the choice of
371 imputed values should be provided, and implications for interpretation of results
372 discussed.

373 5. Include descriptive summaries of vRNA by treatment arm and timepoint. We suggest
374 including two figures (see Figure 1): distributions of quantitative levels (e.g., box and
375 whisker plots) and distribution of vRNA categories (e.g., <LLOQ versus ≥LLOQ).

376

377 Analytic strategies to estimate differences between arms we recommend:

378 1. Methods that address censoring without imputation, such as tobit or median regression,
379 or LMEC [23,24] be prioritized. But with increased censoring:

380 a. Normality assumptions underlying regression analysis for censored data cannot
381 be evaluated over the full range of the distribution, and dropping timepoints with
382 high levels of censoring from analysis may be appropriate.

383 b. Differences in medians (and their CIs) between arms might not be estimable from
384 quantile regression.

385 2. Alternatively, consider non-parametric tests to analyze quantitative vRNA, such as the
386 censored version of the Wilcoxon test (Gehan-Wilcoxon) which is implementable in
387 standard software as a stratified test to account for baseline vRNA.

388 3. Comparing the proportion of participants with vRNA <LLOQ between arms over time may
389 be preferred if there are high amounts of censoring.

390 4. With early, frequent measurements (e.g., daily), more complex extensions of LMEC that
391 evaluate viral dynamics (e.g., estimating initial increases and subsequent vRNA decay)
392 [20,31–34], or time-to-viral clearance via methods for time-to-event data [4–7,10,35,36]
393 might be used.

394 **ACKNOWLEDGMENTS**

395 We thank the study participants, site staff, site investigators, and the entire ACTIV-2/A5401
396 study team; the AIDS Clinical Trials Group, including Lara Hosey, Jhoanna Roa, and Nilam
397 Patel; the ACTG Laboratory Center, including Grace Aldrovandi and William Murtaugh; Frontier
398 Science, including Marlene Cooper, Howard Gutzman, and Kevin Knowles; the Harvard Center
399 for Biostatistics in AIDS Research (CBAR) and ACTG Statistical and Data Analysis Center
400 (SDAC) including Ronald Bosch and Linda Harrison; the National Institute of Allergy and
401 Infectious Diseases (NIAID) / Division of AIDS (DAIDS); Bill Erhardt; the Foundation for the
402 National Institutes of Health and the Accelerating COVID-19 Therapeutic Interventions and
403 Vaccines (ACTIV) partnership, including Stacey Adams; and the PPD clinical research business
404 of Thermo Fisher Scientific. We also thank AstraZeneca for partnership in the ACTIV-2 study
405 and for donation of study drug for the cohort used as the illustrative example in this analysis.

406

407 **REFERENCES**

- 408 1. COVID-19: Developing Drugs and Biological Products for Treatment or Prevention. **2021**; .
- 409 2. Weinreich DM, Sivapalasingam S, Norton T, et al. REGEN-COV Antibody Combination
410 and Outcomes in Outpatients with Covid-19. *N Engl J Med*. Massachusetts Medical
411 Society; **2021**; 385(23):e81.
- 412 3. Weinreich DM, Sivapalasingam S, Norton T, et al. REGN-COV2, a Neutralizing Antibody
413 Cocktail, in Outpatients with Covid-19. *New England Journal of Medicine*. Massachusetts
414 Medical Society; **2021**; 384(3):238–251.
- 415 4. Dougan M, Azizad M, Mocherla B, et al. A randomized, placebo-controlled clinical trial of
416 bamlanivimab and etesevimab together in high-risk ambulatory patients with COVID-19
417 and validation of the prognostic value of persistently high viral load. *Clin Infect Dis*. **2021**;
418 :ciab912.
- 419 5. Dougan M, Nirula A, Azizad M, et al. Bamlanivimab plus Etesevimab in Mild or Moderate
420 Covid-19. *New England Journal of Medicine*. Massachusetts Medical Society; **2021**;
421 385(15):1382–1392.
- 422 6. Dougan M, Azizad M, Chen P, et al. Bebtelovimab, alone or together with bamlanivimab
423 and etesevimab, as a broadly neutralizing monoclonal antibody treatment for mild to
424 moderate, ambulatory COVID-19 [Internet]. medRxiv; 2022 [cited 2023 Feb 9]. p.
425 2022.03.10.22272100. Available from:
426 <https://www.medrxiv.org/content/10.1101/2022.03.10.22272100v1>
- 427 7. Fischer WA, Eron JJ, Holman W, et al. A phase 2a clinical trial of molnupiravir in patients
428 with COVID-19 shows accelerated SARS-CoV-2 RNA clearance and elimination of
429 infectious virus. *Science Translational Medicine*. American Association for the
430 Advancement of Science; **2021**; 14(628):eabl7430.
- 431 8. Montgomery H, Hobbs FDR, Padilla F, et al. Efficacy and safety of intramuscular
432 administration of tixagevimab–cilgavimab for early outpatient treatment of COVID-19
433 (TACKLE): a phase 3, randomised, double-blind, placebo-controlled trial. *The Lancet*
434 *Respiratory Medicine*. Elsevier; **2022**; 10(10):985–996.
- 435 9. Gupta A, Gonzalez-Rojas Y, Juarez E, et al. Effect of Sotrovimab on Hospitalization or
436 Death Among High-risk Patients With Mild to Moderate COVID-19: A Randomized Clinical
437 Trial. *JAMA*. **2022**; 327(13):1236–1246.
- 438 10. Jagannathan P, Andrews JR, Bonilla H, et al. Peginterferon Lambda-1a for treatment of
439 outpatients with uncomplicated COVID-19: a randomized placebo-controlled trial. *Nat*
440 *Commun*. Nature Publishing Group; **2021**; 12(1):1967.
- 441 11. Alemany A, Millat-Martinez P, Corbacho-Monné M, et al. High-titre methylene blue-treated
442 convalescent plasma as an early treatment for outpatients with COVID-19: a randomised,
443 placebo-controlled trial. *The Lancet Respiratory Medicine*. Elsevier; **2022**; 10(3):278–288.

