

Appendix

Effect of Tocilizumab, Sarilumab, and Baricitinib on Mortality Among Patients Hospitalized for COVID-19 Treated with Corticosteroids: A Systematic Review and Meta-Analysis

Arthur M. Albuquerque¹, Igor Eckert BS², Lucas Tramujas MD³, Guillaume Butler-Laporte MD^{4,5}, Emily G. McDonald MD MSc^{6,7}, James M. Brophy MD PhD^{5,8}, and Todd C. Lee MD MPH^{4,5,6}

¹School of Medicine, Universidade Federal do Rio de Janeiro, Rio de Janeiro, Brazil

²Department of Nutrition, Universidade Federal de Ciências da Saúde de Porto Alegre, Porto Alegre, Brazil

³HCor Research Institute, São Paulo, Brazil

⁴Division of Infectious Diseases, Department of Medicine, McGill University, Montréal, Canada

⁵Department of Epidemiology, Occupational Health, and Biostatistics, McGill University, Montréal, Canada

⁶Clinical Practice Assessment Unit, Department of Medicine, McGill University, Montréal, Canada

⁷Division of General Internal Medicine, Department of Medicine, McGill University, Montréal, Canada

⁸Division of Cardiology, Department of Medicine, McGill University, Montréal, Canada

Table of Contents

Appendix Methods	3
Search Strategies	3
Data Synthesis and Analysis	6
Meta-analysis	6
Posterior Predictive Distribution	8
Meta-regression	9
Priors	10
Post-hoc Analysis	17
Noninferiority Analysis	18
Appendix References	19
Appendix Results	21
Appendix Table 1: Trial Characteristics (Part 1)	22
Appendix Table 1: Trial Characteristics (Part 2)	23
Appendix Figure 1: Risk of Bias Assessment	24
Appendix Figure 2: Risk of Bias Proportion	25
Appendix Figure 3: Posterior Probabilities (Odds Ratio)	26
Appendix Figure 4: Baricitinib Sensitivity Meta-Analysis	27
Appendix Figure 5: Posterior Between-Study Standard Deviations	28
Appendix Table 2: Posterior Probabilities of Heterogeneity	29
Appendix Table 3: Meta-Analyses Summary Results	30
Appendix Figure 6: Funnel Plots	31
Appendix Table 4: Certainty of Evidence (GRADE)	32
Appendix Figure 7: Sarilumab vs. Tocilizumab Sensitivity Analysis	33
Appendix Table 5: Sarilumab vs. Tocilizumab Sensitivity Analysis	34
Appendix Table 6: Baricitinib vs. Tocilizumab Sensitivity Analysis	35

Appendix Methods

Search Strategies

PubMed:

Query	Terms
#1	(randomized controlled trial[pt] OR controlled clinical trial[pt] OR randomized[tiab] OR placebo[tiab] OR drug therapy[sh] OR randomly[tiab] OR trial[tiab] OR groups[tiab] NOT (animals [mh] NOT humans [mh]))
#2	coronavirus[Title] OR "corona virus" [Title] OR "corona pandemic"[Title] OR coronavirinae[Title] OR coronaviridae[Title] OR betacoronavirus[Title] OR covid19[Title] OR covid[Title] OR nCoV[Title] OR "CoV 2"[Title] OR CoV2[Title] OR sars2[Title] OR sarscov2[Title] OR 2019nCoV[Title] OR "novel CoV"[Title] OR "wuhan virus"[Title] OR coronavirus[Title/Abstract] OR "corona virus" [Title/Abstract] OR "corona pandemic"[Title/Abstract] OR coronavirinae[Title/Abstract] OR coronaviridae[Title/Abstract] OR betacoronavirus[Title/Abstract] OR covid19[Title/Abstract] OR covid[Title/Abstract] OR nCoV[Title/Abstract] OR "CoV 2"[Title/Abstract] OR CoV2[Title/Abstract] OR sars2[Title/Abstract] OR sarscov2[Title/Abstract] OR 2019nCoV[Title/Abstract] OR "novel CoV"[Title/Abstract] OR "wuhan virus"[Title/Abstract] OR "COVID-19" [Appendix Concept] OR "severe acute respiratory syndrome coronavirus 2" [Appendix Concept] OR ((wuhan[Title] OR hubei[Title] OR huanan[Title]) OR (wuhan[Title/Abstract] OR hubei[Title/Abstract] OR huanan[Title/Abstract]) AND ("severe acute respiratory"[Title] OR pneumonia[Title])) OR ("severe acute respiratory"[Title/Abstract] OR pneumonia[Title/Abstract]) AND (outbreak[Title] OR outbreak[Title/Abstract])
#3	sarilumab[Title/Abstract]
#4	baricitinib[Title/Abstract]
#5	tocilizumab[Title/Abstract]

Tocilizumab: #1 AND #2 AND #5

Sarilumab: #1 AND #2 AND #3

Baricitinib: #1 AND #2 AND #4

CENTRAL:

Query	Terms
#1	MeSH descriptor: [COVID-19] explode all trees
#2	sarilumab with PubYear from 2021 to 2021, in Trials
#3	baricitinib
#4	tocilizumab with PubYear from 2021 to 2021, in Trials

Tocilizumab: #1 AND #4

Sarilumab: #1 AND #2

Baricitinib: #1 AND #3

EMBASE:

Query	Terms
#1	('crossover procedure'/exp OR 'crossover procedure') AND [embase]/lim OR (('prospective study'/exp OR 'prospective study') AND [embase]/lim) OR (('follow up'/exp OR 'follow up') AND [embase]/lim) OR (('placebo'/exp OR 'placebo') AND [embase]/lim) OR (('clinical trial'/exp OR 'clinical trial') AND [embase]/lim) OR (('single blind procedure'/exp OR 'single blind procedure') AND [embase]/lim) OR (('double blind procedure'/exp OR 'double blind procedure') AND [embase]/lim) OR (('randomization'/exp OR 'randomization') AND [embase]/lim) OR (('controlled clinical trial'/exp OR 'controlled clinical trial') AND [embase]/lim) OR (('randomized controlled trial'/exp OR 'randomized controlled trial') AND [embase]/lim)
#2	baricitinib':ti,ab
#3	sarilumab':ti,ab
#4	tocilizumab':ti,ab
#5	coronavirus infection' OR 'coronavirus disease 2019' OR 'severe acute respiratory syndrome coronavirus 2'

