Analysis plan

Does Sarilumab Reduce Mortality in Hospitalized Patients for COVID-19? A Bayesian Reanalysis

Arthur M. Albuquerque[1], Todd C. Lee[2], Donald R. Williams[3], Lucas Tramujas[4], and James M. Brophy[5]

[1]School of Medicine, Universidade Federal do Rio de Janeiro, Brazil
[2]Division of Infectious Diseases, Department of Medicine, McGill University, Canada
[3]University of California, Davis, U.S.A.
[4]HCor Research Institute, Brazil
[5]McGill University Health Center, Canada

July 13, 2021

Abstract

This is the preregistered protocol for a meta-analysis on the effect of sarilumab in mortality in COVID-19. We followed the PRISMA-P guidelines for preparing this document.[1]

Introduction

Rationale

COVID-19 is an inflammatory mediated disease. In light of this correlation, several randomized clinical trials (RCTs) have tested whether anti-inflammatory medications, such as interleukin-6 antagonists, can reduce mortality in COVID-19 patients.

On July 6, 2021, a meta-analysis authored by the WHO Rapid Evidence Appraisal for COVID-19 Therapies Working Group was published evaluating the association between IL-6 antagonists and all-cause mortality in patients hospitalized for COVID-19.[2] Results showed that this drug class might reduce mortality, which led WHO to start recommending the use of sarilumab and tocilizumab in patients hospitalized for COVID-19.[3]

Tocilizumab's overall effect was 0.83 odds ratio (95% confidence interval [CI] 0.74 - 0.92), which contributed to a total of 81.61% of weight to the meta-analysis on all-cause mortality ([2], Figure 1). On the other hand, sarilumab's overall effect was 1.08 (95%CI 0.86 - 1.36). Thus, current evidence on sarilumab is equally compatible with 14% mortality odds reduction and 36% mortality odds increase compared to the control group.[4]

Although tocilizumab and sarilumab are both IL-6 antagonists, one could argue that both drugs are not equivalent and thus must be separately analyzed. According to this interpretation and current evidence on sarilumab,[2] future research should determine whether sarilumab, in fact, reduces mortality in patients hospitalized for COVID-19. Therefore, we sought to analyze sarilumab's effect further while considering possible differences between this drug and tocilizumab. In this document, we will describe our analysis plan.

Objectives

In this meta-analysis, we will address the following questions about COVID-19 patients on supplemental oxygen:

1. What is the effect of sarilumab in reducing mortality?
2. Is sarilumab noninferior to tocilizumab in reducing mortality?
3. Is sarilumab equivalent to tocilizumab in reducing mortality?

Methods

Eligibility criteria

Types of studies

Randomized controlled trials.

Types of participants

Adult COVID-19 hospitalized patients on supplemental oxygen.

Types of outcome variables

28-day all-cause mortality.

Information sources

We will retrieve randomized controlled trials (RCT) included in a recent meta-analysis.[2]

Study records

Data management

We will store data in csv format.

Selection process

Two independent reviewers (AMA and LT) will identify the RCTs included in a recent meta-analysis.[2]

Data collection process

We will extract the number of deaths and total patients in each treatment arm from the meta-analysis.[2]

Data items

Regarding general trial characteristics:

1. Trial name
2. Registration number
3. Year
4. Trial design

5. Type of control treatment, such as placebo or standard of care
6. Tocilizumab and/or sarilumab dosage

Regarding overall patient characteristics in each treatment arm:

1. Use of corticosteroids (%)
2. Supplemental oxygen flow rate $\leq$ 15 L/min at randomization (%)
3. Non-invasive ventilation at randomization (%)
4. Invasive mechanical ventilation at randomization (%)

Regarding the outcome in each treatment arm:

1. Number of deaths and total patients in each treatment arm

Risk of bias in individual studies

We will retrieve RCTs already included in a published meta-analysis.[2] The risk of bias of each RCT has already been evaluated in this meta-analysis with the Cochrane Risk of Bias version 2 (RoB 2) tool.[5] Thus, we will not assess the risk of bias of each RCT again.

Data synthesis

We will conduct our analyses using R and fit our Bayesian models using *Stan* through the *brms* package.[6, 7]

Estimate of interest

We will use the odds ratio as our estimate of interest. In this work, treatment benefit will be indicated by an odds ratio lower than 1.

