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 SARS-CoV-2 antiviral drugs. The most widely used measure of virological response in clinical trials to date is the time to viral clearance assessed by serial qPCR. We posited that rate of viral clearance would have better discriminatory value. Using a pharmacodynamic model fit to individual SARS-CoV-2 qPCR data from 46 prospectively followed uncomplicated COVID-19 infections, we simulated viral load data and showed that rate of viral clearance is a uniformly superior endpoint as compared to time to clearance with respect to type 2 error. Rate of clearance is not dependent on initial viral load or assay sensitivity. We apply our approach to data from a randomised controlled trial of ivermectin. Pharmacometric antiviral assessments should be conducted in early illness with serial qPCR samples taken over one week. Response-adaptive trials using rate of clearance can identify promising antiviral interventions rapidly.

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 of infected individuals ($< 5\%$, strongly age dependent [2]) progress to severe pneumonia and some die. While low dose dexamethasone has been shown to reduce mortality at this late stage by approximately 30% [3], and IL-6 receptor antagonists provide additional benefit, 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.

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 being time-to-clearance (the time to reach the lower limit of detection for a given assay) and 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-clearance is a uniformly better endpoint as compared to time-to-clearance (with respect to type 2 error), and necessitates shorter follow-up. We propose a response-adaptive randomisation trial design for phase 2 evaluations of antiviral drugs in which rate-of-clearance is the primary endpoint. We apply our Bayesian analytical framework to data from a randomised placebo-controlled trial of ivermectin and show considerable posterior uncertainty around any *in vivo* antiviral effect.

Results

Model of viral dynamics

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 of the main component can be estimated from the slope of a 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 (which 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. 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 the treatment effect is a multiplicative effect on the clearance rate. Figure 1B shows how effect sizes (in terms of % increases of the slope) 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). These effect sizes can be put into perspective with respect to recent results from 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 (7, 10 or 14 days), rate-of-clearance results in a lower type 2 error (greater power) than time-to-clearance. While 7 days only of 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 small molecule antiviral drugs which are unequivocally effective in-vivo at present, the pharmacometric properties of an effective treatment are uncertain. If 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 biases 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 and a sample size of 50 individuals per arm. Figure 3A shows that the marginal gains tail off after 9-10 days for rate-of-clearance, and tail off at 12-13 days for time-to-clearance. We note that these estimates 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 results in more accurate estimates. Measuring viral densities more than once daily adds logistical difficulties, but could be worthwhile if the information gained was substantial. Figure 3B 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 3B. 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

For a trial that compares the efficacy of multiple candidate antiviral drugs, in order to choose the most promising candidate, the most efficient design with the goal of determining the best candidate for a subsequent phase 3 trial uses response-adaptive randomisation (including a control arm). In this the randomisation ratios are adjusted based on the estimated drug effects. Response-adaptive