Tocilizumab: #1 AND #4 AND #5 AND 2021:py

Sarilumab: #1 AND #3 AND #5 AND 2021:py

Baricitinib: #1 AND #2 AND #5

Medrxiv:

Query	Terms
#1	(sarilumab OR baricitinib OR tocilizumab) AND (covid-19 OR coronavirus OR covid) AND random* and abstract or title "random randomized randomly randomised placebo-controlled"

Data Synthesis and Analysis

Meta-analysis

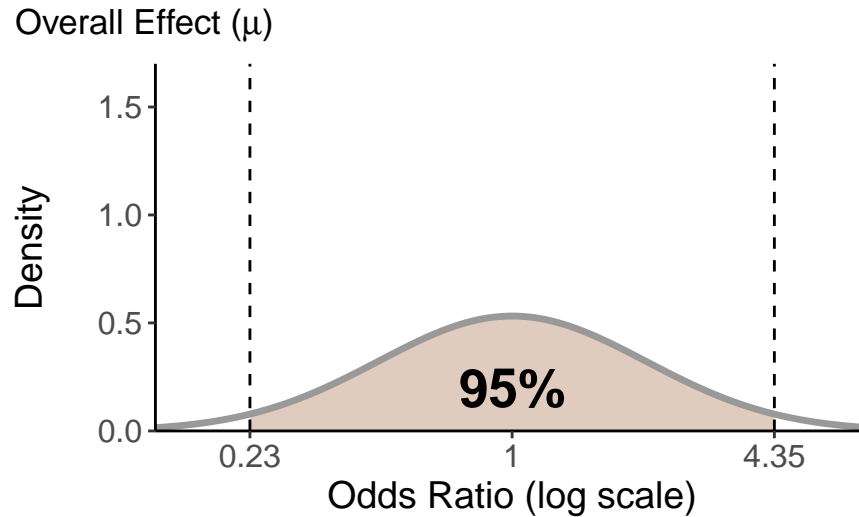
For each set of studies (tocilizumab, baricitinib, or sarilumab vs. control), we fitted a Bayesian random-effects meta-analysis. Each model is defined as:

$$\begin{aligned}y_i &\sim \text{Normal}(\theta_i, \sigma_i^2) \\ \theta_i &\sim \text{Normal}(\mu, \tau^2)\end{aligned}$$

where y_i is the observed mean log odds ratio in study i for either tocilizumab, sarilumab, or baricitinib versus control treatment. 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.

Because these are Bayesian models, we need to specify prior distributions for all parameters. We applied a prior for μ that cover plausible values, assigning limited density to unlikely values, and thereby exerting little influence on the results in all models. A simulation study supported by empirical evidence suggests that a prior with a 95% interval between 0.23 and 4.35 is suitable to this case.(1) In the log scale, such prior is described as:

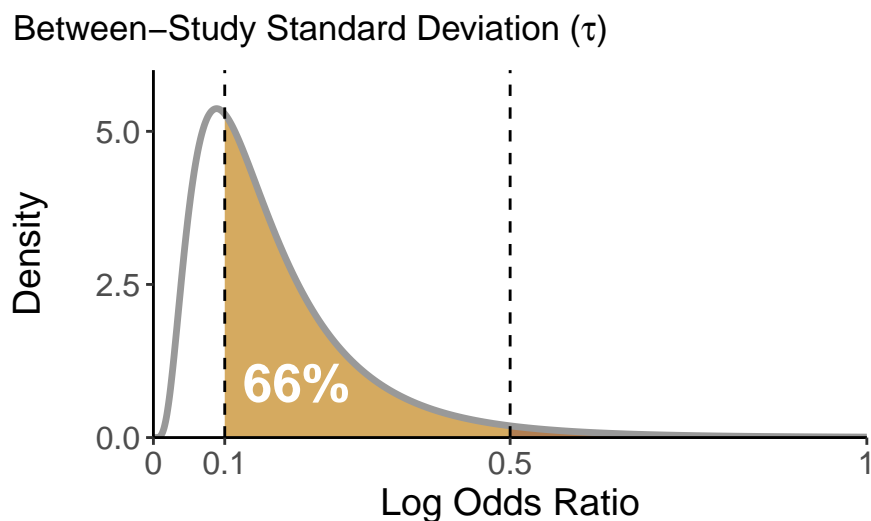
$$\mu \sim \text{Normal}(0, 0.75^2)$$



Regarding the between-study standard deviation parameter (τ), we used an informative prior based on previous evidence. We applied the predictive distribution on all-cause mortality for “placebo / control” vs. “pharmacological” comparisons derived by Turner et al.(2)

Although the definition of small or large between-study heterogeneity is arbitrary, previous work suggests cutoff values (“reasonable” heterogeneity between 0.1 and 0.5, and “fairly high” between 0.5 and 1.0).(3) We added a category for low heterogeneity (between 0 and 0.1):

$$\tau \sim \text{Log-Normal}(-1.975, 0.67^2)$$

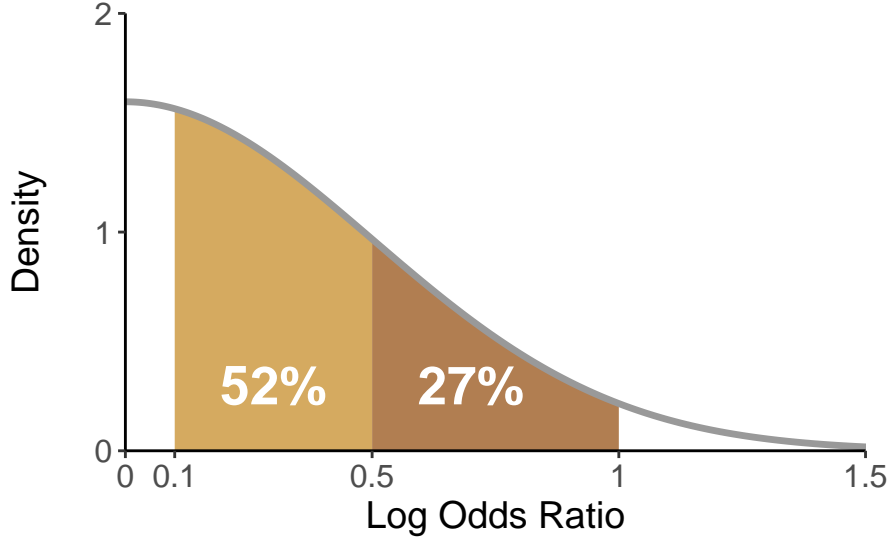


This distribution concentrates 31% of probability density between 0 and 0.1 log odds ratio (‘low heterogeneity’), 66% between 0.1 and 0.5 (‘reasonable’), and only 4% between 0.5 and 1.0 (‘fairly high’).

