medRxiv preprint doi: https://doi.org/10.1101/2021.01.06.21249368; this version posted January 8, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

Optimal design for phase 2 studies of SARS-CoV-2 antiviral drugs

James A Watson[1,a,b],
Stephen Kissler[c],
Nicholas PJ Day[a,b],
Yonatan Grad[c],
Nicholas J White[a,b]

[1]Corresponding author: jwatowatson@gmail.com

[a]Mahidol-Oxford Tropical Medicine Research Unit, Faculty of Tropical Medicine, Mahidol University, Bangkok, 10400, Thailand; [b]Centre for Tropical Medicine and Global Health, Nuffield Department of Medicine, University of Oxford, New Richards Building, Old Road Campus, Roosevelt Drive, Oxford, OX3 7LG, UK; [c]Department of Immunology and Infectious Diseases, Harvard T. H. Chan School of Public Health, Boston, MA 02115, USA;

Abstract

There is no agreed methodology for pharmacometric assessment of candidate antiviral drugs in COVID-19. The most widely used measure of virological response in clinical trials so far is the time to viral clearance assessed by qPCR of viral nucleic acid in eluates from serial nasopharyngeal swabs. We posited that the rate of viral clearance would have better discriminatory value. Using a pharmacodynamic model fit to individual SARS-CoV-2 virus clearance data from 46 uncomplicated COVID-19 infections in a cohort of prospectively followed adults, we simulated qPCR viral load data to compare type 2 errors when using time to clearance and rate of clearance under varying antiviral effects, sample sizes, sampling frequencies and durations of follow-up. The rate of viral clearance is a uniformly superior endpoint as compared to time to clearance with respect to type 2 error, and it is not dependent on initial viral load or assay sensitivity. For greatest efficiency pharmacometric assessments should be conducted in early illness and daily qPCR samples should be taken over 7 to 10 days in each patient studied. Adaptive randomisation and early stopping for success permits more rapid identification of active interventions.

Background

SARS-CoV-2 infection can be characterised roughly as two overlapping clinical stages. The first stage comprises uncontrolled viral replication. In individuals who have symptomatic COVID-19 illness, timing of peak viral load in the nasopharynx corresponds approximately to the time of symptom onset [1]. The second stage comprises a decrease in the viral load resulting initially from host defence mechanisms. During this second stage a small subset ($< 5\%$, strongly age dependent [2]) of infected individuals progress to severe pneumonia and some of these die. While low dose dexamethasone has been shown to reduce mortality at this late stage by approximately 30% [3], no antiviral drugs have shown unequivocal benefit. New effective therapeutics are urgently needed to treat COVID-19 earlier in the course of infection and thereby prevent hospitalisation and death.

NOTE: This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.

medRxiv preprint doi: https://doi.org/10.1101/2021.01.06.21249368; this version posted January 8, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

It has been hypothesised that effective antiviral drugs administered during the first stage of infection would accelerate virus clearance and thereby reduce the overall viral load. In theory this should reduce the probability of progression to severe COVID-19 illness. Accelerated viral clearance has been used in several preliminary reports of therapeutic interventions as a proxy measure of benefit (i.e. a reduced chance of developing severe COVID-19) [4, 5, 6]. By contrast, if measures of clinical response or clinical progression are used as the primary endpoint then large sample sizes are required to detect meaningful effects as only a minority of symptomatic individuals progress to severe disease. If accelerated viral clearance is indeed predictive of clinical benefit then measurement of viral clearance is much more efficient as a method of 'phase 2' candidate drug screening and dose-finding.

Viral clearance can be summarised in different ways, the most common summaries being the time-to-clearance (the time to reach the lower limit of detection for a given assay) and the rate-of-clearance (the slope of the decline of the log viral load). To optimise the design of a phase 2 trial that uses a summary statistic of viral clearance as the primary endpoint, it is necessary to assess viral clearance dynamics in untreated individuals and to characterise the main sources of variation. We use prospectively gathered data from 46 individuals who had frequent qPCR sampling before and after peak viral load to simulate data and determine optimal trial design parameters [7]. We use a pharmacodynamic model fit to these data to simulate viral clearance data in untreated and treated individuals with varying effect sizes for the hypothetical treatments. The simulations demonstrate that rate of viral clearance is a uniformly better endpoint as compared to time to viral clearance (where better is quantified by the resulting type 2 error), and necessitates shorter follow-up for accurate characterisation. We propose an adaptive randomisation phase 2 trial design for antiviral drugs in which rate of viral clearance is the primary endpoint.

Results

The simulation results presented here are based on the posterior predictive distribution of viral dynamics from a Bayesian pharmacodynamic model fitted to cycle threshold (Ct) data from 46 individuals prospectively followed over time with frequent qPCR measurements [7]. In general the decline in qPCR estimated viral loads in nasopharyngeal swab samples from peak values was first order and monoexponential (and therefore characterised adequately by a rate constant), although in some individuals a second slower phase of reduction was evident. Thus the clearance rate can be estimated from the slope of a simple linear regression onto the qPCR Ct values (the Ct value is proportional to the log viral load). The posterior predictive distribution takes into account measurement variability (note that this could be reduced by using extra information such as urea correction for extracellular fluid volume in nasopharyngeal or oropharyngeal swabs, or host DNA to account for cell content). We simulate only from the posterior predictive distribution for symptomatic individuals, as enrolment in pharmacometric assessments is usually from passive case detection and symptomatic patients represent the main target group for therapeutics. In addition, these data showed that symptomatic individuals have slower viral clearance than asymptomatic individuals [7]. Figure 1A shows example Ct clearance curves for 10 individuals simulated at random. If the antiviral drug reduces viral replication, then the simplest mechanistic model of an effective antiviral treatment effect is parameterised as a multiplicative effect on the clearance rate. Therefore, the effect of an antiviral drug is parameterised as increasing the slope of the clearance curve (inversely related to the clearance half-life). Figure 1B shows how effect sizes (in terms of % increases) translate into average clearance curves. For an effect size of 15%, Figure 1B and 1C show how this translates into endpoint distributions when there is daily sampling for 21 days after the peak viral load. We make the simplifying assumption that all individuals are recruited into the study at peak viral load (although when fitting to real data this assumption would need verification). The effect sizes used can be put into perspective with respect to recent results from

medRxiv preprint doi: https://doi.org/10.1101/2021.01.06.21249368; this version posted January 8, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

monoclonal antibody treatments. The trial of the neutralising monoclonal antibody LY-CoV555 [4] estimated an approximate 15% increase in viral clearance (rate of viral clearance was estimated in the phase 2 trial using the baseline viral load and a single time point at day 11); the trial of the human IgG1 antibody cocktail REGN-COV2 estimated an approximate 20% increase in viral clearance [5].