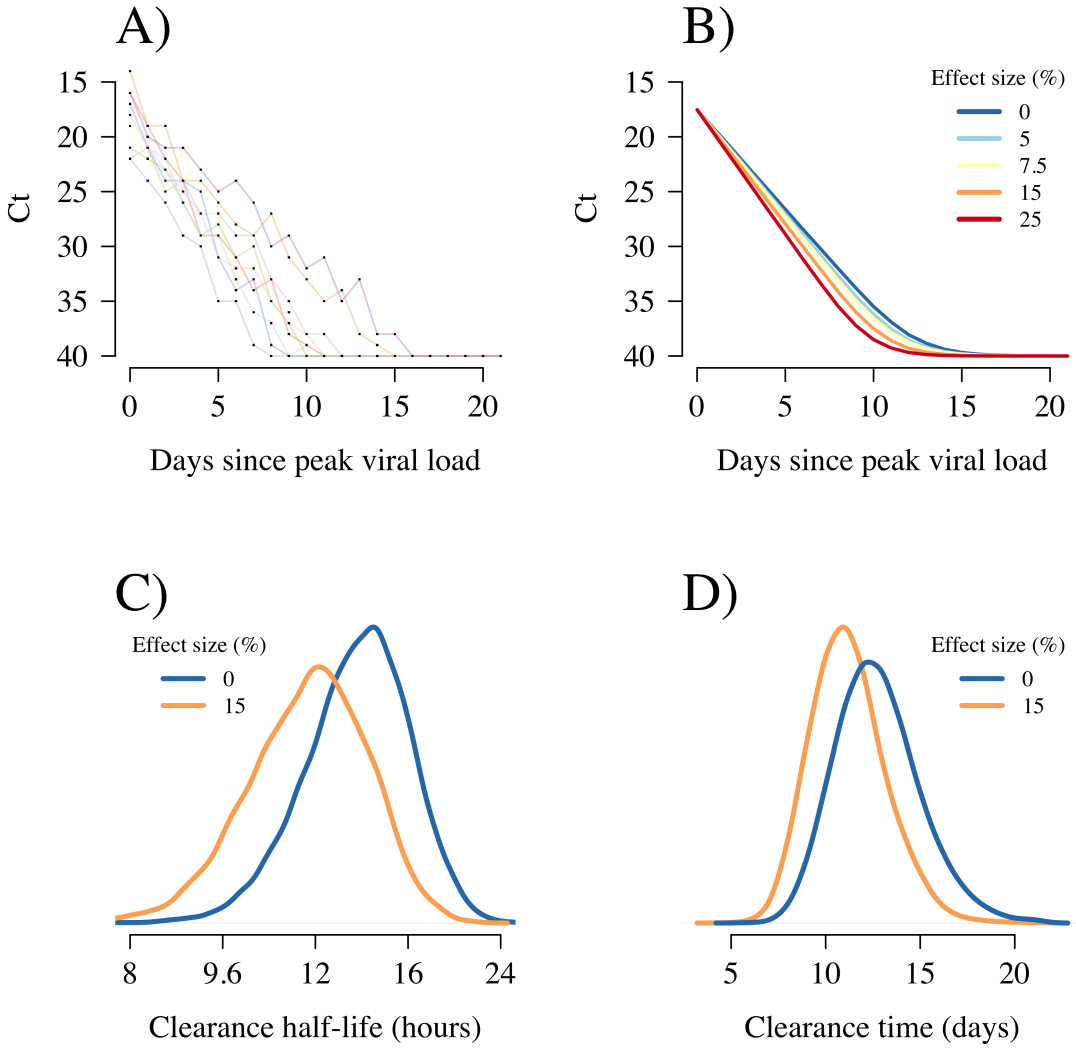

Figure 1: Example of simulated outcomes under the Bayesian pharmacodynamic model of viral dynamics fitted to 47 prospectively studied individuals [7]. A: example Ct curves in 10 simulated individuals. B: how multiplicative effect sizes translate into average clearance curves. C: distribution of clearance 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.