We will perform our main analyses using crude log odds ratio (mean y + standard error σ), which will be derived from number of deaths and total patients in each treatment arm (tocilizumab vs. control or sarilumab vs. control).

Sarilumab's effect

Our first research question is:

- What is the effect of sarilumab in reducing mortality?

To answer this question, **we will fit a Bayesian random-effects meta-analysis**,[8] described as:

$$\begin{aligned} y_i &\sim Normal(\theta_i, \sigma_i^2) \\ \theta_i &\sim Normal(\mu_{sarilumab}, \tau^2) \end{aligned} \tag{1}$$

where y_i is the observed mean log odds ratio of sarilumab versus control in study i. We assume these effect sizes are normally distributed around the study-specific mean θ_i along with a known sampling variance, represented by the observed σ_i^2. We also assume θ_i is drawn from a normal distribution

where $\mu_{sarilumab}$ is the average effect of interest and τ^2 is the between-study heterogeneity. Thus, there are two sources of variation: σ^2 and τ^2.

Because we will fit a Bayesian model, we should assign a prior distribution to each parameter. We will apply weakly informative priors for μ and τ,[9, 10] visually displayed below:

$$\begin{aligned} \mu_{sarilumab} &\sim \text{Normal}(0, 1.5^2) \\ \tau &\sim \text{Half-Normal}(0.5) \end{aligned} \tag{2}$$

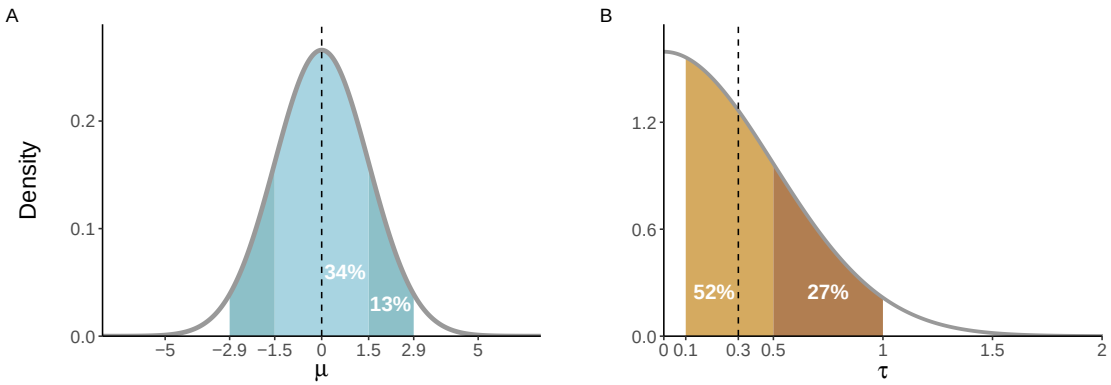

Figure 1: Prior distributions for $\mu_{sarilumab}$ and τ. Vertical dashed lines depict the median value. (A) The Normal(0, 1.5^2) distribution approximately concentrates 34% of probability density between -1.5 and 0 (0.22 and 1 in the normal scale, respectively) and 34% of probability density between 0 and 1.5 (1 and 4.48 in the normal scale, respectively). Moreover, this distribution approximately concentrates 13% between -2.9 and -1.5 (0.06 and 0.22 in the normal scale, respectively) and 13% between 1.5 and 2.9 (4.48 and 18.2 in the normal scale, respectively). (B) The Half-Normal(0.5) distribution concentrates 52% of probability density between 0.1 and 0.5, which Spiegelhalter et al.[11] categorized as a "reasonable heterogeneity" range and 27% between 0.5 and 1, categorized as a "fairly high heterogeneity" range.

Noninferiority analysis

- Is sarilumab noninferior to tocilizumab in reducing mortality?