- 444 12. Reis G, Moreira Silva EAS, Medeiros Silva DC, et al. Early Treatment with Pegylated
445 Interferon Lambda for Covid-19. *New England Journal of Medicine*. Massachusetts
446 Medical Society; **2023**; 388(6):518–528.
- 447 13. Ramakrishnan S, Nicolau DV, Langford B, et al. Inhaled budesonide in the treatment of
448 early COVID-19 (STOIC): a phase 2, open-label, randomised controlled trial. *The Lancet*
449 *Respiratory Medicine*. Elsevier; **2021**; 9(7):763–772.
- 450 14. Mitjà O, Corbacho-Monné M, Ubals M, et al. Hydroxychloroquine for Early Treatment of
451 Adults with Mild Covid-19: A Randomized-Controlled Trial. *Clin Infect Dis*. **2020**; :ciaa1009.
- 452 15. Hammond J, Leister-Tebbe H, Gardner A, et al. Oral Nirmatrelvir for High-Risk,
453 Nonhospitalized Adults with Covid-19. *New England Journal of Medicine*. Massachusetts
454 Medical Society; **2022**; 386(15):1397–1408.
- 455 16. Jayk Bernal A, Gomes da Silva MM, Musungaie DB, et al. Molnupiravir for Oral Treatment
456 of Covid-19 in Nonhospitalized Patients. *New England Journal of Medicine*. Massachusetts
457 Medical Society; **2022**; 386(6):509–520.
- 458 17. Gottlieb RL, Vaca CE, Paredes R, et al. Early Remdesivir to Prevent Progression to
459 Severe Covid-19 in Outpatients. *New England Journal of Medicine*. Massachusetts
460 Medical Society; **2022**; 386(4):305–315.
- 461 18. Mukae H, Yotsuyanagi H, Ohmagari N, et al. Efficacy and Safety of Ensitrelvir in Patients
462 With Mild-to-Moderate Coronavirus Disease 2019 (COVID-19): The Phase 2b Part of a
463 Randomized, Placebo-Controlled, Phase 2/3 Study. *Clinical Infectious Diseases*. **2022**;
464 :ciac933.
- 465 19. Marschner IC, Betensky RA, DeGruttola V, Hammer SM, Kuritzkes DR. Clinical trials using
466 HIV-1 RNA-based primary endpoints: statistical analysis and potential biases. *J Acquir*
467 *Immune Defic Syndr Hum Retrovirol*. **1999**; 20(3):220–227.
- 468 20. Chew KW, Moser C, Daar ES, et al. Antiviral and clinical activity of bamlanivimab in a
469 randomized trial of non-hospitalized adults with COVID-19. *Nat Commun*. Nature
470 Publishing Group; **2022**; 13(1):4931.
- 471 21. Bender Ignacio RA, Chew KW, Moser C, et al. Tixagevimab/Cilgavimab IM and IV for
472 COVID-19: A Randomized Controlled ACTIV-2 Trial. Seattle, WA: Conference on
473 Retroviruses and Opportunistic Infections; 2023.
- 474 22. Berg MG, Zhen W, Lucic D, et al. Development of the RealTime SARS-CoV-2 quantitative
475 Laboratory Developed Test and correlation with viral culture as a measure of infectivity. *J*
476 *Clin Virol*. **2021**; 143:104945.
- 477 23. Hughes JP. Mixed effects models with censored data with application to HIV RNA levels.
478 *Biometrics*. **1999**; 55(2):625–629.
- 479 24. Vaida F, Liu L. Fast Implementation for Normal Mixed Effects Models With Censored
480 Response. *J Comput Graph Stat*. **2009**; 18(4):797–817.