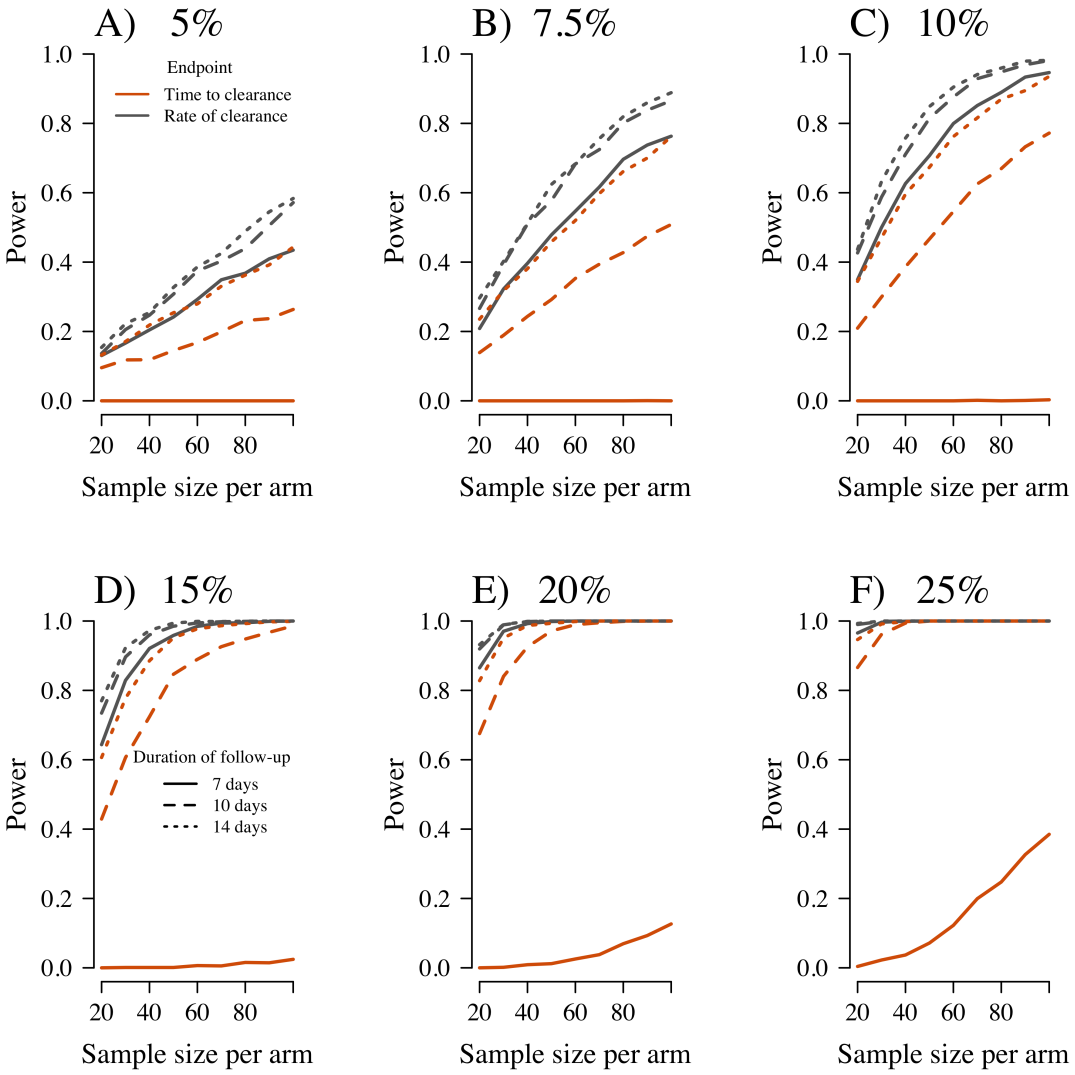

Figure 2: Characterising the statistical power (1-type 2 error) as a function of sample size for the primary trial endpoints time-to-clearance (orange) versus rate-of-clearance (black). For each endpoint, we calculate the power with follow-up durations of 7 days (thick lines), 10 days (dashed lines),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

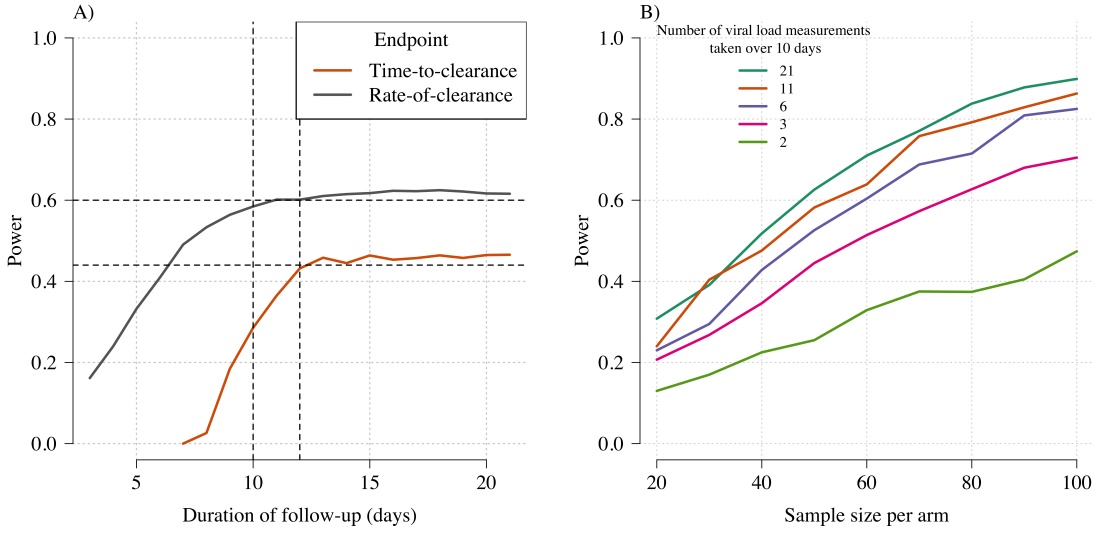

Figure 3: Comparing the effect on statistical power of the duration of follow-up for rate-of-clearance (black) and time-to-clearance (orange). The sample size was fixed at 50 per arm with an effect size of 7.5%. Relationship between the statistical power (1-type 2 error) and the number of qPCR samples 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).

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 [clearance rate of $T_k >$ 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:

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

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 an ad hoc approach could be used (for example using only the Ct values post observed nadir), or a more complicated piece-wise linear model could be fitted to the data [7].

The model should right censor explicitly Ct values above or equal to the lower limit of quantification. This is usually set at a cycle threshold of 40, however, it may be justified to use lower cycle thresholds in order to discount the effect of persistent viral shedding (i.e. a slower second phase - see the ivermectin analysis in the next section). 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).