Rate-of-clearance results in uniformly lower type 2 error than time-to-clearance

We simulated individual patient viral clearance data for varying sample sizes and varying durations of follow-up and compared the power (1 - type 2 error) to detect effects using either time-to-clearance (defined as the first time at which the qPCR Ct value reaches 40) or rate-of-clearance (defined as the slope coefficient estimated from a linear regression of time onto the Ct values < 40). Figure 2 shows that in all scenarios, for the three durations of follow-up considered here (7, 10 or 14 days), rate-of-clearance results in a lower type 2 error (greater power) than time-to-clearance. While the 7 days follow-up misses the clearance time in nearly all individuals, 7 days is sufficient to obtain a reasonable estimate of the clearance slope. Daily qPCR measurements for 10 instead of 7 days can substantially reduce the type 2 error at moderate sample sizes for intermediate effects, but extending this to 14 days results in very small further gains in terms of power.

As there are no unequivocally effective antiviral drugs at present, the pharmacometric properties of an effective treatment are uncertain. Assuming that there is not viral density dependence in clearance rates then the higher the initial viral load the longer it will take for virus to become undetectable. The dependency between baseline viral load and time to clearance can bias results, particularly in small studies (where by chance one group may have higher starting viral burdens). Rate-of-clearance is independent of baseline viral load and so removes this finite sample bias.

Duration of follow-up and frequency of sampling

Figure 2 shows that the marginal gain between 14 versus 10 days of follow-up when using rate-of-clearance as the trial endpoint is small. We looked at marginal gains in more detail, simulating for an effect size of 7.5% acceleration in clearance (increase in slope) and a sample size of 50 individuals per arm. Figure 3 shows that the marginal gains tail off after 9-10 days for the rate-of-clearance endpoint, and tail off at 12-13 days for the time-to-clearance endpoint. We note that these values assume that individuals are recruited at peak viral load.

In SARS-CoV-2 viral clearance rate estimations, more detailed sampling reduces the variation from measurement error and so results in more accurate estimates. Measuring viral densities more than once daily adds logistical difficulties in a clinical trial, but could be worth it if the information gained was substantial. Figure 4 shows small marginal gains in statistical power for twice versus once daily qPCR, suggesting that once daily measurement of qPCR is likely to be sufficient (i.e. an acceptable trade-off between statistical error and resource costs). However, large gains are made by having at least one interim follow-up measurement (estimating the rate from 3 samples versus 2, Figure 4. For small effect sizes (e.g. 7.5% increase in viral clearance rate), estimating clearance rates from daily qPCRs over 10 days considerably increases power over using a single day 10 measurement ($> 80\%$ power compared to less than 50% power for 100 patients per arm).

Adaptive randomisation to identify active antiviral drugs efficiently

Our model simulations suggest that the rate of viral clearance can be estimated from a short duration of follow-up (e.g. 7 to 10 days). For a trial aiming to compare the efficacy of multiple candidate antiviral drugs, the most efficient design to determine the best candidate for a subsequent phase 3 trial is adaptive randomisation (including a control arm), whereby the randomisation

medRxiv preprint doi: https://doi.org/10.1101/2021.01.06.21249368; this version posted January 8, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