One major criteria in noninferiority analyses is that superiority of active comparator — e.g., tocilizumab — in comparison to control — e.g., usual care — has been shown.[12]. Previous meta-analyses have presented results in favor of tocilizumab's superiority.[2, 13] Thus, we will use the following formula to estimate our noninferiority margin:[14]

$$\gamma = \left(\frac{1}{\theta}\right)^{1-x} \tag{3}$$

where γ is the estimated noninferiority margin, θ is the tocilizumab's mean effect size, and x is the percentage of tocilizumab's effect that is desired to be preserved. We derived θ from the most recent meta-analysis assessing tocilizumab's effect.[2] More specifically, we derived the mean odds

ratio regarding tocilizumab's effect in the subgroup of patients using corticosteroids. We chose this subgroup because current guidelines suggest that all COVID-19 patients on supplemental oxygen should also be on corticosteroids. We will follow the U.S. Department of Health and Human Services Food and Drug Administration's recommendation and define x as 50% for our primary noninferiority analysis.[12, 15] Thus, our noninferiority margin was estimated as:

$$\gamma = \left(\frac{1}{0.77}\right)^{0.5} = 1.139606 \tag{4}$$

Next, **we will fit a Bayesian random-effects meta-regression model**, described as:

$$\begin{aligned} y_i &\sim Normal(\theta_i, \sigma_i^2) \\ \theta_i &\sim Normal(\mu, \tau^2) \\ \mu &= \beta_0 + \beta_1 x \end{aligned} \tag{5}$$

where y_i is the observed mean log odds ratio in study i for either tocilizumab versus control or sarilumab versus control. We assume these effect sizes are normally distributed around the study-specific mean θ_i along with a known sampling variance, represented by the observed σ_i^2. We also assume θ_i is drawn from a normal distribution where μ is the average effect and τ^2 is the between-study heterogeneity. In contrast to Equation 1, we will estimate μ as a function of the estimated population effect size β_0 and the moderator x multiplied by the coefficient β_1. x is dummy-coded, where 0 indicates that the observed y_i is an effect size estimate of tocilizumab compared to control, while 1 indicates that y_i is an effect size estimate of sarilumab versus control.

For this model, we chose weakly informative prior distributions for β_0 (analogous to μ in Figure 1), β_1 (Figure 2) and τ (Figure 1):

$$\begin{aligned} \beta_0 &\sim \text{Normal}(0, 1.5^2) \\ \beta_1 &\sim \text{Normal}(0, 1^2) \\ \tau &\sim \text{Half-Normal}(0.5) \end{aligned} \tag{6}$$

Although standard deviations for β_0 and β_1 are distinct, we consider both prior distributions weakly informative. As further explained below, β_0 is tocilizumab's effect and β_1 is the difference between sarilumab and tocilizumab. In the odds ratio scale, $\beta_1 = \exp(-1.96)$ represents that sarilumab's effect is 0.14× smaller than tocilizumab's effect, while $\beta_1 = \exp(1.96)$ represents that sarilumab's effect is 7.1× larger than tocilizumab's effect. Thus, a Normal$(0, 1^2)$ distribution only excludes extremely implausible values, which defines a weakly-informative distribution.[8]

In summary, by updating prior beliefs with the observed effects in RCTs, we will estimate the marginal posterior distribution of overall tocilizumab's effect, sarilumab's effect and the between-study heterogeneity:

1. The marginal posterior distribution of β_0 will yield the overall log odds ratio of tocilizumab's effect in comparison to the control group

2. The sum of β_0 and β_1 marginal posterior distributions will yield the overall log odds ratio of sarilumab's effect in comparison to the control group

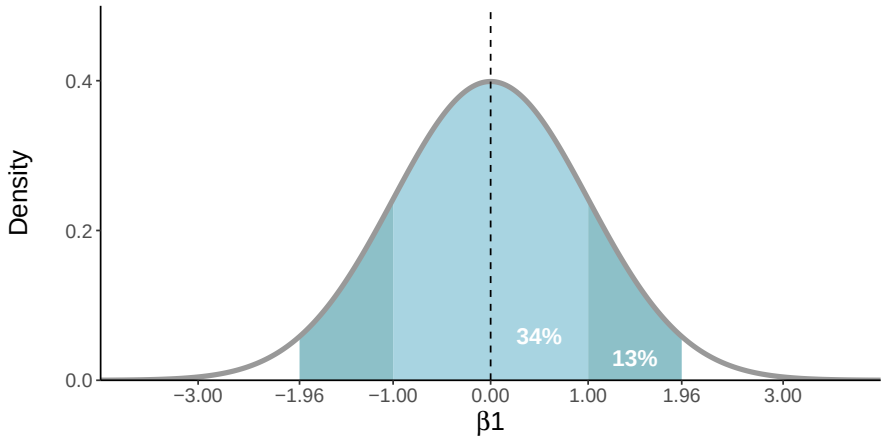

Figure 2: Prior distribution for β_1. The dashed line depict the median value. The Normal(0, 1^2) distribution approximately concentrates 34% of probability density between -1 and 0 (0.37 and 1 in the normal scale, respectively) and 34% of probability density between 0 and 1 (1 and 2.72 in the normal scale, respectively). Moreover, this distribution approximately concentrates 13% between -1.96 and -1 (0.14 and 0.37 in the normal scale, respectively) and 13% between 1 and 1.96 (2.72 and 7.10 in the normal scale, respectively).