Figure 4 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 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).

Bayesian analysis of a randomised controlled trial of ivermectin

Ivermectin is a broad spectrum antiparasitic, anthelmintic and endectocide drug with *in vitro* antiviral activity that includes activity against SARS-CoV-2 [8]. Many countries have recommended ivermectin for the treatment of COVID-19, notably in Latin America, but there is as yet no convincing evidence that it has a significant antiviral effect *in vivo*. We used serial qPCR data from a recent pilot randomised controlled trial of ivermectin versus placebo in 24 symptomatic COVID-19 patients [9] to demonstrate how individual rates of clearance can be estimated under a hierarchical Bayesian linear model and used to calibrate the resulting uncertainty around putative antiviral effects. These data also demonstrate how time to clearance as a measure of virological effect is highly sensitive to the lower limit of detection.

Figure 5A shows the individual viral load data (mean Ct values from two separate PCRs on two different genes). There is a clear bi-exponential decline in viral loads, with the majority of patients showing persistent viral shedding on days 13 and 20. All the Ct values for the day 20 qPCR are between 32 and 40. This bi-exponential decline is not captured by our simple log-linear model and thus clearance rate estimates would be biased downwards if we fit a simple linear model using all the data. We fit a Bayesian hierarchical linear model to the observed Ct values censoring all values ≥ 32, and using only the data from days 0-6. In this dataset, single exponential decline is a reasonable approximation of viral clearance in the first week. We used an informative prior on the treatment effect (parameterised as a multiplicative effect on the individually estimated slope coefficient) defined as a normal distribution with mean 0 and standard deviation 0.1. This prior implies that treatment effects greater than 20% increase in the slope of viral clearance had less than 5% prior mass. This informative prior is justified from the SARS-CoV-2 antibody studies, suggesting maximum effects of around 20% [4, 5]. Our analysis of the qPCR data from these 24 patients (12 were randomised to ivermectin, 12 to placebo) results in a posterior distribution over the effect of ivermectin with 75% of the mass greater than 0 (i.e. increased clearance, Figure 5B) and a mean effect of 5% increase in viral clearance. This does suggest the possibility of a small *in vivo* antiviral effect that could justify further phase 2 clinical trials, but it highlights the large uncertainty around this effect. Posterior individual patient fits and comparison of prior and posterior distributions are shown in Figures S1 and S2.

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 [10], but did not reduce mortality in the large SOLIDARITY trial [11], nor was there a trend to faster viral clearance in the initial clinical trial in China [12]. This has led the WHO to recommend against its use in COVID-19 [13]. Indeed, no antiviral small molecule drug has yet shown life saving benefit in COVID-19 infections, nor is there any definitive evidence yet that any antiviral drugs accelerate viral clearance [14], although a randomised placebo controlled trial of peginterferon lambda (a large molecule biological antiviral)