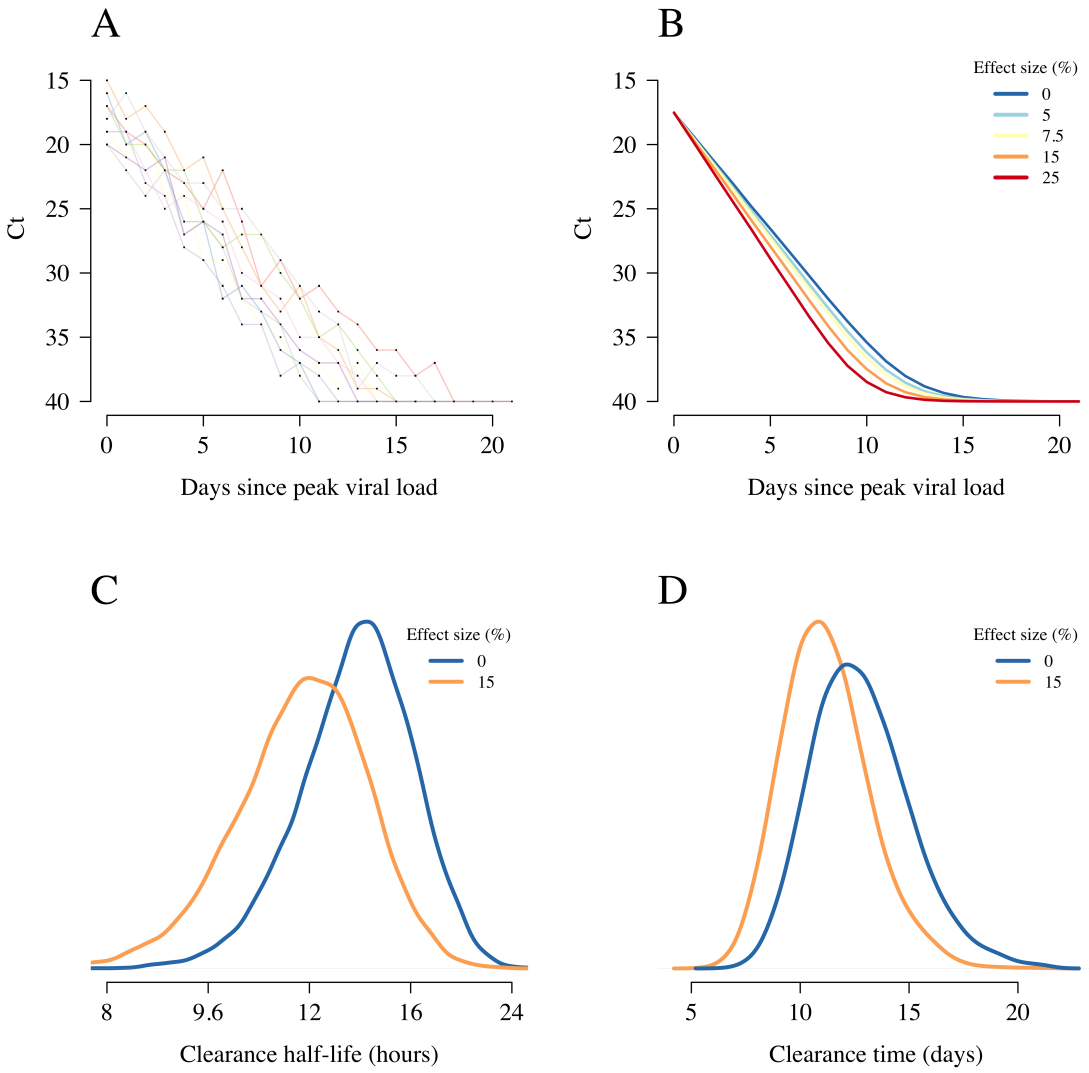

Figure 1: Example of simulated outcomes under the Bayesian pharmacodynamic model of viral dynamics fitted to 47 prospectively studied individuals [7]. A: example Ct curves in 10 simulated individuals. B: how multiplicative effect sizes translate into average clearance curves. C: distribution of clearance half-lives for no drug (blue: effect size equal to 0) and a drug that increases the slope by 15% (orange). D: corresponding distribution of clearance times for the same effect sizes.

medRxiv preprint doi: https://doi.org/10.1101/2021.01.06.21249368; this version posted January 8, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

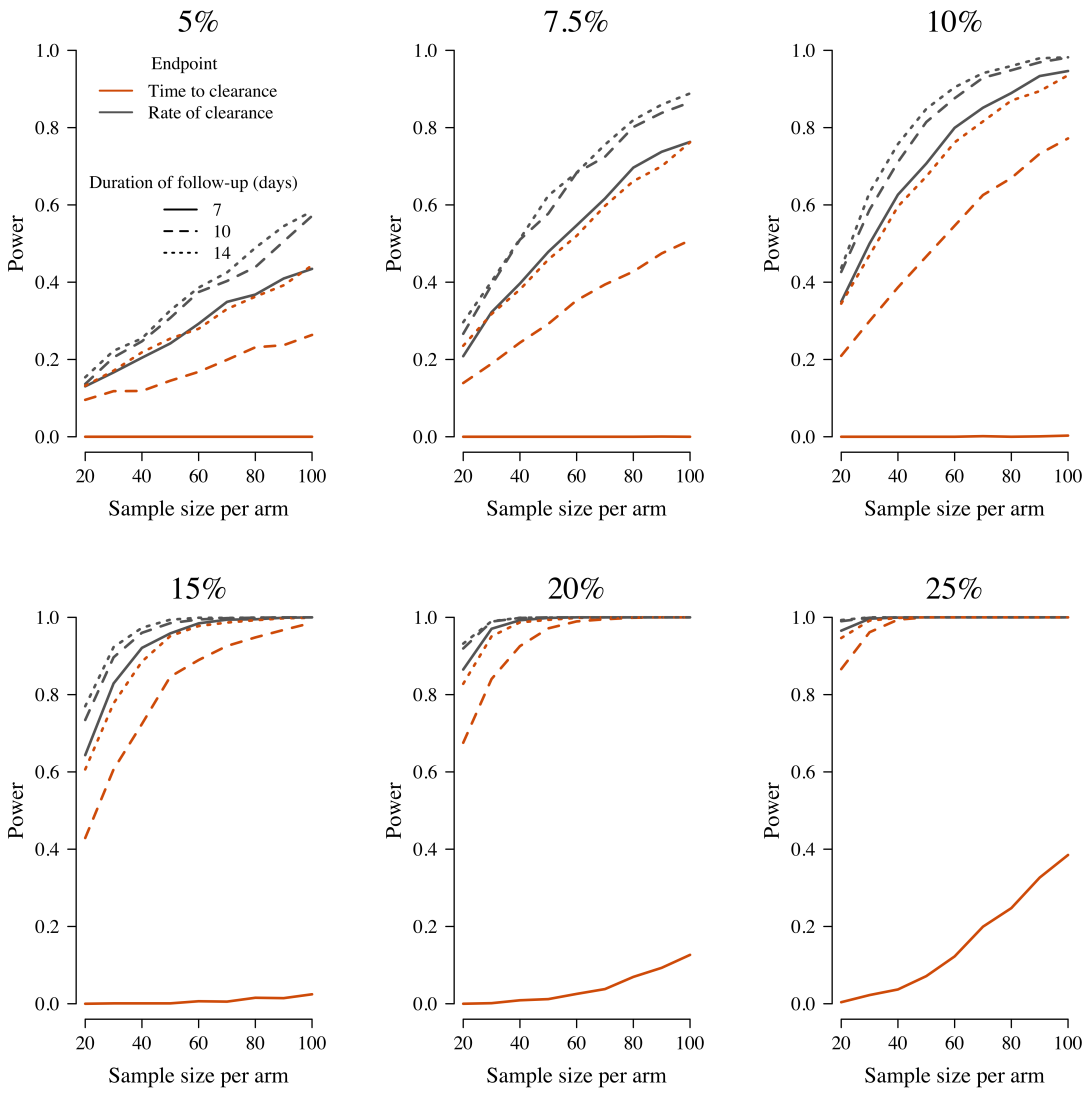

Figure 2: Characterising the statistical power (1-type 2 error) as a function of sample size for the primary trial endpoints time-to-clearance (orange) versus rate-of-clearance (black). For each endpoint, we calculate the power with follow-up durations of 7 days (thick lines), 10 days (dashed lines), or 14 days (dotted lines). The effect sizes are given above each plot as percentage increases in clearance rates.