While maintaining the same prior for μ , we added a post-hoc sensitivity analysis using a less informative prior for τ :(4)

$$\tau \sim \text{Half-Normal}(0, 0.5^2)$$

This distribution yields higher heterogeneity probabilities, while limiting less plausible high levels.



This distribution concentrates 16% of probability density between 0 and 0.1 log odds ratio (‘low heterogeneity’), 52% between 0.1 and 0.5 (‘reasonable’), and 27% between 0.5 and 1.0 (‘fairly high’).

Posterior Predictive Distribution

To further explore the impact of between-study heterogeneity, we evaluated the posterior predictive distribution in each meta-analysis. This distribution allows inference about “... what we would expect to see in a new study population that is exchangeable with the studies included in our meta-analysis.”(5)

$$\theta_{new} \sim \text{Normal}(\mu, \tau^2)$$

The posterior predictive distribution estimates true treatment effects (study populations, θ_{new}) expected in future settings. This distribution is vital to inform probable values for the true treatment association in future settings, valuable for power calculations or prior distribution elicitation in future RCTs.(3) This generates new trial population parameter estimates, independently of sample size and other population characteristics.(6)

Meta-regression

Our primary research question is whether baricitinib and sarilumab are noninferior to tocilizumab. A sensitive approach to answer this question is to separately meta-analyze direct comparisons between tocilizumab vs. baricitinib and tocilizumab vs. sarilumab. However, a preliminary search by our research group showed a limited number of studies that report such comparisons.(7) Thus, we decided to only estimate the noninferiority probability of sarilumab and baricitinib vs. tocilizumab through indirect (across-trials) comparisons. Because of the fragility of such comparisons, it is of great interest to test the robustness of these estimates with multiple prior distributions.

Bayesian network meta-analyses are historically used to estimate direct, indirect, and mixed-effects when more than two treatment comparisons are available.(8) However, this approach does not allow one to assign prior distributions to specific comparisons within the network, such as an indirect effect. To overcome this limitation, we fitted Bayesian random-effect meta-regression models to estimate the indirect differences mentioned above (two separate sets of models: tocilizumab vs. baricitinib; tocilizumab vs. sarilumab).(9) These models are identical to meta-regressions with a dichotomous covariate used to evaluate subgroup differences.(10) This approach allows one to apply a plethora of prior distributions specifically to the indirect comparison between treatments.

Both baricitinib and sarilumab models are identical, and are described as:(9)

$$\begin{aligned} y_i &\sim Normal(\theta_i, \sigma_i^2) \\ \theta_i &\sim Normal(\mu, \tau^2) \\ \mu &= \beta_0 + \beta_{1,T}x_i \end{aligned}$$

where y_i is the observed mean log odds ratio in study i for either tocilizumab, sarilumab, or baricitinib versus control treatment. 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. We will estimate μ as a function of the estimated population effect size β_0 and the moderator x multiplied by the coefficient $\beta_{1,T}$. 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 the **sarilumab** ($T = Sari$) or **baricitinib** ($T = Bari$) compared to control.

In summary, we can estimate the overall log odds ratio for tocilizumab, sarilumab, and baricitinib (depending on which model) as:

- Tocilizumab = β_0
- Sarilumab = $\beta_0 + \beta_{1,Sari}$
- Baricitinib = $\beta_0 + \beta_{1,Bari}$

$\beta_{1,T}$ is the parameter that estimates the indirect difference between tocilizumab vs. sarilumab or tocilizumab vs. baricitinib. When exponentiated, $\beta_{1,T}$ yields the ratio of odds ratios, as demonstrated below. We will reference sarilumab and baricitinib as “comparator” hereafter.

$$\beta_{1,T} = \ln(OR_{comparator}) - \ln(OR_{tocilizumab}) = \ln\left(\frac{OR_{comparator}}{OR_{tocilizumab}}\right)$$

$$\exp(\beta_{1,T}) = \frac{OR_{comparator}}{OR_{tocilizumab}}$$

Priors

β_0 represents the log odds ratio for tocilizumab vs. control ($\ln[OR_{tocilizumab}]$). Here, we applied the same prior mentioned before (regarding μ) for β_0 :(1)

$$\beta_0 \sim \text{Normal}(0, 0.75^2)$$

Regarding the between-study standard deviation parameter (τ), we also used the same informative prior mentioned before:(2)

$$\tau \sim \text{Log-Normal}(-1.975, 0.67^2)$$

In addition to the parameters mentioned above, we should also assign a prior distribution for the indirect comparison parameter ($\beta_{1,T}$). Inspired by the guidelines provided by Zampieri et al.,(11) we fitted different models to represent distinct beliefs. Before discussing these prior distributions, we will first elucidate the underlying rationale of this parameter.

A log odds ratio lower than 0 represents lower mortality in the intervention arm (tocilizumab, sarilumab, or baricitinib) in comparison to control. Moreover, the indirect comparison parameter represents the absolute difference between comparator (sarilumab or baricitinib) and tocilizumab studies in the log scale. Therefore, when the mortality reduction due to intervention is greater in comparator studies (lower $\ln[OR_{comparator}]$ than $\ln[OR_{tocilizumab}]$), $\beta_{1,T}$ is negative in the log scale:

$$\downarrow \ln(OR_{comparator}) - \uparrow \ln(OR_{tocilizumab}) \rightarrow -\beta_{1,T}$$

Here is the same rationale in the linear scale to further facilitate $\beta_{1,T}$'s prior elicitation:

$$\exp(\beta_{1,T}) = ROR = \frac{OR_{comparator}}{OR_{tocilizumab}}$$

where ROR is the ratio of odds ratios, i.e. the interaction parameter exponentiated ($\exp[\beta_{1,T}]$). Thus, a ROR lower than 1 also indicates greater mortality reduction due to intervention in comparator studies:

$$\frac{\downarrow OR_{comparator}}{\uparrow OR_{tocilizumab}} \rightarrow ROR < 1$$

We separately applied 4 normally distributed prior distributions for $\beta_{1,T}$ in each model (sarilumab and baricitinib’s models), which can be described based on the “prior belief” and “belief strength.” While the former reflects the distribution’s mean, the latter reflects the distribution’s standard deviation.