- 481 25. Zou G. A modified poisson regression approach to prospective studies with binary data.
482 Am J Epidemiol. **2004**; 159(7):702–706.
- 483 26. Taiwo BO, Chew KW, Moser C, et al. Phase 2 safety and antiviral activity of SAB-185, a
484 novel polyclonal antibody therapy for non-hospitalized adults with COVID-19. J Infect Dis.
485 **2023**; :jjad013.
- 486 27. Moser C, Li JZ, Eron JJ, et al. Predictors of SARS-CoV-2 RNA From Nasopharyngeal
487 Swabs and Concordance With Other Compartments in Nonhospitalized Adults With Mild to
488 Moderate COVID-19. Open Forum Infect Dis. **2022**; 9(11):ofac618.
- 489 28. Uh H-W, Hartgers FC, Yazdanbakhsh M, Houwing-Duistermaat JJ. Evaluation of
490 regression methods when immunological measurements are constrained by detection
491 limits. BMC Immunol. **2008**; 9:59.
- 492 29. Lubin JH, Colt JS, Camann D, et al. Epidemiologic Evaluation of Measurement Data in the
493 Presence of Detection Limits. Environ Health Perspect. **2004**; 112(17):1691–1696.
- 494 30. Hughes MD. Analysis and design issues for studies using censored biomarker
495 measurements with an example of viral load measurements in HIV clinical trials. Stat Med.
496 **2000**; 19(23):3171–3191.
- 497 31. Watson JA, Kissler SM, Day NPJ, Grad YH, White NJ. Characterizing SARS-CoV-2 Viral
498 Clearance Kinetics to Improve the Design of Antiviral Pharmacometric Studies. Antimicrob
499 Agents Chemother. 66(7):e00192-22.
- 500 32. Perelson AS, Ke R. Mechanistic Modeling of SARS-CoV-2 and Other Infectious Diseases
501 and the Effects of Therapeutics. Clin Pharmacol Ther. **2021**; 109(4):829–840.
- 502 33. Wang S, Pan Y, Wang Q, Miao H, Brown AN, Rong L. Modeling the viral dynamics of
503 SARS-CoV-2 infection. Math Biosci. **2020**; 328:108438.
- 504 34. Ke R, Zitzmann C, Ho DD, Ribeiro RM, Perelson AS. In vivo kinetics of SARS-CoV-2
505 infection and its relationship with a person’s infectiousness. Proc Natl Acad Sci U S A.
506 **2021**; 118(49):e2111477118.
- 507 35. Hoffmann-La Roche. A Phase II Randomized, Double-Blind, Placebo-Controlled Study to
508 Evaluate the Antiviral Activity, Safety, Pharmacokinetics, and Efficacy of RO7496998 (AT-
509 527) in Non-Hospitalized Adult Patients With Mild or Moderate COVID-19 [Internet].
510 clinicaltrials.gov; 2022 Oct. Report No.: results/NCT04709835. Available from:
511 <https://clinicaltrials.gov/ct2/show/results/NCT04709835>
- 512 36. Golan Y, Campos JAS, Woolson R, et al. Favipiravir in Patients With Early Mild-to-
513 moderate Coronavirus Disease 2019 (COVID-19): A Randomized Controlled Trial. Clinical
514 Infectious Diseases. **2023**; 76(3):e10–e17.
- 515