medRxiv preprint doi: https://doi.org/10.1101/2021.01.06.21249368; this version posted January 8, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

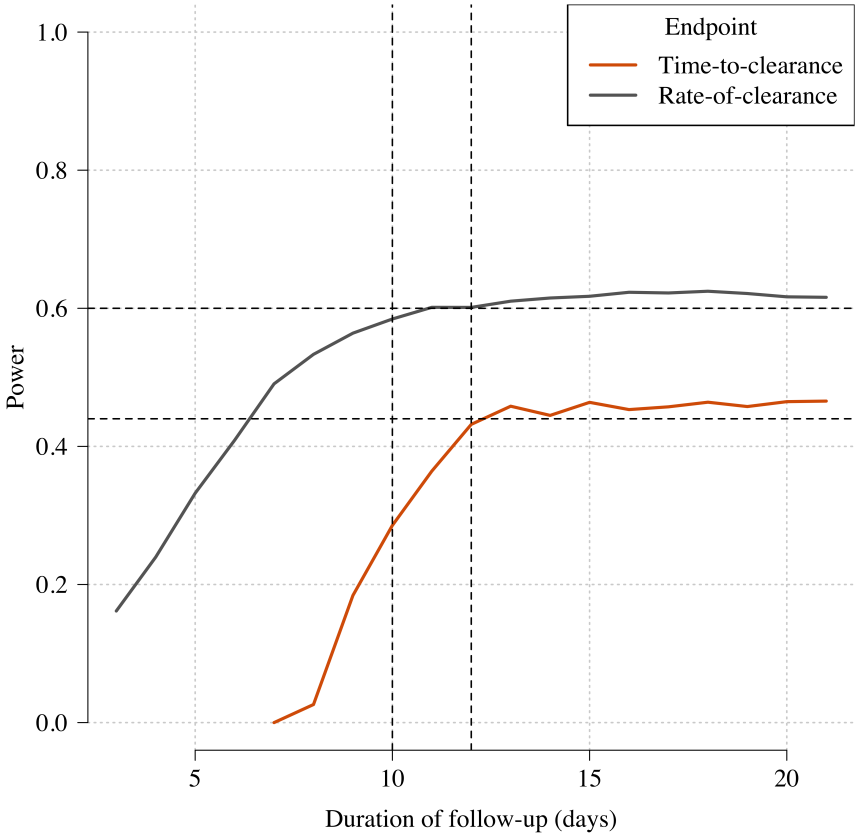

Figure 3: Comparing the effect on statistical power of the duration of follow-up for rate-of-clearance (black) and time-to-clearance (orange). The sample size was fixed at 50 per arm with an effect size of 7.5%.

medRxiv preprint doi: https://doi.org/10.1101/2021.01.06.21249368; this version posted January 8, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