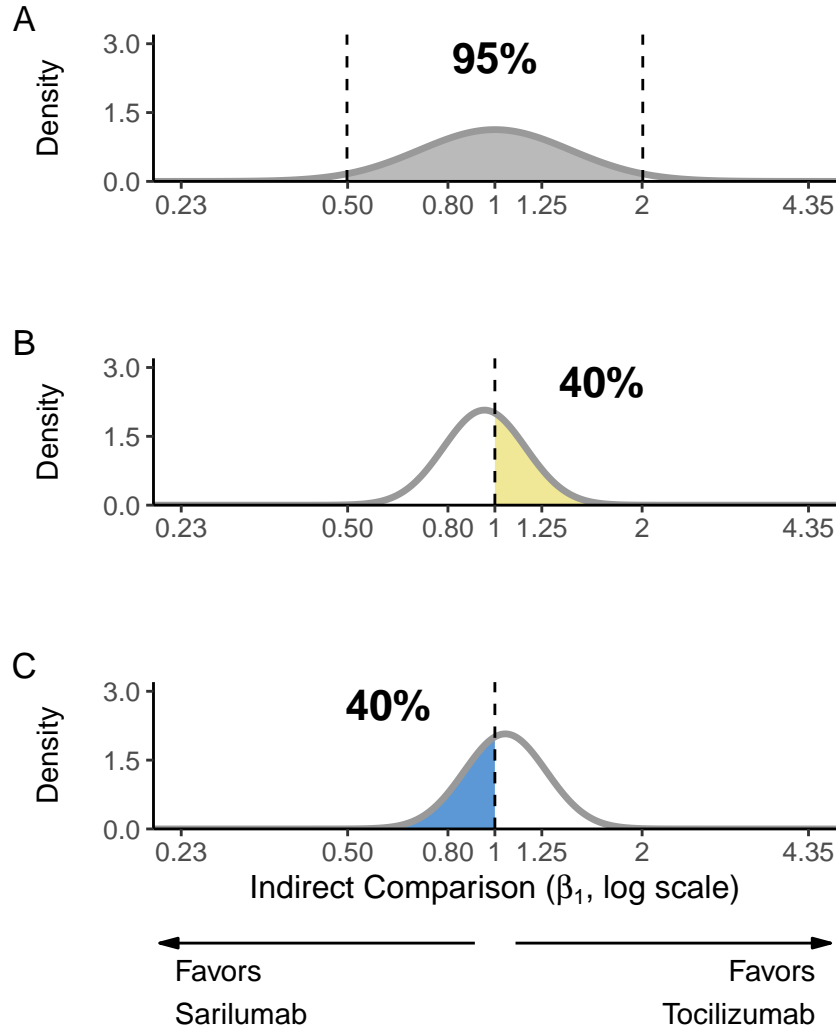
In the model comparing **sarilumab** to tocilizumab, the beliefs are:

- “Skeptical,” which expects no difference between sarilumab and tocilizumab studies (mean centered at 1) with 95% probability between 0.5 and 2 (linear scale, ratio of odds ratios). These values represent a weak belief strength,(11) based on the little information currently available on this topic. Normal(0, 0.354²). Results from this prior take part in the main meta-regression figure in this article.
- “Optimistic for Sarilumab,” which expects a greater mortality reduction due to sarilumab use in comparison to tocilizumab. The mean is centered at 0.952 (linear scale, ratio of odds ratios), which was based on the single study that directly compared sarilumab to tocilizumab (REMAP-CAP).(7) In this study, they assessed “hospital survival” and an odds ratio greater than 1 indicated greater in-hospital mortality reduction due to sarilumab use in comparison to tocilizumab. As shown in their Figure S3, the median odds ratio of sarilumab vs. tocilizumab was 1.05, indicating a greater odds of hospital survival due to sarilumab. However, in the present study, we will assess mortality (not hospital survival), and an odds ratio lower (not greater) than 1 represents greater mortality reduction due to sarilumab use. Thus, to guide this belief based on the study mentioned above, we calculated the reciprocal of 1.05 (1/1.05), which yields 0.952. Normal(−0.049, 0.193²). Results from this prior take part in the Appendix meta-regression figure in this article.

Further, we considered that there is 40% of probability that the ratio of odds ratios is greater than 1. We chose 40% because the single study that directly compared sarilumab to tocilizumab found 34% of probability density for tocilizumab’s superiority over sarilumab.(7) However, this was the first and only study and therefore likely to be an over-estimate and not to include potential between study variability. We then elected to see a 40% of probability for tocilizumab’s superiority over sarilumab in this prior as being a “realistic optimistic for Sarilumab” prior.(7) Although not identical to the 30% suggested by Zampieri et al. – which we consider too informative in this case, given that it would be more informative than the study mentioned above – we believe this prior represents weak belief strength.(11)

- “Optimistic for Tocilizumab,” which expects a greater mortality reduction due to tocilizumab use in comparison to sarilumab. The mean is centered at 1.05, the reciprocal of 0.952. Similarly, this prior belief will also have a weak strength, represented by 40% of probability that the ratio of odds ratios is lower than 1.(11) $\text{Normal}(0.049, 0.193^2)$ Results from this prior take part in the Appendix meta-regression figure in this article.
- “Vague,” which represents no prior belief with no strength. $\text{Normal}(0, 4^2)$. Results from this prior take part in the main meta-regression figure in this article.