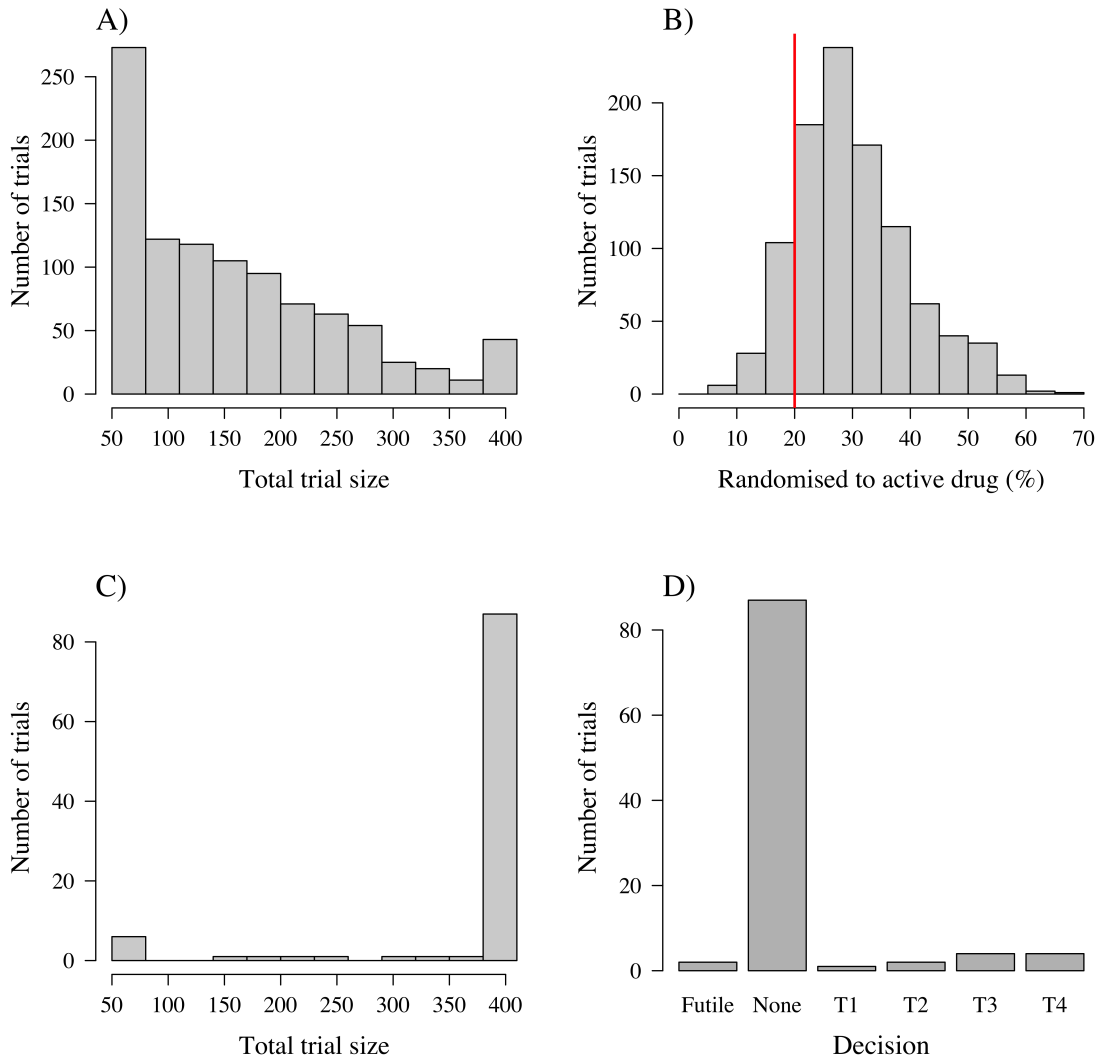

Figure 4: Simulating phase 2 trials with adaptive randomisation, where rate of clearance is the proxy measure of drug efficacy used to specify the randomisation probabilities for the intervention arms. Panels A & B show results from scenario 1 (one active drug out of 4 intervention arms); panels C & 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%.

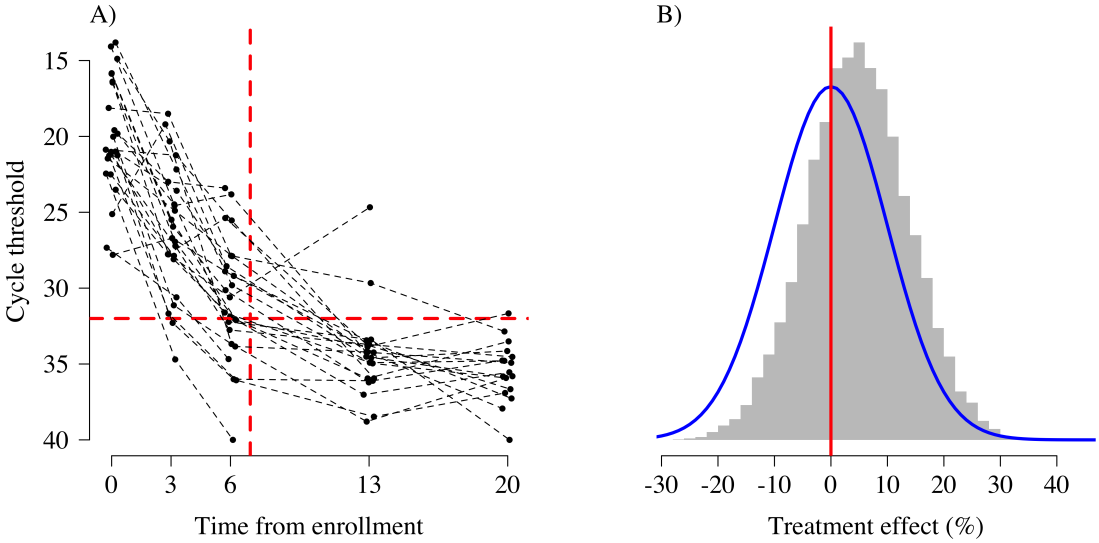

Figure 5: Post-hoc Bayesian analysis of serial viral load qPCR data from a pilot randomised controlled trial of ivermectin versus placebo ($n = 24$) [9]. Panel A shows the mean Ct values (average of qPCRs on two different genes) at each time point (random jitter added on the x-axis for visual clarity). The vertical and horizontal lines show the thresholds used to analyse the data. All Ct values $\geq$32 are considered right censored; only data from days 0-6 are used to fit the model. Panel B shows the posterior distribution over the treatment effect (histogram) with the prior distribution (blue density line).