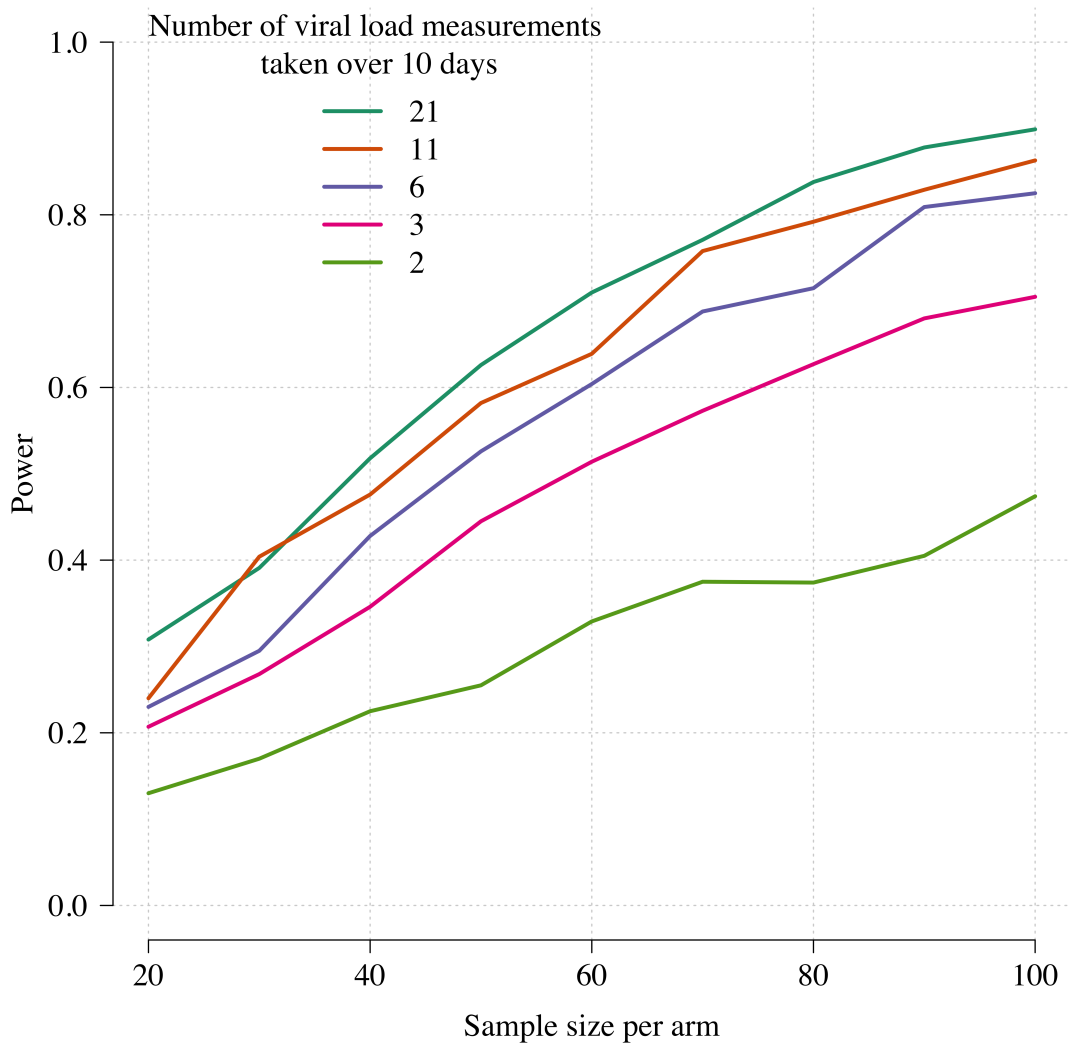

Figure 4: Relationship between the statistical power (1-type 2 error) and the number of qPCR samples taken for an effect size of 7.5% when using rate of decrease in nasopharyngeal viral densities as the trial endpoint. In each simulation, follow-up was 10 days. Viral load measurements vary from twice daily (21 measurements in total, dark green) to once every ten days (i.e. baseline plus day 10 measurement, light green).

medRxiv preprint doi: https://doi.org/10.1101/2021.01.06.21249368; this version posted January 8, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

ratios are adjusted as the trial progresses, based on the estimated drug effects. Adaptive randomisation favours the most promising candidates by increasing the number of individuals who are randomised to that drug. For example, the randomisation probability of the control arm could remain fixed throughout the trial, but the randomisation probabilities of the intervention arms $T_k, k = 1..K$ could be set as proportional to $P\,[\text{clearance rate of } T_k > \text{clearance rate of control}]$. These probabilities can be estimated from a hierarchical model that accounts for inter-individual variability in both peak viral load and rate of clearance with independent additive terms on the clearance rate for each intervention arm. For example, this could specified as:

$$y_{i,t} \sim N\left[\alpha_i + (\beta_i + \beta_{T_i} + \beta)t, \sigma^2_{Ct}\right] \tag{1}$$

where $y_{i,t}$ is the observed Ct value in individual i at time t; T_i is the randomised treatment given to individual i; α_i is the random intercept for individual i (Ct value at time 0); $\beta_i + \beta_{T_i} + \beta$ is the slope estimated for individual i which decomposes as the population mean slope β, a treatment specific random effect β_{T_i} (we set $\beta_{control} = 0$) and an individual random effect β_i; σ_{Ct} is the standard deviation that determines the independent and identically distributed model error. If the individuals included in the trial are all symptomatic, then we can assume that all observed Ct values are above the nadir (i.e., viral loads below the peak) and thus we are estimating a single slope (as opposed to [7] where viral dynamics before and after the peak are inferred). In reality this assumption may not hold and either an ad hoc approach could be used (for example using only the Ct values post observed nadir), or a more complicated piecewise linear model could be fit to the data (as done in [7]).

Adjustment for right truncation should be applied for Ct values above or equal to the limit of quantification (usually set at 40). In addition, the adaptive randomisation design can be improved by specifying a minimum effect size $\lambda > 0$ that is the lower bound of clinically relevant increases in viral clearance. Denoting the probability of being assigned treatment arm k as $\pi(k)$, we propose an adaptive randomisation scheme whereby $\pi(k) \propto P(\beta_{T_k} > \lambda)$, keeping the control arm randomisation ratio fixed (e.g. $\frac{1}{K+1}$ for K intervention arms). $P(\beta_{T_k} > \lambda)$ is estimated from all data accrued so far. As qRT-PCR is widely available, a rapid "turn-around" in reporting clearance rates allowing timely adaption should be feasible.

Figure 5 shows the results from simulation of two trial scenarios which used adaptive randomisation with four intervention arms (5 arms in total including the control). In scenario 1 (top two panels), one of the treatment arms increased the viral clearance rate by 10%, whereas the other three arms had no effect relative to control. In scenario 2 (bottom two panels), all four intervention arms were ineffective (identical to control). We specified a stopping rule for success, defined as a probability greater than 99% that one of the intervention arms had a clearance rate more than 1% faster than the control arm (i.e. $\lambda = 0.01$). Stopping for futility occurred if none of the intervention arms had greater than 10% probability of having more than a 1% faster clearance rate compared to control. Randomisation ratios for the four intervention arms were proportional to the posterior probability that the viral clearance was more than 1% faster than the control arm. These randomisation ratios were updated in batches of 30 patients with an initial 'burn-in' of 50 patients. The maximum trial size was set at 400.

For scenario 1, the median trial size was 140 patients, with a median of 43 patients randomised to the active arm (30% of patients across trials) and 21 patients to each of the inactive arms. The trial identified the active arm in 93% of cases; did not reach the success or futility thresholds by 400 patients in 3.4% of cases, and stopped for futility or identified the wrong intervention arm in 3.7% of cases (from a total of 1000 independent simulations). This gives a type 2 error of 7% for an effect size of 10% increase in viral clearance rate. For scenario 2, stopping for futility by 400 patients occurred in only 2% of cases, no decision was reached in 87% of cases, and one of the four interventions arms was classified wrongly as meeting the success probability threshold in 11% of cases. This gives a type 1 error of 11%. The type 1 and type 2 errors are partly controlled by the

medRxiv preprint doi: https://doi.org/10.1101/2021.01.06.21249368; this version posted January 8, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

value of λ (the clinical significance threshold). Varying λ trades type 1 versus type 2 errors. We chose a low value of $\lambda = 0.01$ in order to have high power but with the drawback of higher type 1 error. This corresponds to the usual goal of a phase 2 trial, which is not to miss effective therapies (as compared to a phase 3 trial, which is not to recommend ineffective therapies).