3. The marginal posterior distribution of τ^2 will estimate the between-study heterogeneity

However, none of the marginal posterior distributions listed above directly answers our research question: is sarilumab noninferior to tocilizumab in reducing mortality? Thus, to answer this question, we will compare the marginal posterior distributions of sarilumab to tocilizumab.

Because x (Equation 5) is dummy-coded and 0 represents that data on tocilizumab is being modeled, β_0 (intercept) yields the marginal posterior of distribution of tocilizumab's overall effect. When data on sarilumab is being modeled, x equals to 1, indicating that β_1 yields the marginal posterior distribution of the difference between sarilumab's effect to tocilizumab in the log odds ratio scale:

$$\beta_1 = \beta_\Delta = ln(OR_{sarilumab}) - ln(OR_{tocilizumab}) \tag{7}$$

where OR is odds ratio. This equation can be simplified to:

$$exp(\beta_\Delta) = \frac{OR_{sarilumab}}{OR_{tocilizumab}} \tag{8}$$

Thus, we can compare the effect of sarilumab vs. control to tocilizumab vs. control by either estimating the difference of marginal posterior distributions in the log odds ratio scale (Equation 7) or the ratio of marginal posterior distributions in the odds ratio scale (Equation 8).

Instead of assessing if sarilumab is noninferior or not from tocilizumab (dichotomous conclusion), we will leverage the Bayesian framework and estimate the posterior probability of noninferiority. To quantify this posterior probability, we will calculate the area under the curve of $exp(\beta_\Delta)$ (Equation 8) below our estimated noninferiority margin (Equation 4).

Equivalence analysis

- Is sarilumab equivalent to tocilizumab in reducing mortality?

To quantify the equivalence between both drugs, we will calculate the area under the curve of β_{Δ} (Equation 7) within a range of practical equivalence (ROPE).[16] The choice of boundaries for a ROPE are subjective and should be based on clinical judgment.[17] We have decided that a 15% mortality odds reduction or increase (i.e., ratio of odds ratio of 0.85 and $\frac{1}{0.85}$, respectively [Equation 8]) is a reasonable equivalence range for our main analyses.

As previously discussed, sarilumab and tocilizumab are drugs with similar characteristics and are usually considered analogous. Thus, one could assume they are equal and perform pooled analyses. However, this approach ignores possible differences between these drugs and, for example, could overestimate sarilumab's effect by excessively leveraging from tocilizumab's evidence.

Instead of assuming these drugs are entirely exchangeable, we will perform a more thorough analysis. In summary, we will use data from tocilizumab as a prior belief for analyses on sarilumab. Yet, we will also account for the difference between these drugs by increasing the uncertainty around tocilizumab's data.

To this end, we will increasingly augment the uncertainty around log odds ratio estimates on tocilizumab vs. control from each study i. Notably, σ_i^2 is inversely related to the number of events in a study i. Therefore, we will dampen the number of events by multiplying them by a weight factor (W) ranging from 1 to 0.01, where 1 means 100% weight, i.e., no dampening, and 0.01 means only 1% weight, i.e., 99% of dampening. Consequently, we will increase σ_i^2 and thus augment the uncertainty around log odds ratio estimates for tocilizumab vs. control.

We will use the following formula to approximately augment σ_i^2 for each study i, adapted from Spiegelhalter et al.[11]:

$$\sigma_{i[W]}^2 = \frac{1}{Wa_i + \frac{1}{2}} + \frac{1}{Wb_i + \frac{1}{2}} + \frac{1}{Wc_i + \frac{1}{2}} + \frac{1}{Wd_i + \frac{1}{2}} \tag{9}$$

where a_i, b_i, c_i and d_i represent the number of events in a study i following this 2x2 table:

	Tocilizumab	Control
Death	a_i	b_i
No death	c_i	d_i

and W is the weight factor ranging from $W = 1$ to $W = 0.01$. Therefore, we will have sets of $\sigma_{i[W]}^2$ between $\sigma_{i[1]}^2$ and $\sigma_{i[0.01]}^2$.