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 Timepoints	Handling Values <LLoQ	Pros	Caveats/Issues
Analysis of covariance (ANCOVA)/linear regression	1	Single Imputation	<p>Easy to implement in standard software.</p> <p>With small proportion <LLoQ, impact of imputation is likely modest.</p>	<p>Using imputation results in biased estimates of differences between randomized arms in mean change.</p> <p>Normality assumption in model may be violated</p> <p>Need to restrict to those \geqLLoQ at baseline to calculate changes.</p>
Linear regression for censored data (tobit regression)	1	Not Required	<p>Easy to implement in standard software.</p> <p>Analyses considering censored measurements avoids bias that may be created by using imputed</p>	<p>Normality assumption in model cannot be confirmed when large proportion of data are censored.</p> <p>Need to restrict to those \geqLLoQ at baseline to calculate changes.</p>

			values.	
Median regression for censored data	1	Not Required	<p>Easy to implement in standard software.</p> <p>Distribution free model removes assumptions about distribution of the errors.</p>	<p>Model cannot be fitted when large proportion of data are censored.</p> <p>Need to restrict to those \geq LLoQ at baseline to calculate changes.</p>
Mixed models for repeated measures (MMRM)	> 1	Single Imputation	<p>Easy to implement in standard software.</p> <p>Global test of, no difference between randomized arms across timepoints, can be easily generated.</p>	<p>Using imputation results in biased estimates of the difference between randomized arms in mean change, with the bias at one time dependent on the proportion $<$ LLoQ at other times (as information is shared among times through an assumed correlation structure).</p> <p>Multivariate normality assumption may be violated.</p> <p>Need to restrict to those \geq LLoQ at baseline to calculate changes.</p>
MMRM for Censored Data (Linear mixed effects models for censored data,	> 1	Not Required	Analyses considering censored measurements avoids bias that may be	<p>Increase complexity in implementing model in standard software as the number of timepoints increases.</p> <p>Multivariate normality assumption difficult to verify,</p>

LMEC)			<p>created by using imputed values.</p> <p>Global test of, no difference between randomized arms across timepoints, can be easily generated.</p> <p>Possible improved precision by sharing information over timepoints through an assumed model.</p>	<p>particularly when large proportion of data are censored at one or more times.</p> <p>Need to restrict to those \geq LLoQ at baseline to calculate changes.</p>
Binary Regression	≥ 1	Not Required	<p>Easy to implement in standard software.</p> <p>Includes all participants, regardless of baseline value.</p> <p>Estimation of treatment effects not influenced by the</p>	<p>Loss of statistical power when dichotomizing outcome from continuous variable to a binary variable.</p>

			proportion <LLoQ.	
--	--	--	-------------------	--

LLoQ = Lower Limit of Quantification; ANCOVA = Analysis of Covariance; MMRM = Mixed Model Repeated Measures; LMEC =

Linear Mixed Effects Models with Censored Response

Table 2: Distribution of SARS-CoV-2 RNA by Study Visit in each Treatment Arm in overall cohort (A) and among those with vRNA \geq LLoQ at Baseline/Day 0 (B)

A: All participants in cohort (N=114)			
Visit		Active (N=58)	Placebo (N=56)
Baseline	Median (quartiles)	4.0 (<LLoQ, 6.6)	3.0 (<LLoQ, 5.9)
	<LLoQ, n (%)	14 (28)	19 (38)
	Missing, n	9	6
Day 3	Median (quartiles)	<LLoQ (<LLoQ, 3.9)	<LLoQ (<LLoQ, 3.9)
	<LLoQ, n (%)	26 (52)	24 (52)
	Missing, n	8	10
Day 7	Median (quartiles)	<LLoQ (<LLoQ, 2.2)	<LLoQ (<LLoQ, 2.2)
	<LLoQ, n (%)	37 (74)	35 (71)
	Missing, n	8	7
Day 14	Median (quartiles)	<LLoQ (<LLoQ, <LLoQ)	<LLoQ (<LLoQ, <LLoQ)
	<LLoQ, n (%)	49 (98)	45 (97)
	Missing, n	11	7
B: All participants with vRNA \geq LLoQ at baseline (N=66)			
Visit		Active (N=35)	Placebo (N=31)
Baseline	Median (quartiles)	5.5 (3.7, 8.0)	5.0 (3.1, 6.7)
	<LLoQ, n (%)	0 (0)	0 (0)
	Missing, n	0	0
Day 3	Median (quartiles)	3.0 (<LLoQ, 4.5)	3.4 (<LLoQ, 5.9)
	<LLoQ, n (%)	8 (27)	7 (28)
	Missing, n	5	6
Day 7	Median (quartiles)	<LLoQ (<LLoQ, 2.5)	<LLoQ (<LLoQ, 3.3)
	<LLoQ, n (%)	18 (62)	14 (54)