Discussion

An effective, well tolerated, safe and affordable treatment of COVID-19 that prevented progression to severe disease would be of enormous global health benefit. There have been a large number of small clinical trials assessing the efficacy of repurposed candidate antivirals, based usually on moderate activity in virus cell cultures, but actionable evidence has come only from the relatively few large randomised controlled trials such the RECOVERY and SOLIDARITY platform trials. Among the small molecule antiviral candidates the greatest interest has surrounded remdesivir, a viral RNA-dependent RNA polymerase inhibitor. Remdesivir was associated with a shorter duration of hospitalisation in a randomised controlled trial [8], but did not reduce mortality in the large SOLIDARITY trial [9], nor was there a trend to faster viral clearance in the initial clinical trial in China [10]. This has led the WHO to recommend against its use in COVID-19 [11]. Indeed, no antiviral small molecule drug has yet shown life saving benefit in COVID-19 infections, nor is there any convincing evidence yet that any antiviral drugs do accelerate viral clearance [12]. In most studies describing viral clearance 'rates', it is time to a negative nasopharyngeal swab that is being described without adjustment for the initial viral load. For example, a recent study in Brazil reported faster viral clearance after treatment with nitazoxanide (an antiparasitic drug) compared to placebo [13]. However, the reported comparison did not adjust for the large difference in baseline viral loads between the two groups (median viral densities in nasopharyngeal samples were 0.5 $\log_{10}$ copies per mL higher in the placebo group compared to the nitazoxanide group). The difference in clearance rates estimated from the median values implies only an approximate 1.15% increase in viral clearance rate, which is of uncertain clinical significance.

The most promising evidence for accelerated viral clearance from a therapeutic currently comes from antibody therapies, two of which have recently been given US FDA emergency use authorisations [14, 15]. The monoclonal antibody LY-CoV555 (trade name bamlanivimab) reduced the viral load by day 11 in a phase 2 trial of 452 patients (effect size of approximately a 15% increase in viral clearance rate for the best observed dosing subgroup) [4]. The antibody cocktail REGN-COV2 (a combination of casirivimab and imdevimab) increased viral clearance by approximately 20% in an extended phase 2 trial of 275 patients, with the largest effect on clearance seen in patients with the highest viral loads. Our simulation results show that these large effect sizes could have been characterised with much smaller patient sample sizes using viral clearance as the primary endpoint. There are a large number of both repurposed and novel antiviral drugs either in development or under consideration for COVID-19 prophylaxis or treatment, but there is no agreed methodology for testing them *in vivo* or for determining the optimum dosage. Although COVID-19 is a systemic infection, viral clearance from the upper respiratory tract is the only readily accessible measure of a patient's virological response. Viral clearance reflects both host-defence and any additional contribution from an antiviral therapeutic. Thus, the pattern of clearance will depend on the stage of disease and the host's response. Peak viral densities in nasopharyngeal or oropharyngeal fluids occur around the time of COVID-19 illness onset [1, 12]. Deterioration in oxygenation (reflecting pneumonitis) necessitating hospital admission occurs in a minority of patients approximately one week later at a time when viral densities are usually much lower. Thus assessment of antiviral activity is best performed in subjects at initial presentation of uncomplicated illness. In some individuals, studied at a very early stage of illness, viral densities will rise but once densities fall the pattern is generally first order (log-linear), as it is for most pathogens. In some patients a second slower phase of elimination is evident. In extreme cases there is long lasting viral persistence [16].

medRxiv preprint doi: https://doi.org/10.1101/2021.01.06.21249368; this version posted January 8, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