Next, we will fit a frequentist random-effects meta-analyses similar to Equation 1 using the R package *metafor*,[18] but now we will fit one model per set of $\sigma_{i[W]}^2$ regarding tocilizumab:

$$\begin{aligned} y_i &\sim Normal(\theta_{i[W]}, \sigma_{i[W]}^2) \\ \theta_{i[W]} &\sim Normal(\mu_{tocilizumab[W]}, \tau_{[W]}^2) \end{aligned} \tag{10}$$

where y_i is the observed mean log odds ratio of tocilizumab versus control in study i. We assume these effect sizes are normally distributed around the study-specific mean $\theta_{i[W]}$ along with a known sampling variance, represented by the observed $\sigma_{i[W]}^2$ for each study i according to the weight factor W. We also assume $\theta_{i[W]}$ is drawn from a normal distribution where $\mu_{tocilizumab[W]}$ is the average effect of interest and $\tau_{[W]}^2$ is the between-study heterogeneity.

Therefore, we will fit one model per weight factor W and each will yield the an estimate of tocilizumab's overall effect ($\mu_{tocilizumab[W]}$).

Lastly, because our ultimate goal is to use tocilizumab data as the prior for sarilumab, we will fit a model similar to the one mentioned in the beginning of this section in which sarilumab's effect was estimated (Equation 1), but now the **prior** distribution on $\mu_{sarilumab}$ (sarilumab's mean effect) will

be described by the estimated $\mu_{tocilizumab[W]}$'s mean and standard error in Model 10. We will fit one model per W, each described as:

$$\begin{aligned} y_i &\sim Normal(\theta_i, \sigma_i^2) \\ \theta_i &\sim Normal(\mu_{sarilumab[W]}, \tau^2) \end{aligned} \tag{11}$$

Here are the prior distributions, where $M_{[W]} = mean(\mu_{tocilizumab[W]})$, and $M_{[W]} = SE(\mu_{tocilizumab[W]})$:

$$\begin{aligned} \mu_{sarilumab[W]} &\sim \text{Normal}(M_{[W]}, S_{[W]}) \\ \tau &\sim \text{Half-Normal}(0.5) \end{aligned} \tag{12}$$

In conclusion, we will estimate marginal posterior distributions for $\mu_{sarilumab}$ while borrowing data from tocilizumab in different degrees.

To guide future research, we will also calculate the posterior predictive distributions for θ using different $\mu_{sarilumab[W]}$ while taking into account the between-study heterogeneity provided by τ^2's marginal posterior distribution.

Sensitivity analyses

1. We will refit our models only, including patients that were on the use of corticosteroids.
2. We will re-estimate the posterior probability of sarilumab's noninferiority to tocilizumab using different noninferiority margins.
3. We will re-estimate the posterior probability of equivalence by testing different ROPEs.
4. We will refit our models, excluding unpublished studies while including peer-reviewed articles and preprints.
5. We will refit our models including peer-reviewed articles only.
6. We will refit our models using vague priors for μ, β_0 and β_1.
7. We will refit model 11 using a prior for $\mu_{sarilumab[W]}$ centered around 0 ($= ln[1]$), reflecting a skeptical belief.
8. We will refit our models using different prior distributions for τ, ranging from weakly informative[19] to informative,[20], as visually displayed in Figure 3.

$$\begin{aligned} \tau &\sim \text{Half-Normal}(1) \\ \tau &\sim \text{Half-Cauchy}(0.5) \\ \tau &\sim \text{Half-Normal}(0.2) \\ \tau &\sim \text{Log-Normal}(-1.975, 0.67^2) \end{aligned} \tag{13}$$

Meta-bias

We will assess if publication bias is present by using funnel plot and Egger's test.