	Missing, n	6	5
Day 14	Median (quartiles)	<LLoQ (<LLoQ, <LLoQ)	<LLoQ (<LLoQ, <LLoQ)
	<LLoQ, n (%)	27 (93)	24 (89)
	Missing, n	6	4

LLoQ = Lower Limit of Quantification

Table 3: Differences between Treatment Arms in SARS-CoV-2 RNA (log₁₀ copies/ml) change from baseline– Mean/Median^a, 95% CI and p-value among those with quantifiable baseline vRNA

Imputation	Day 3	Day 7	Day 14
A: Linear regression model with imputation, separate model by day – unadjusted			
LLoQ imputation	-0.87 (-1.70, -0.06) p=0.037	-0.82 (-1.79, 0.15) p=0.09	-0.25 (-1.30, 0.81) p=0.64
½LLoQ imputation	-0.86 (-1.69, -0.04) 0.041	-0.90 (-1.82, 0.01) p=0.053	-0.29 (-1.32, 0.74) p=0.58
B: Linear regression model with imputation, separate model by day – adjusted for baseline			
LLoQ imputation	-0.74 (-1.41, -0.06) p=0.034	-0.56 (-1.01, -0.11) p=0.015	-0.06 (-0.18, 0.07) p=0.38
½LLoQ imputation	-0.77 (-1.53, 0.002) 0.050	-0.69 (-1.29, -0.09) p=0.024	-0.11 (-0.37, 0.16) p=0.42
C: Linear regression model for censored data (tobit regression), separate model by day – adjusting for baseline			
N/A	-0.97 (-1.81, -0.13) p=0.023	-1.36 (-2.31, -0.41) p=0.005	Not Obtained ^b
D: Median regression model for censored data, separate model by day – adjusting for baseline			
N/A	-1.17 (-2.42, 0.07) p=0.07	-0.96 (NE, NE) NE	NE
E: MMRM across all three days (Day 3, 7 and 14) with imputation – adjusting for baseline			
LLoQ imputation	-0.39 (-1.23, 0.45) p=0.36	-0.49 (-0.95, -0.04) p=0.032	-0.07 (-0.20, 0.06) p=0.27
½LLoQ imputation	-0.52 (-1.44, 0.40)	-0.60 (-1.21, 0.01)	-0.13 (-0.40, 0.14)

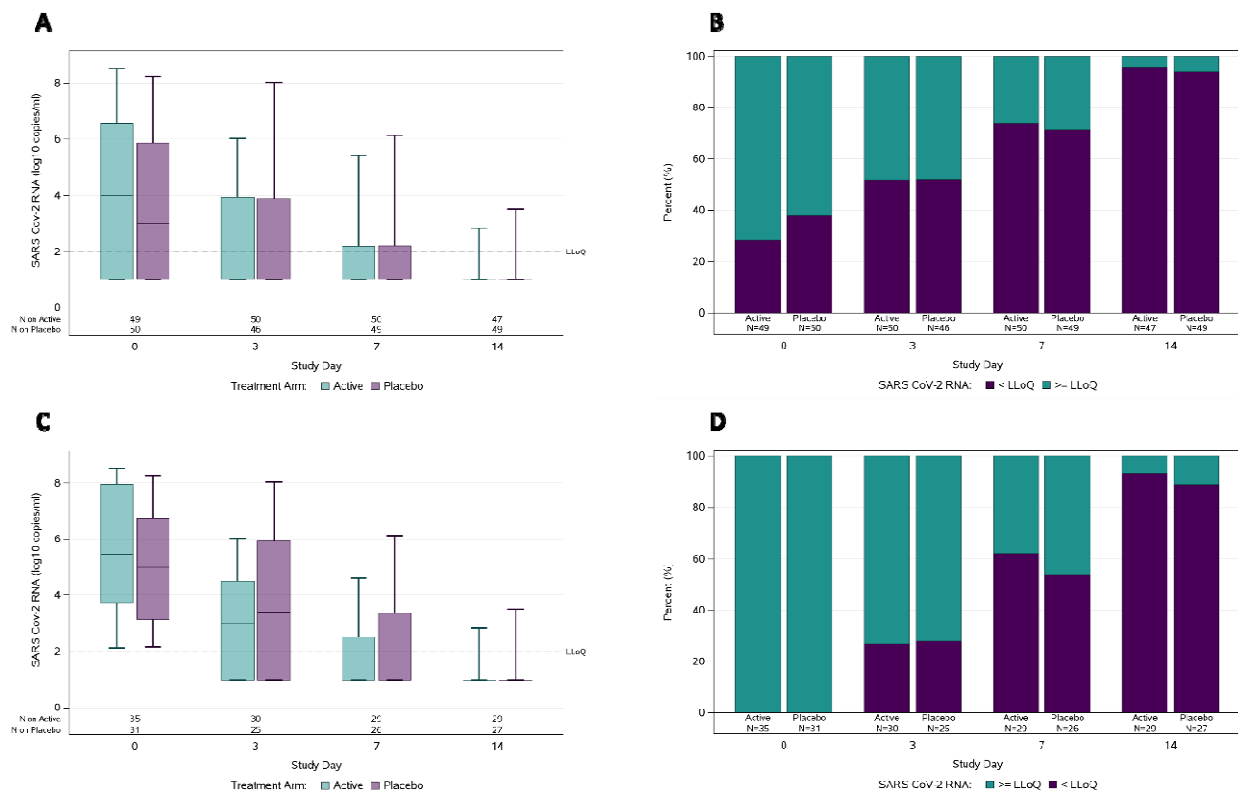
	p=0.26	p=0.052	p=0.33
F: MMRM across Days 3 and 7 with imputation – adjusting for baseline			
LLoQ imputation	-0.65 (-1.36, 0.07) p=0.08	-0.58 (-1.01, -0.15) p=0.009	--
½LLoQ imputation	-0.72 (-1.50, 0.06) p=0.07	-0.71 (-1.29, -0.13) p=0.018	--
G: MMRM for censored data across Days 3 and 7– adjusting for baseline			
N/A	-1.10 (-1.94, -0.26) p=0.011	-1.33 (-2.23, -0.43) p=0.004	--