in 60 COVID-19 outpatients showed greater viral decline in the peginterferon group [15]. We note that the primary endpoint in this trial was the proportion qPCR negative 7 days after enrolment, an endpoint which has the same disadvantages as time to clearance. In addition, 15 out of the 60 patients had undetectable virus at baseline [15] emphasising the need to select patients shortly after symptom onset. The final analysis did adjust for baseline viral load but most studies describing viral clearance 'rates' are reporting the time to a negative nasopharyngeal swab without adjustment for the baseline viral load. For example, a recent study in Brazil reported faster viral clearance after treatment with nitazoxanide (an antiparasitic drug) compared to placebo [16]. 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 [17, 18]. 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. We propose using the slope of the initial log-linear decline in viral densities in nasopharyngeal or oropharyngeal secretions (i.e. the rate constant of the major decline) as the primary endpoint in phase 2 studies. The difference between the slopes with the putative antiviral and without represents the drug effect. Time-to-clearance has been reported widely in clinical trials, but this pharmacodynamic measure has substantial disadvantages. It is dependent both on the baseline viral load 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. [19]). Time-to-clearance results in a substantially higher type 2 error (lower power) than rate-of-clearance, and therefore requires larger sample sizes, at least daily sampling, and longer duration of follow-up. Thus, the time-to-clearance endpoint is imprecise, inefficient, and expensive. 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 rate-of-clearance is measured.

Our model based simulations suggest that pharmacometric assessment of candidate antivirals for COVID-19 should measure virus clearance rate [20], and not the much more widely used time to clearance [21], as their primary endpoint. If performed in early uncomplicated illness, reasonable precision can be obtained with daily qPCR samples taken over the course of one week after enrolment in 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 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.

Simulations of type 2 error

We simulated serial qPCR data for varying durations, treatment effects and frequency of qPCR sampling. Clearance rates were estimated by simply fitting a linear model to each individual independently, using all Ct values less than 40. Difference in clearance rates were tested using a t-test comparing the distributions of individual clearance rates in the treated and not treated groups. Time to clearance was defined as the first timepoint such that the Ct value was equal to 40. If this occured after the maximum follow-up duration, the time was considered right censored. Differences in time to clearance were tested using a log-rank test on the Kaplan-Meier survival curves from the time to event data. In both cases a significance level $\alpha = 5\%$ was used to define a significant difference.

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* [22] 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 (the median trial size was 410 which is the maximum trial size) whereas under scenario 1 (one active treatment), the majority of trials stop early (median trial size was 140). The computational complexity of the posterior fitting is proportional to the number of individuals and so takes longer at each interim analysis.

Bayesian analysis of a randomised controlled trial of ivermectin in COVID-19

Data from the randomised controlled trial of ivermectin versus placebo are publically available with the published study report [9]. All details of the clinical study can be found in the published report. Briefly, outpatients with symptomatic COVID-19 (less than 72 hours since onset of symptoms) were randomised 1:1 to receive ivermectin as a single dose 400 μg/kg (n = 12) or placebo (n = 12). The primary trial endpoint was qPCR positivity at day 7. All 24 patients remained qPCR positive at day 7.

We fit a simple hierarchical linear model to the serial Ct values (these are the mean Ct values from PCRs on the SARS-CoV-2 E gene and N gene [9]) where each patient has a random slope and a random intercept (parameterised as independent random effects). Let y_t^i be the Ct value for patient i at time t. We define $C_t^i = \mathbb{1}(y_i^t < 32)$ ($\mathbb{1}$ is the indicator function, taking the value 1 if $y_i^t < 32$, 0 otherwise). $T_i \in \{0, 1\}$, where 1 is treated (randomised to ivermectin) and 0 is control (randomised to placebo).

We define:

$$\mu_t^i = \beta_0 + \beta_0^i + (\beta_t + \beta_t^i)(1 + \beta_T T_i)t$$

μ_t^i is the expected Ct value under the model, whereby the treatment effect β_T has a fixed multiplicative effect on the individually estimated slope. $100\beta_T$ is therefore the percentage increase in slope in the treated individuals.