The first three distributions are shown below:



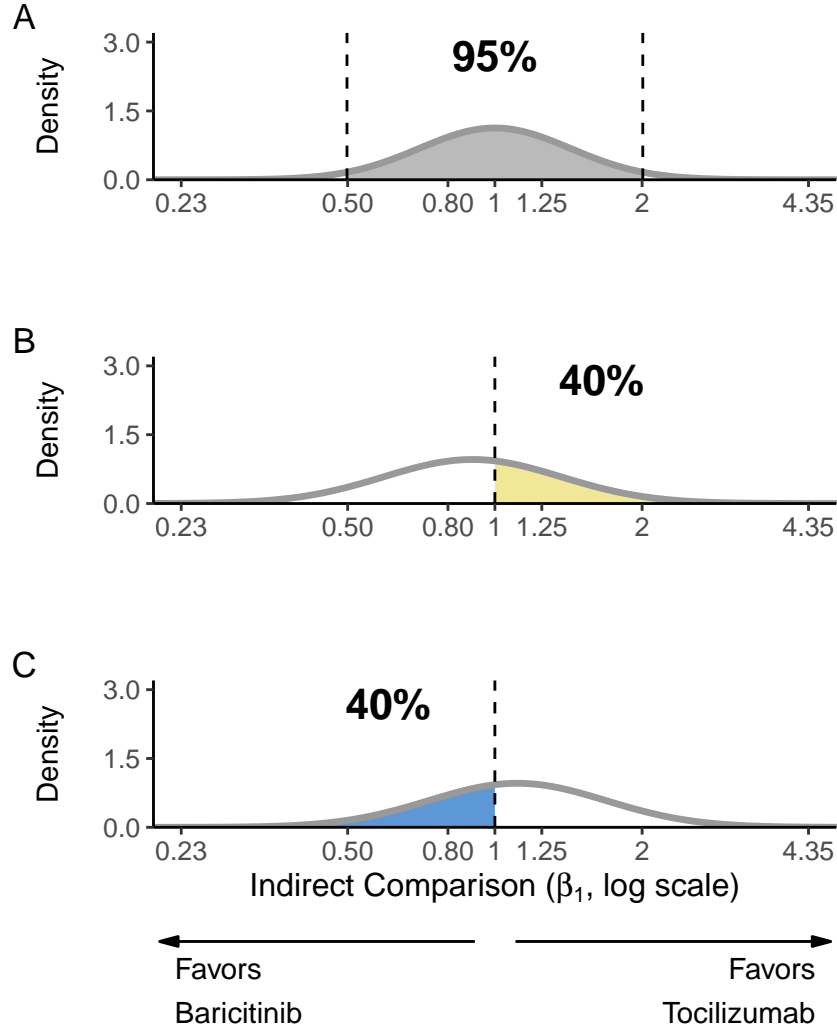
Prior distributions for β_1, Sari (ratio of odds ratios) in models comparing sarilumab to tocilizumab. Panel A: ‘Skeptical’ prior. The $\text{Normal}(0, 0.354^2)$ distribution approximately concentrates 95% of probability density between 0.5 and 2.0 in the linear scale (-0.69 and 0.69 in the log scale, respectively). Panel B: ‘Optimistic for Sarilumab’ prior. The $\text{Normal}(-0.049, 0.193^2)$ distribution approximately concentrates 40% of probability

density above 1. Panel C: ‘Optimistic for Tocilizumab’ prior. The $\text{Normal}(0.049, 0.193^2)$ distribution approximately concentrates 40% of probability density below 1.

In the model comparing **baricitinib** to tocilizumab, the beliefs are:

- “Skeptical,” which expects no difference between baricitinib and tocilizumab studies (mean centered at 1) with 95% probability between 0.5 and 2 (linear scale, ratio of odds ratios). These values represent a weak belief strength,(11) based on the little information currently available on this topic. $\text{Normal}(0, 0.354^2)$. Results from this prior take part in the main meta-regression figure in this article.
- “Optimistic for Baricitinib,” which expects a greater mortality reduction due to baricitinib use in comparison to tocilizumab. The mean is centered at 0.9 (linear scale, ratio of odds ratios), which was arbitrarily chosen based on clinical judgment that a 10% relative difference is clinically important. Because there was no information previously available, we considered that there is 40% of probability that the ratio of odds ratios is greater than 1, representing a weak belief strength.(11) $\text{Normal}(-0.105, 0.416^2)$. Results from this prior take part in the main meta-regression figure in this article.
- “Optimistic for Tocilizumab,” which expects a greater mortality reduction due to tocilizumab use in comparison to sarilumab. The mean is centered at 1.11, the reciprocal of 0.9. Similarly, this prior belief will also have a weak strength, represented by 40% of probability that the ratio of odds ratios is lower than 1.(11) $\text{Normal}(0.105, 0.416^2)$. Results from this prior take part in the main meta-regression figure in this article.
- “Vague,” which represents no prior belief with no strength. $\text{Normal}(0, 4^2)$. Results from this prior take part in the main meta-regression figure in this article.

The first three distributions are shown below:



Prior distributions for β_1 , Bari (ratio of odds ratios) in models comparing baricitinib to tocilizumab. Panel A: 'Skeptical' prior. The Normal($0, 0.354^2$) distribution approximately concentrates 95% of probability density between 0.5 and 2.0 in the linear scale (-0.69 and 0.69 in the log scale, respectively). Panel B: 'Optimistic for Baricitinib' prior. The Normal($-0.105, 0.41^2$) distribution approximately concentrates 40% of probability density above 1. Panel C: 'Optimistic for Tocilizumab' prior. The Normal($0.105, 0.41^2$) distribution approximately concentrates 40% of probability density below 1.

In summary, both sarilumab and baricitinib models can be fully described as:

$$\begin{aligned}y_i &\sim \text{Normal}(\theta_i, \sigma_i^2) \\ \theta_i &\sim \text{Normal}(\mu, \tau^2) \\ \mu &= \beta_0 + \beta_{1,T}x_i\end{aligned}$$

while the priors in sarilumab's models are:

$$\begin{aligned}\beta_0 &\sim \text{Normal}(0, 0.75^2) \\ \tau &\sim \text{Log-Normal}(-1.975, 0.67^2) \\ \beta_{1,\text{Sari}[\text{Skeptical}]} &\sim \text{Normal}(0, 0.354^2) \\ \beta_{1,\text{Sari}[\text{Sarilumab}]} &\sim \text{Normal}(-0.049, 0.193^2) \\ \beta_{1,\text{Sari}[\text{Tocilizumab}]} &\sim \text{Normal}(0.049, 0.193^2) \\ \beta_{1,\text{Sari}[\text{Vague}]} &\sim \text{Normal}(0, 4^2)\end{aligned}$$

and the priors in baricitinib's models are:

$$\begin{aligned}\beta_0 &\sim \text{Normal}(0, 0.75^2) \\ \tau &\sim \text{Log-Normal}(-1.975, 0.67^2) \\ \beta_{1,\text{Bari}[\text{Skeptical}]} &\sim \text{Normal}(0, 0.354^2) \\ \beta_{1,\text{Bari}[\text{Baricitinib}]} &\sim \text{Normal}(-0.105, 0.416^2) \\ \beta_{1,\text{Bari}[\text{Tocilizumab}]} &\sim \text{Normal}(0.105, 0.416^2) \\ \beta_{1,\text{Bari}[\text{Vague}]} &\sim \text{Normal}(0, 4^2)\end{aligned}$$

where all parameters are in the log scale.