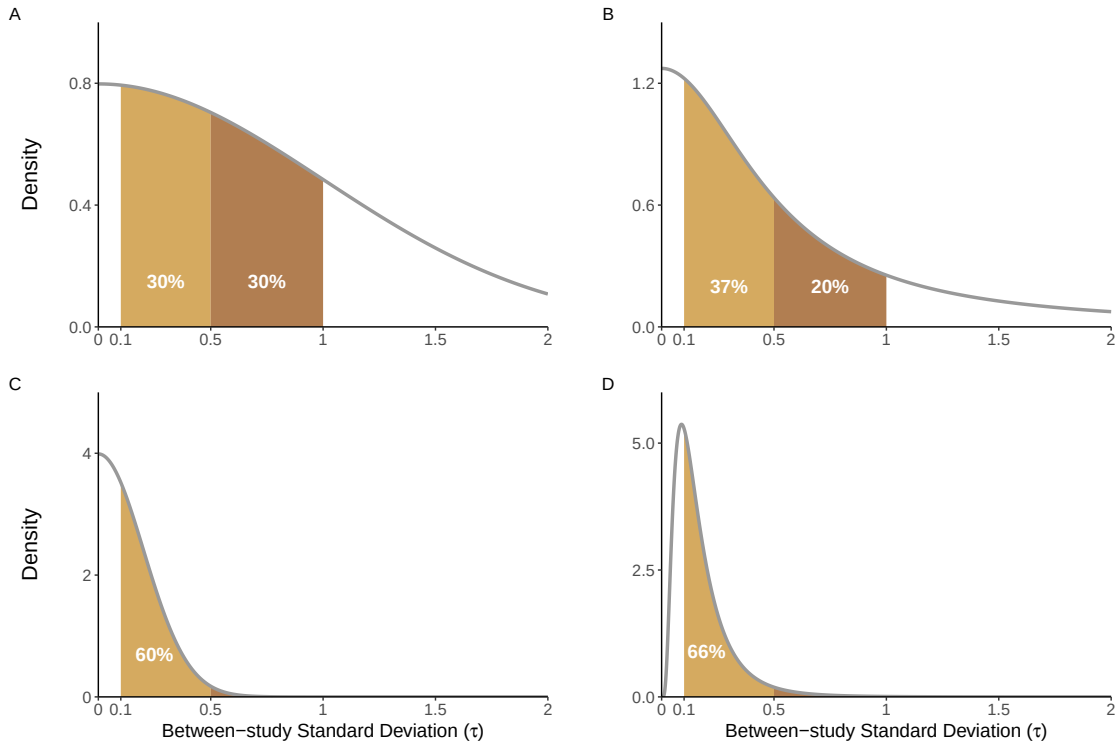

Figure 3: Prior distributions for τ in sensitivity analyses. (A) The Half-Normal(1) distribution concentrates 30% of probability density between 0.1 and 0.5, which Spiegelhalter et al.[11] categorized as a "reasonable heterogeneity" range, and also 30% between 0.5 and 1, categorized as a "fairly high heterogeneity" range. (B) The Half-Cauchy(0.5) distribution concentrates 37% of probability density between 0.1 and 0.5, and 20% between 0.5 and 1. (C) The Half-Normal(0.2) distribution concentrates 60% of probability density between 0.1 and 0.5, and 1% between 0.5 and 1. (D) The Log-Normal$(-1.975, 0.67^2)$ distribution concentrates 66% of probability density between 0.1 and 0.5, and 3% between 0.5 and 1.

Confidence in cumulative evidence

The certainty of evidence of the mortality outcome in each RCT has already been evaluated in the meta-analysis with the GRADE approach.[2] Thus, we will not evaluate again the certainty of evidence.

References

[1] Larissa Shamseer et al. "Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015: elaboration and explanation". In: *BMJ* 349 (2015). DOI: 10.1136/bmj.g7647. URL: https://www.bmj.com/content/349/bmj.g7647.

[2] The WHO Rapid Evidence Appraisal for COVID-19 Therapies (REACT) Working Group. "Association Between Administration of IL-6 Antagonists and Mortality Among Patients Hospitalized for COVID-19: A Meta-analysis". In: *JAMA* (July 2021). DOI: 10.1001/jama.2021.11330. URL: https://doi.org/10.1001/jama.2021.11330.

[3] *Therapeutics and COVID-19: living guideline*. Accessed: 2021-07-13. URL: https://www.who.int/publications/i/item/WHO-2019-nCoV-therapeutics-2021.2.