The likelihood is defined as:

$$y_t^i \sim \text{Student}(\mu_t^i, \sigma, 7)$$

For censored values ($C_t^i = 0$) the likelihood is defined as $1 - F(32; \mu_t^i, \sigma, 7)$ where F is the student-t cumulative distribution function. We specify the residual error model as a student-t distribution with 7 degrees of freedom as a default choice for robust heavy-tailed regression.

The priors are defined as (these are also shown graphically in Figure S2):

$$\beta_0 \sim \text{Normal}(20, 3)$$

$$\beta_0^i \sim \text{Normal}(0, \tau_0)$$

$$\tau_0 \sim \text{Normal}(3, 2)$$

$$\beta_t \sim \text{Normal}(2, 1)$$

$$\beta_t^i \sim \text{Normal}(0, \tau_t)$$

$$\tau_t \sim \text{Normal}(0.5, 1)$$

$$\sigma \sim \text{Normal}(1, 1)$$

$$\beta_T \sim \text{Normal}(0, 0.1)$$

All of these priors are informative, using knowledge from the prospectively gathered data [7]. The informative prior on the treatment effect is important (we put 95% of the prior mass on beneficial or negative effects between 20% decrease in viral clearance and 20% increase in viral clearance.

The posterior sampler was written in *stan* implemented in *rstan* using *R* version 4.0.2 [23].

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] Caly L, Druce JD, Catton MG, Jans DA, Wagstaff KM. The FDA-approved drug ivermectin inhibits the replication of SARS-CoV-2 in vitro. Antiviral research. 2020;178:104787.

[9] Chaccour C, Casellas A, Blanco-Di Matteo A, Pineda I, Fernandez-Montero A, Ruiz-Castillo P, et al. The effect of early treatment with ivermectin on viral load, symptoms and humoral response in patients with non-severe COVID-19: A pilot, double-blind, placebo-controlled, randomized clinical trial. EClinicalMedicine. 2021;32. doi:10.1016/j.eclinm.2020.100720.

[10] 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.

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

[12] 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;.

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

[14] 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.

[15] Feld JJ, Kandel C, Biondi MJ, Kozak RA, Zahoor MA, Lemieux C, et al. Peginterferon lambda for the treatment of outpatients with COVID-19: a phase 2, placebo-controlled randomised trial. The Lancet Respiratory Medicine. 2021;doi:10.1016/S2213-2600(20)30566-X.

[16] 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.

[17] 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.

[18] 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.

[19] 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;.

[20] 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.

[21] 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.

[22] Goodrich B, Gabry J, Ali I, Brilleman S. rstanarm: Bayesian applied regression modeling via Stan.; 2020. Available from: https://mc-stan.org/rstanarm.

[23] Stan Development Team. RStan: the R interface to Stan; 2020. Available from: http://mc-stan.org/.