One could argue that baricitinib's priors are stronger than sarilumab's given differences on the mean values ($|0.105| > |0.049|$). We note that, although the mean values are different because we only found prior evidence for the sarilumab vs. tocilizumab comparison, these prior discrepancies are likely to be of little relevance. As depicted in the figures above, the baricitinib's priors have a greater mean but lower density than sarilumab's priors. Therefore, we do not expect these priors' impact on each marginal posterior distribution to be relevantly different.

Post-hoc Analysis

Regarding the sarilumab vs. tocilizumab meta-regression model, we also performed a post-hoc analysis with a different set of priors.

We fitted the exact same model, but with different “Optimistic for Sarilumab” and “Optimistic for Tocilizumab” priors (hereafter, “Optimistic for Sarilumab (REMAP-CAP)” and “Optimistic for Tocilizumab (inverse REMAP-CAP)”):

$$y_i \sim \text{Normal}(\theta_i, \sigma_i^2) \quad (\text{Likelihood})$$

$$\theta_i \sim \text{Normal}(\mu, \tau^2)$$

$$\mu = \beta_0 + \beta_{1,\text{Sari}}x_i$$

$$\beta_0 \sim \text{Normal}(0, 0.75^2) \quad (\text{Priors})$$

$$\tau \sim \text{Log-Normal}(-1.975, 0.67^2)$$

$$\beta_{1,\text{Sari}[\text{Sarilumab}]} \sim \text{Normal}(-0.049, 0.118^2)$$

$$\beta_{1,\text{Sari}[\text{Tocilizumab}]} \sim \text{Normal}(0.049, 0.118^2)$$

These priors were constructed to reflect the exact same strength of the single study that directly compared sarilumab to tocilizumab.(7)

The new standard deviation of 0.118 was obtained through the formula:(12)

$$\sigma = \frac{\ln(\text{UCL}) - \ln(\text{LCL})}{3.92}$$

where UCL is 1.35, and LCL is 0.85, as depicted in the Figure S3 of the RCT cited above (REMAP-CAP). In comparison to the 40% probability density above/below 1.0 yielded by the main priors in the section before, these sensitivity priors yield 34%, highlighting its greater strength.

Results from these priors take part in the main meta-regression figure in this article.

Noninferiority Analysis

Our primary goal is to assess whether sarilumab and baricitinib are noninferior to tocilizumab. Thus, we calculated the posterior probabilities of noninferiority based on the marginal posterior distributions of $\beta_{1,T}$ in all models and scenarios described above.

One major criteria in noninferiority analyses is that superiority of active comparator — in this case, tocilizumab — in comparison to control has been shown.(13) The WHO meta-analysis presented results in favor of tocilizumab’s superiority.(14) Thus, we will use the following formula to estimate our noninferiority margin:(15)

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

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. Regarding θ , we extracted the mean odds ratio regarding tocilizumab’s effect in the subgroup of patients using corticosteroids (page 14 in their Supplement 2 (14)). We chose this subgroup because it represents our population of interest. Moreover, 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.(13,16) Thus, our main noninferiority margin (in the linear scale) will be:

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

Appendix References

1. Pedroza C, Han W, Thanh Truong VT, Green C, Tyson JE. Performance of informative priors skeptical of large treatment effects in clinical trials: A simulation study. *Stat Methods Med Res* [Internet]. 2018 Jan 1 [cited 2021 Sep 11];27(1):79–96. Available from: <https://doi.org/10.1177/0962280215620828>
2. Turner RM, Jackson D, Wei Y, Thompson SG, Higgins JPT. Predictive distributions for between-study heterogeneity and simple methods for their application in Bayesian meta-analysis. *Statistics in Medicine* [Internet]. 2015 [cited 2021 May 8];34(6):984–98. Available from: <https://onlinelibrary.wiley.com/doi/abs/10.1002/sim.6381>
3. Spiegelhalter DJ, Abrams KR, Myles JP. Bayesian approaches to clinical trials and health care evaluation. Chichester ; Hoboken, NJ: Wiley; 2004. 391 p. (Statistics in practice).
4. Röver C, Bender R, Dias S, Schmid CH, Schmidli H, Sturtz S, et al. On weakly informative prior distributions for the heterogeneity parameter in Bayesian random-effects meta-analysis. *Research Synthesis Methods* [Internet]. 2021 [cited 2021 Sep 2];12(4):448–74. Available from: <https://onlinelibrary.wiley.com/doi/abs/10.1002/jrsm.1475>
5. Welton NJ, Jones HE, Dias S. Bayesian Methods for Meta-Analysis. In: *Bayesian Methods in Pharmaceutical Research*. Chapman and Hall/CRC; 2020.
6. Higgins JPT, Thompson SG, Spiegelhalter DJ. A re-evaluation of random-effects meta-analysis. *Journal of the Royal Statistical Society: Series A (Statistics in Society)* [Internet]. 2009 [cited 2021 Jun 28];172(1):137–59. Available from: <https://rss.onlinelibrary.wiley.com/doi/abs/10.1111/j.1467-985X.2008.00552.x>
7. Effectiveness of Tocilizumab, Sarilumab, and Anakinra for critically ill patients with COVID-19 The REMAP-CAP COVID-19 Immune Modulation Therapy Domain Randomized Clinical Trial. :21.
8. Dias S, editor. Network meta-analysis for decision making. Hoboken, NJ: Wiley; 2018. 1 p. (Wiley Series in Statistics in Practice).
9. Pitchforth JO, Mengersen KL. Bayesian Meta-Analysis. In: *Case Studies in Bayesian Statistical Modelling and Analysis* [Internet]. John Wiley & Sons, Ltd; 2012 [cited 2022 Mar 5]. p. 118–40. Available from: <https://onlinelibrary.wiley.com/doi/abs/10.1002/9781118394472.ch7>
10. Thompson SG, Higgins JPT. How should meta-regression analyses be undertaken and interpreted? *Statist Med* [Internet]. 2002 Jun 15 [cited 2021 Sep 7];21(11):1559–73. Available from: <https://onlinelibrary.wiley.com/doi/10.1002/sim.1187>
11. Zampieri FG, Casey JD, Shankar-Hari M, Harrell FE, Harhay MO. 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. *Am J Respir Crit Care Med* [Internet]. 2021 Mar 1 [cited 2021 Mar 1];203(5):543–52. Available from: <https://www.atsjournals.org/doi/10.1164/rccm.202006-2381CP>