[4] Zad Rafi and Sander Greenland. "Semantic and Cognitive Tools to Aid Statistical Science: Replace Confidence and Significance by Compatibility and Surprise". In: *BMC Medical Research Methodology* 20.1 (Sept. 2020). DOI: 10.1186/s12874-020-01105-9.

[5] Jonathan A. C. Sterne et al. "RoB 2: A Revised Tool for Assessing Risk of Bias in Randomised Trials". en. In: *BMJ* 366 (Aug. 2019). DOI: 10.1136/bmj.l4898.

[6] Stan Development Team. *RStan: the R interface to Stan*. R package version 2.21.2. 2020. URL: http://mc-stan.org/.

[7] Paul-Christian Bürkner. "brms: An R Package for Bayesian Multilevel Models Using Stan". In: *Journal of Statistical Software* 80.1 (2017), pp. 1–28. DOI: 10.18637/jss.v080.i01.

[8] Andrew Gelman et al. *Bayesian Data Analysis*. 2020. URL: http://www.stat.columbia.edu/~gelman/book/.

[9] Donald R Williams, Philippe Rast, and Paul-Christian Bürkner. "Bayesian Meta-Analysis with Weakly Informative Prior Distributions". In: *PsyArXiv* (2018). DOI: 10.31234/osf.io/7tbrm. URL: psyarxiv.com/7tbrm.

[10] Christian Röver et al. "On Weakly Informative Prior Distributions for the Heterogeneity Parameter in Bayesian Random-Effects Meta-Analysis". In: *Research Synthesis Methods* n/a.n/a (). DOI: 10.1002/jrsm.1475.

[11] D. J. Spiegelhalter, K. R. Abrams, and Jonathan P. Myles. *Bayesian Approaches to Clinical Trials and Health Care Evaluation*. en. Statistics in Practice. Chichester ; Hoboken, NJ: Wiley, 2004.

[12] Michael Tsui et al. "Most Noninferiority Trials Were Not Designed to Preserve Active Comparator Treatment Effects". en. In: *Journal of Clinical Epidemiology* 110 (June 2019). DOI: 10.1016/j.jclinepi.2019.03.003.

[13] "Interleukin-6 Blocking Agents for Treating COVID-19: A Living Systematic Review". en. In: *Cochrane Database of Systematic Reviews* 3 (2021). DOI: 10.1002/14651858.cd013881.

[14] Jane-Chloé Trone et al. "Assessment of Non-Inferiority with Meta-Analysis: Example of Hypofractionated Radiation Therapy in Breast and Prostate Cancer". en. In: *Scientific Reports* 10.1 (Sept. 2020). DOI: 10.1038/s41598-020-72088-2.

[15] Welch VA Higgins JPT. *U.S. Department of Health and Human Services, Food and Drug Administration, Center for Drug Evaluation and Research, Center for Biologics Evaluation and Research. Non-inferiority clinical trials to establish effectiveness: guidance for industry*. 2016. URL: https://www.fda.gov/downloads/Drugs/Guidances/UCM202140.pdf.

[16] John K. Kruschke. *Doing Bayesian Data Analysis: A Tutorial with R, JAGS, and Stan*. Academic Press, 2015. URL: https://sites.google.com/site/doingbayesiandataanalysis/.

[17] Fernando G. Zampieri et al. "Using Bayesian Methods to Augment the Interpretation of Critical Care Trials. An Overview of Theory and Example Reanalysis of the Alveolar Recruitment for Acute Respiratory Distress Syndrome Trial". en. In: *American Journal of Respiratory and Critical Care Medicine* 203.5 (Mar. 2021). DOI: 10.1164/rccm.202006-2381cp.

[18] Wolfgang Viechtbauer. *Conducting meta-analyses in R with the metafor package*. 2010. URL: https://www.jstatsoft.org/v36/i03/.

[19] Christian Röver. "Bayesian Random-Effects Meta-Analysis Using the bayesmeta R Package". In: *Journal of Statistical Software, Articles* 93.6 (2020), pp. 1–51. DOI: 10.18637/jss.v093.i06.

[20] Rebecca M. Turner et al. "Predictive distributions for between-study heterogeneity and simple methods for their application in Bayesian meta-analysis". In: *Statistics in Medicine* 34.6 (2015), pp. 984–998. DOI: https://doi.org/10.1002/sim.6381. URL: https://onlinelibrary.wiley.com/doi/abs/10.1002/sim.6381.