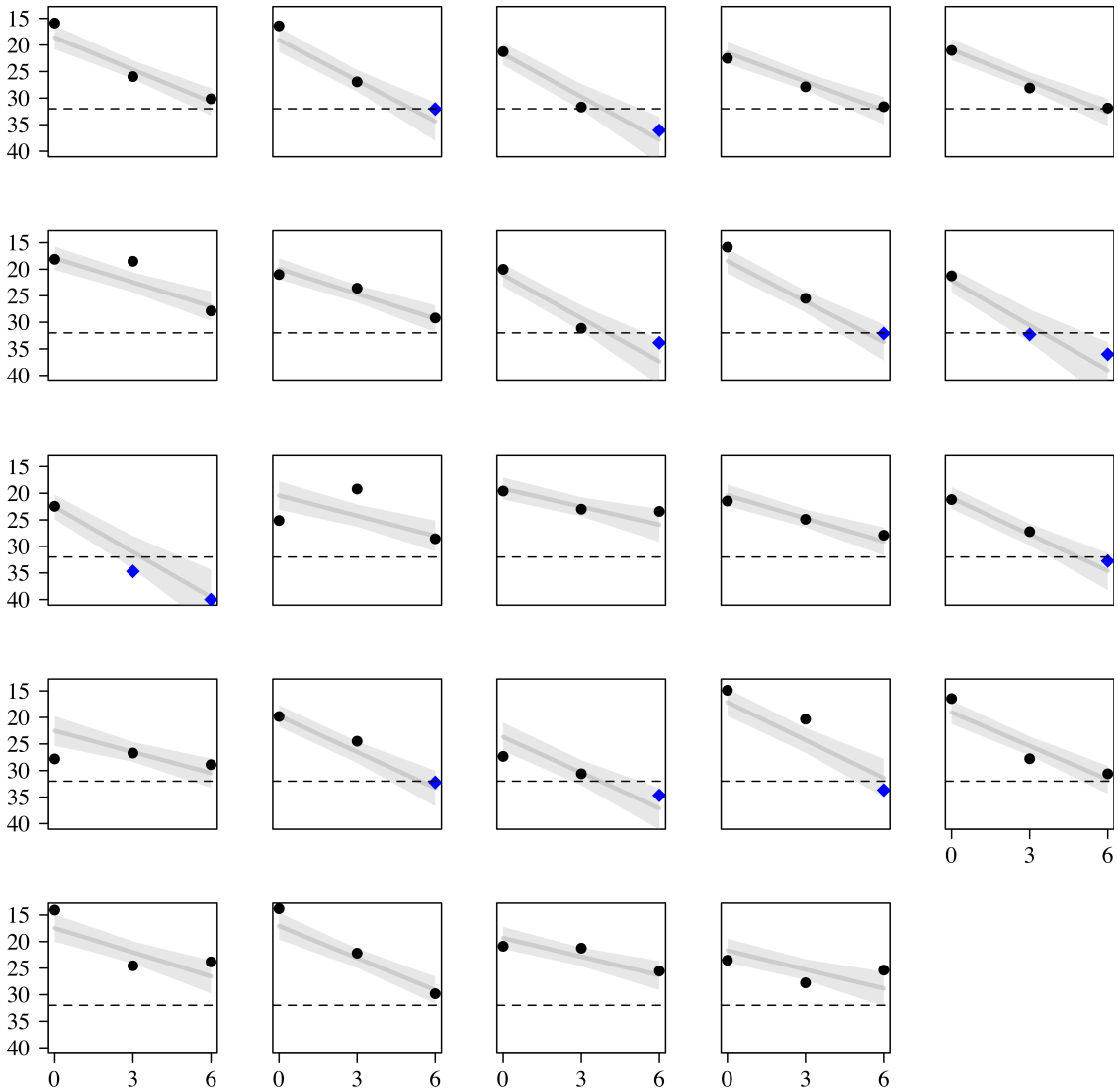

Figure S1: Individual model fits from the Bayesian post hoc analysis of individual patient data ($n = 24$) from a randomised controlled trial of ivermectin versus placebo [9]. In each panel the x-axis is the time since enrolment (starting at day 0) and the y-axis is the Ct values. The black circles show observed mean Ct values < 32 and the blue diamonds are mean Ct values ≥ 32 which are treated as censored observations in the model. The grey lines (grey shaded areas) show the mean (90% posterior uncertainty interval) individual clearance slopes.

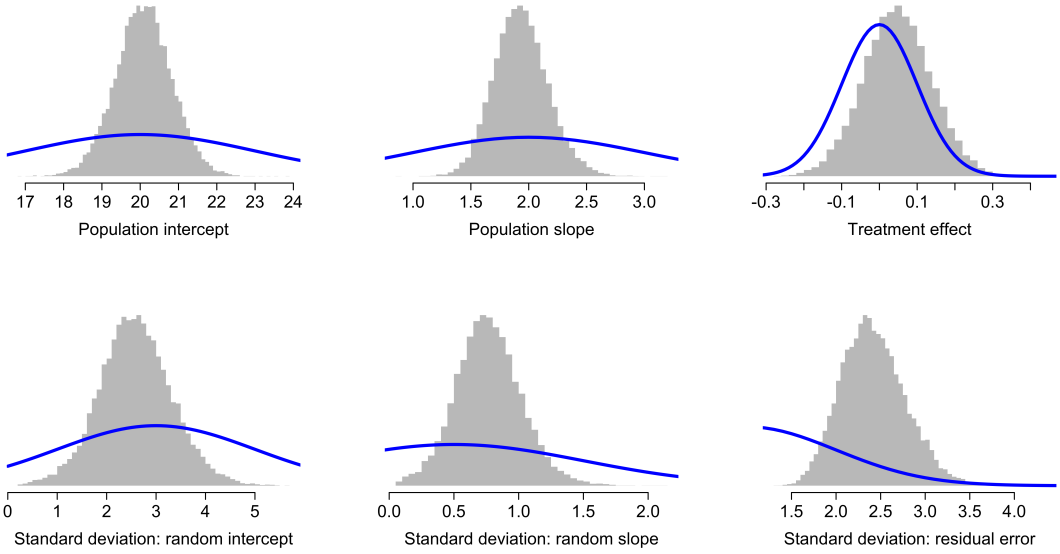

Figure S2: Prior and posterior distributions for the Bayesian hierarchical linear model fit the ivermectin data [9].