12. Higgins J, Thomas J, Chandler J, Cumpston M, Li T, Page M, et al., editors. Cochrane Handbook for Systematic Reviews of Interventions version 6.2 (updated February 2021) [Internet]. Cochrane; 2021 [cited 2022 Jan 15]. Available from: <https://training.cochrane.org/handbook>
13. Tsui M, Rehal S, Jairath V, Kahan BC. Most noninferiority trials were not designed to preserve active comparator treatment effects. *Journal of Clinical Epidemiology* [Internet]. 2019 Jun 1 [cited 2022 Mar 5];110:82–9. Available from: [https://www.jclinepi.com/article/S0895-4356\(18\)30599-7/fulltext](https://www.jclinepi.com/article/S0895-4356(18)30599-7/fulltext)
14. 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. *JAMA* [Internet]. 2021 Aug 10 [cited 2021 Sep 2];326(6):499–518. Available from: <https://doi.org/10.1001/jama.2021.11330>
15. Trone JC, Ollier E, Chapelle C, Mismetti P, Cucherat M, Magné N, et al. Assessment of non-inferiority with meta-analysis: example of hypofractionated radiation therapy in breast and prostate cancer. *Sci Rep* [Internet]. 2020 Sep 22 [cited 2022 Mar 5];10(1):15415. Available from: <https://www.nature.com/articles/s41598-020-72088-2>
16. U.S. Department of Health and Human Services Food and Drug Administration. Non-Inferiority Clinical Trials to Establish Effectiveness Guidance for Industry [Internet]. 2016. Available from: <https://www.fda.gov/media/78504/download>

Appendix Results

Appendix Table 1: Trial Characteristics (Part 1)

Study	Article Type	Number of Sites	Follow-up Length ¹	Dose
Tocilizumab				
RECOVERY Toci	Peer-reviewed	131	28	Based on weight - 800 mg (maximum dose)
REMAP-CAP	Peer-reviewed	113	28	8 mg/kg - 800 mg (maximum dose)
REMDACTA	Peer-reviewed	53	28	8 mg/kg - 800 mg (maximum dose)
PreToViD	Non-peer-reviewed	NA	28	8 mg/kg - 800 mg (maximum dose)
EMPACTA	Peer-reviewed	63	28	8 mg/kg - 800 mg (maximum dose)
COVACTA	Peer-reviewed	62	28	8 mg/kg - 800 mg (maximum dose)
TOCIBRAS	Peer-reviewed	9	28	8 mg/kg - 800 mg (maximum dose)
HMO-020-0224	Non-peer-reviewed	NA	28	8 mg/kg - 800 mg (maximum dose)
COV-AID	Peer-reviewed	16	28	8 mg/kg - 800 mg (maximum dose)
ImmCOVA	Non-peer-reviewed	NA	28	8 mg/kg - 800 mg (maximum dose)
CORIMUNO-TOCI-ICU	Peer-reviewed	9	28	8 mg/kg
CORIMUNO-TOCI-1	Peer-reviewed	9	28	8 mg/kg
ARCHITECTS	Non-peer-reviewed	NA	28	8 mg/kg - 800 mg (maximum dose)
COVIDOSE2-SS-A	Non-peer-reviewed	1	28	40 mg or 120 mg
BACC-Bay	Peer-reviewed	7	28	8 mg/kg - 800 mg (maximum dose)
CORIMUNO-TOCI-DEX	Peer-reviewed	25	28	8 mg/kg
Baricitinib				
ACTT-2	Peer-reviewed	67	29	4 mg or 2 mg
COV-BARRIER	Peer-reviewed	101	28	4 mg or 2 mg
COV-BARRIER 2	Peer-reviewed	18	28	4 mg or 2 mg
RECOVERY Bari	Preprint	159	28	4 mg or lower
Sarilumab				
CORIMUNO-SARI-1	Peer-reviewed	6	28	400 mg
REGENERON-P2	Preprint	65	28	200 or 400 mg
REGENERON-P3	Preprint	65	28	200 or 400 mg
REMAP-CAP	Peer-reviewed	113	28	400 mg
SARCOVID	Non-peer-reviewed	1	28	400 mg
SARICOR	Peer-reviewed	NA	28	200 or 400 mg
SARTRE	Peer-reviewed	8	28	Based on weight, < 75 kg: 200 mg; >= 400 mg

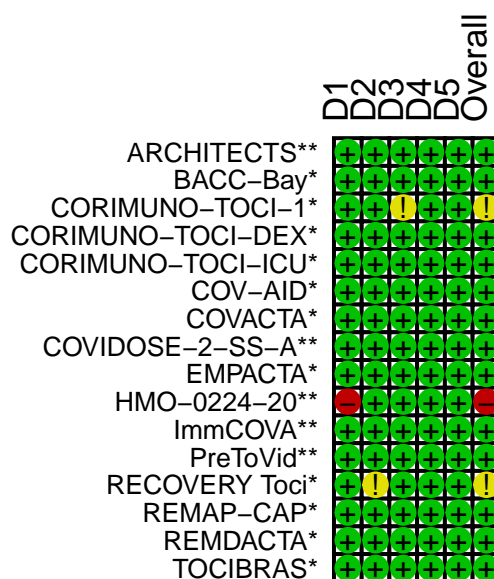
¹Days

Appendix Table 1: Trial Characteristics (Part 2)