^aDifferences in mean change provided except for (D), which is difference in median change.

^bResults 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.

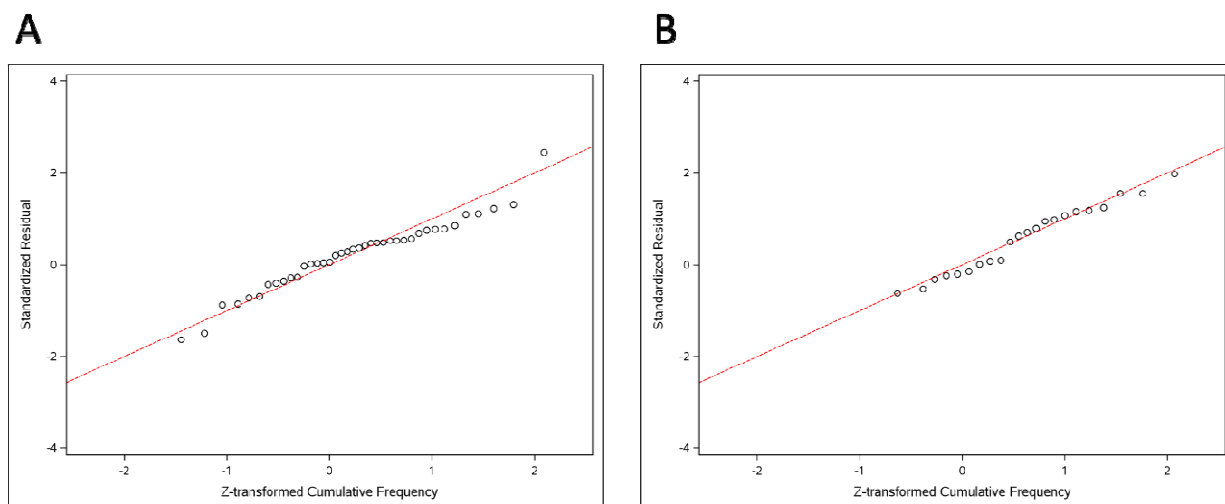
LLoQ = Lower Limit of Quantification; N/A = Not Applicable; NE = Not Estimable; MMRM = Mixed Model for Repeated Measures.

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 \geq LLoQ at Baseline/Day 0 (C and D).



Levels of SARS-CoV-2 RNA (log₁₀ copies/ml) with horizontal line = median, box=interquartile range, whiskers = minimum/maximum (A and C); results below the LLoQ are plotted using an imputed value of 1 log₁₀ copies/ml. Proportion with quantifiable SARS-CoV-2 RNA (green) and unquantifiable (purple) (B and D). LLoQ = Lower Limit of Quantification.

Figure 2: Quantile-quantile (Q-Q) plot for linear regression model for censored data for change in vRNA 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 vRNA



Standardized residuals (for the non-censored 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.