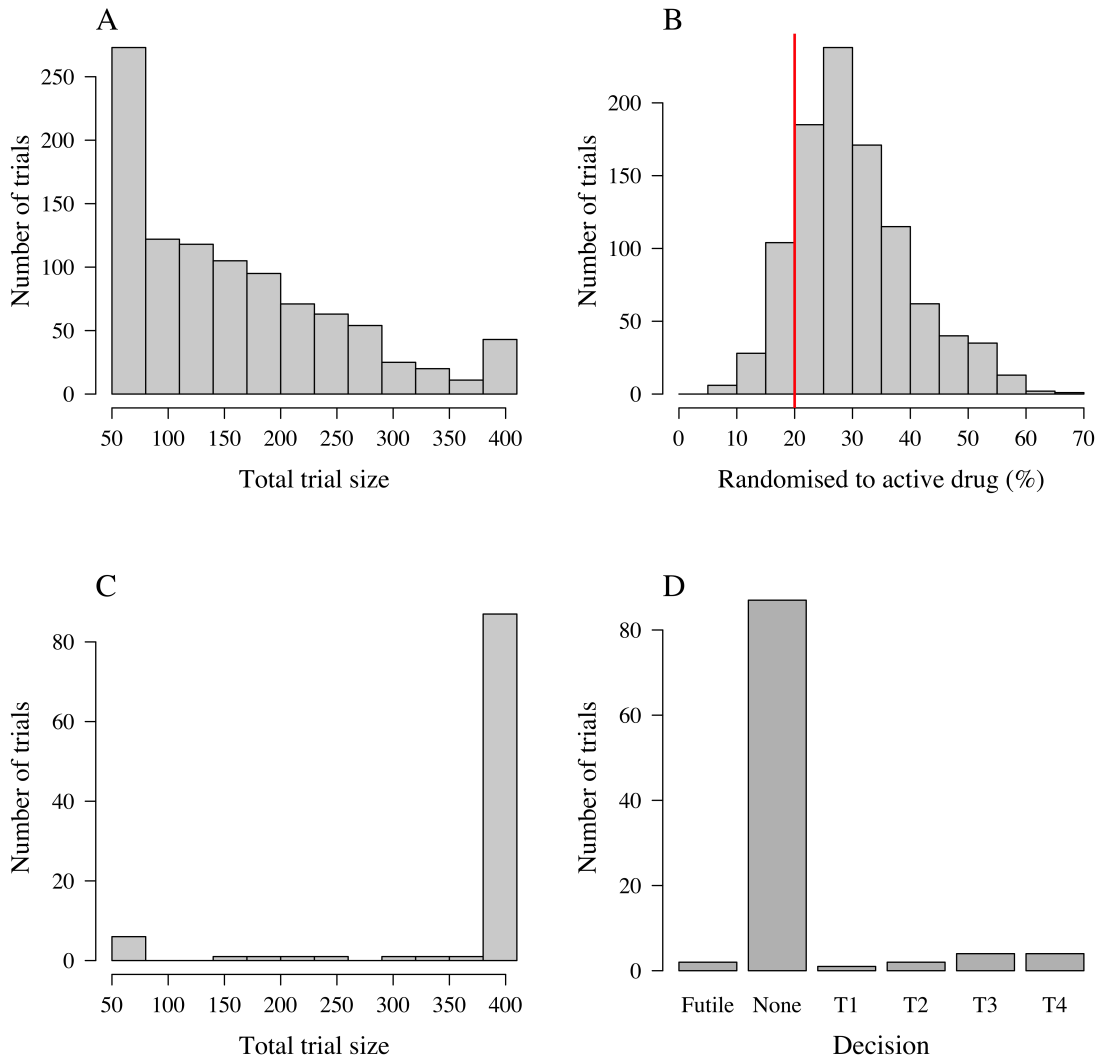

Figure 5: Simulating phase 2 trials with adaptive randomisation, where rate of clearance is the proxy measure of drug efficacy used to specify the randomisation probabilities for the intervention arms. Panels A & B show results from scenario 1 (one active drug out of 4 intervention arms); panels C & D show the results from scenario 2 (no active drugs amongst 4 intervention arms). A: distribution of total trial sample sizes (maximum trial size set to 400) for 1000 scenario 1 simulations. B: distribution of the number of patients randomised to the active arm (randomisation ratio for control was fixed at 20%). The red vertical line shows the randomisation probability under a fixed uniform randomisation scheme. C: distribution of total trial sample sizes (maximum trial size set to 400) in for 100 scenario 2 simulations. D: the decision outcomes of 100 trials simulated under scenario 2. In both scenario 1 & 2, early stopping could occur for (i) success if the probability that one drug resulted in viral clearance greater than 1% increase over control exceeded 99%, or (ii) futility if the probability that any of the drugs resulted in viral clearance greater than 1% increase over control was less than 10%.

medRxiv preprint doi: https://doi.org/10.1101/2021.01.06.21249368; this version posted January 8, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

The slope of the initial log-linear decline in viral densities in nasopharyngeal or oropharyngeal secretions (i.e. the rate constant) is the key measurement of antiviral effect. The difference between the slopes with the putative antiviral and without represents the drug effect. The time to viral clearance has been reported widely in clinical trials, but this pharmacodynamic measure has substantial disadvantages. It is strongly dependent both on the peak or baseline viral densities and the limit of qPCR detection and, by definition, requires follow up until PCR negativity. Evaluations of therapeutic interventions which do not report these values and use time to viral clearance as the primary endpoint are difficult to interpret (e.g. [17]). Time to viral clearance results in a substantially higher type 2 error (lower power) than the clearance rate, and therefore requires much larger sample sizes, at least daily sampling, and a longer duration of follow up. Additionally, unless there is rapid reporting, sampling must continue for much longer than necessary. Thus, use of time to clear as a primary endpoint in pharmacometric assessments is imprecise, inefficient, and expensive. Furthermore, if baseline viral loads differ between treatment groups, then time to viral clearance is biased. Importantly any comparison or meta-analysis of studies which used different sampling techniques, or had different qPCR sensitivities, or different cut-offs, is confounded systematically if time to clearance is used as an endpoint, but not if virus clearance rate is measured.

A swab taken from the nasopharynx or oropharynx contains a variable number of cells and a variable volume of extracellular fluid and nasal secretions. Virus is both intracellular and extracellular. In theory, these factors can be adjusted for by comparison with a human DNA marker and adjusting for urea content respectively. This should reduce intersample variance and improve clearance rate assessments. Nevertheless, even without these adjustments, satisfactory clearance rate estimates can be obtained over a week with as few as six samples. Further studies to examine virus clearance rates at higher peak nasopharyngeal viral densities, and in different patient groups to characterise better the determinants of viral clearance will facilitate interpretation and refine inclusion criteria in phase 2 pharmacometric assessments.

Our model based simulations suggest that pharmacometric assessment of candidate antivirals for COVID-19 should measure virus clearance rate [18], and not the much more widely used time to clearance [19], as their primary endpoint. If performed in early uncomplicated illness, reasonable precision can be obtained with daily qPCR samples over 10 days from each studied patient. Adaptive randomisation using the viral clearance rate can rapidly identify active intervention arms.

Methods

Model of viral dynamics

The pharmacodynamic model of Ct viral load over time is based upon the prospective longitudinal SARS-CoV-2 RT-qPCR testing performed for players, staff, and vendors during the resumption of the 2019-20 National Basketball Association (NBA) season and has been described previously in detail [7]. The code and data are openly accessible on github at `https://github.com/gradlab/CtTrajectories`. The original description of the model had three main individual parameters (using the stan notation): *dp*, *wp*, *wr*. *dp[i]* is the peak viral load for individual i; *wp[i]* is the time until peak viral load; *wr[i]* is the time to recovery (time from peak viral load to time at which viral density is unmeasurable, i.e. Ct ≥ 40). We re-parameterised the stan model (for sampling from the posterior distribution) so that *wr* is replaced by the slope coefficient (equal to -dp/wr). This removes any dependency in the posterior predictive distribution between the clearance rate and the peak viral load. To simulate viral trajectories under the posterior predictive, we draw at random from the top level of the hierarchical model (population distribution of the peak viral load and rate of clearance) and then simulated Ct values from time 0 (time of peak viral load) until a specified maximum time after the peak.

medRxiv preprint doi: https://doi.org/10.1101/2021.01.06.21249368; this version posted January 8, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

Adaptive randomisation

To simulate trials with adaptive randomisation we again used the same model of viral clearance dynamics to simulate trial data, and then used a hierarchical linear model to estimate drug effects. This used *stan_lmer* from the R package *rstanarm* whereby individual random effects were specified for both peak viral loads and viral clearance slopes. This was fitted directly to the simulated Ct values (R formula specification: $y \sim 1 + t * Trt + t - Trt + (1 + t|id)$, where y is the Ct value, t is the timepoint, Trt is the treatment arm, and id is the patient identifier). After a burn-in of 50 patients, randomisation probabilities for the intervention arms were set to be proportional to the posterior probability that the treatment effect was greater than a 1% increase in viral clearance. The posterior distribution over the model parameters was estimated from 1000 draws from 1 chain (to minimise running time).

For scenario 1 we simulated 1000 independent trials (approximately 2 days computation on 8 cores) and for scenario 2 we simulated 100 independent trials (also around 2 days on 8 cores). Trials under scenario 2 (no active drugs) are more computationally expensive because the large majority do not meet any early stopping rules (median trial size is 410, the maximum trial size) whereas under scenario 1 (one active treatment), the majority of trials stop early (median trial size is 140). The posterior fitting is proportional to the number of individuals and so takes longer at each interim analysis.

Code

All the code is available at `https://github.com/jwatowatson/phase2sims`

References

[1] He X, Lau EH, Wu P, Deng X, Wang J, Hao X, et al. Temporal dynamics in viral shedding and transmissibility of COVID-19. Nature Medicine. 2020;26(5):672–675.

[2] Salje H, Kiem CT, Lefrancq N, Courtejoie N, Bosetti P, Paireau J, et al. Estimating the burden of SARS-CoV-2 in France. Science. 2020;.

[3] The RECOVERY Collaborative Group. Dexamethasone in Hospitalized Patients with Covid-19 — Preliminary Report. New England Journal of Medicine. 2020;doi:10.1056/NEJMoa2021436.

[4] Chen P, Nirula A, Heller B, Gottlieb RL, Boscia J, Morris J, et al. SARS-CoV-2 Neutralizing Antibody LY-CoV555 in Outpatients with Covid-19. New England Journal of Medicine. 2020;doi:10.1056/NEJMoa2029849.

[5] Weinreich DM, Sivapalasingam S, Norton T, Ali S, Gao H, Bhore R, et al. REGN-COV2, a Neutralizing Antibody Cocktail, in Outpatients with COVID-19. New England Journal of Medicine. 2020;doi:10.1056/NEJMoa2035002.

[6] Desai A, Gyawali B. Endpoints used in phase III randomized controlled trials of treatment options for COVID-19. EClinicalMedicine. 2020;.

[7] Kissler SM, Fauver JR, Mack C, Tai C, Shiue KY, Kalinich C, et al. Viral dynamics of SARS-CoV-2 infection and the predictive value of repeat testing. medRxiv. 2020;.

[8] Beigel JH, Tomashek KM, Dodd LE, Mehta AK, Zingman BS, Kalil AC, et al. Remdesivir for the Treatment of COVID-19 — Final Report. New England Journal of Medicine. 2020;383(19):1813–1826. doi:10.1056/NEJMoa2007764.

medRxiv preprint doi: https://doi.org/10.1101/2021.01.06.21249368; this version posted January 8, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

[9] WHO Solidarity Trial Consortium. Repurposed antiviral drugs for COVID-19—interim WHO SOLIDARITY trial results. New England Journal of Medicine. 2020;.

[10] Wang Y, Zhang D, Du G, Du R, Zhao J, Jin Y, et al. Remdesivir in adults with severe COVID-19: a randomised, double-blind, placebo-controlled, multicentre trial. The Lancet. 2020;.

[11] Therapeutics and COVID-19: living guideline;. Available from: https://www.who.int/publications/i/item/therapeutics-and-covid-19-living-guideline.

[12] Cevik M, Tate M, Lloyd O, Maraolo AE, Schafers J, Ho A. SARS-CoV-2, SARS-CoV, and MERS-CoV viral load dynamics, duration of viral shedding, and infectiousness: a systematic review and meta-analysis. The Lancet Microbe. 2020;doi:10.1016/S2666-5247(20)30172-5.

[13] Rocco PRM, Silva PL, Cruz FF, Junior MACM, Tierno PFGMM, Moura MA, et al. Early use of nitazoxanide in mild COVID-19 disease: randomized, placebo-controlled trial. medRxiv. 2020;doi:10.1101/2020.10.21.20217208.

[14] Office of the Commissioner. Coronavirus (COVID-19) Update: FDA Authorizes Monoclonal Antibody for Treatment of COVID-19; 2020. Available from: https://www.fda.gov/news-events/press-announcements/coronavirus-covid-19-update-fda-authorizes-monoclonal-antibody-treatment-covid-19.

[15] Office of the Commissioner. Coronavirus (COVID-19) Update: FDA Authorizes Monoclonal Antibodies for Treatment of COVID-19; 2020. Available from: https://www.fda.gov/news-events/press-announcements/coronavirus-covid-19-update-fda-authorizes-monoclonal-antibodies-treatment-covid-19.

[16] Choi B, Choudhary MC, Regan J, Sparks JA, Padera RF, Qiu X, et al. Persistence and Evolution of SARS-CoV-2 in an Immunocompromised Host. New England Journal of Medicine. 2020;383(23):2291–2293.

[17] Mallat J, Hamed F, Balkis M, Mohamed MA, Mooty M, Malik A, et al. Hydroxychloroquine is associated with slower viral clearance in clinical COVID-19 patients with mild to moderate disease: A retrospective study. medRxiv. 2020;.

[18] Trump S, Lukassen S, Anker MS, Chua RL, Liebig J, Thürmann L, et al. Hypertension delays viral clearance and exacerbates airway hyperinflammation in patients with COVID-19. Nature Biotechnology. 2020;doi:10.1038/s41587-020-00796-1.

[19] Kim MS, An MH, Kim WJ, Hwang TH. Comparative efficacy and safety of pharmacological interventions for the treatment of COVID-19: A systematic review and network meta-analysis. PLOS Medicine. 2021;17(12):1–28. doi:10.1371/journal.pmed.1003501.