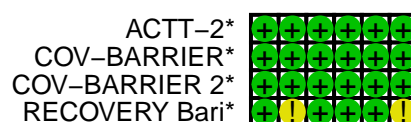
Study	Drug Regimen	Control Type
Tocilizumab		
RECOVERY Toci	Single dose with second dose if the patient had not improved in 12-24h	Usual care
REMAP-CAP	1 or 2 doses at clinical discretion	Usual care
REMDACTA	1 or 2 doses	Placebo + Usual care + Remdesivir
PreToViD	Single dose	Usual care
EMPACTA	Single dose with second dose allowed if symptoms worsened	Placebo + Usual care
COVACTA	1 or 2 doses	Placebo + Usual care
TOCIBRAS	Single dose	Usual care
HMO-020-0224	Single dose	Placebo + Usual care
COV-AID	Single dose	Usual care
ImmCOVA	Single dose	Usual care
CORIMUNO-TOCI-ICU	Single dose with second dose of 400 mg if decrease of oxygen requirement < 50 %	Usual care
CORIMUNO-TOCI-1	Single dose with second dose of 400 mg if decrease of oxygen requirement < 50 %	Usual care
ARCHITECTS	Single dose with 1 additional dose allowed if symptoms worsened	Placebo + Usual care
COVIDOSE2-SS-A	Single dose	Usual care
BACC-Bay	Single dose	Placebo + Usual care
CORIMUNO-TOCI-DEX	Single dose with second dose of 400 mg if decrease of oxygen requirement < 50 %	Usual care
Baricitinib		
ACTT-2	1 dose/day for 14 days or until discharge	Placebo + Usual care + Remdesivir
COV-BARRIER	1 dose/day for 10 days or until discharge	Placebo + Usual care
COV-BARRIER 2	1 dose/day for 10 days or until discharge	Placebo + Usual care
RECOVERY Bari	1 dose/day for 10 days or until discharge	Usual care
Sarilumab		
CORIMUNO-SARI-1	Single dose with second dose of 400 mg if decrease of oxygen requirement < 50 %	Usual care
REGENERON-P2	Multiples doses	Placebo + Usual Care
REGENERON-P3	Multiples doses	Placebo + Usual Care
REMAP-CAP	Single dose	Usual care
SARCOVID	Single dose of 200 mg twice	Usual care
SARICOR	Single dose	Usual care
SARTRE	Single dose	Usual care

Appendix Figure 1: Risk of Bias Assessment

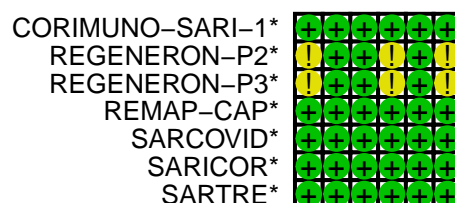
Tocilizumab



Baricitinib



Sarilumab



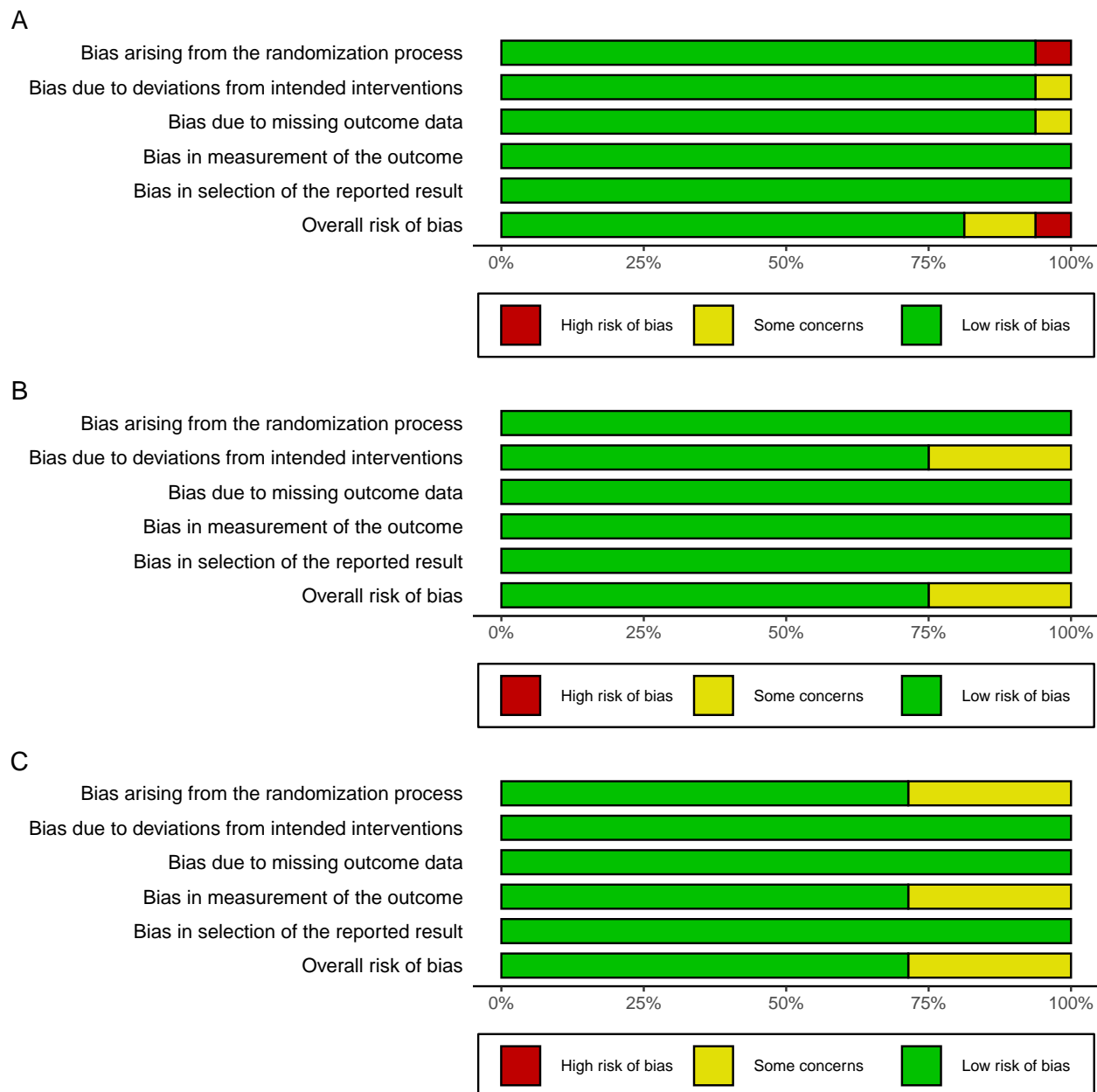
Risk of bias for domain-level as well as overall judgement for each individual study.

D1: bias arising from randomisation process; D2: bias arising from deviations from intended interventions; D3: bias due to missing outcome data; D4: bias in measurement of the outcome; D5: bias in selection of the reported result.

“+” depicts “Low” risk, while “!” and “-” indicate “Some concerns” and “High,” respectively.

Studies with * indicate that the risk of bias assessment was performed by our research group. In contrast, ** indicates that the assessment was performed by the WHO React Working Group, as explained in our Methods section.

Appendix Figure 2: Risk of Bias Proportion

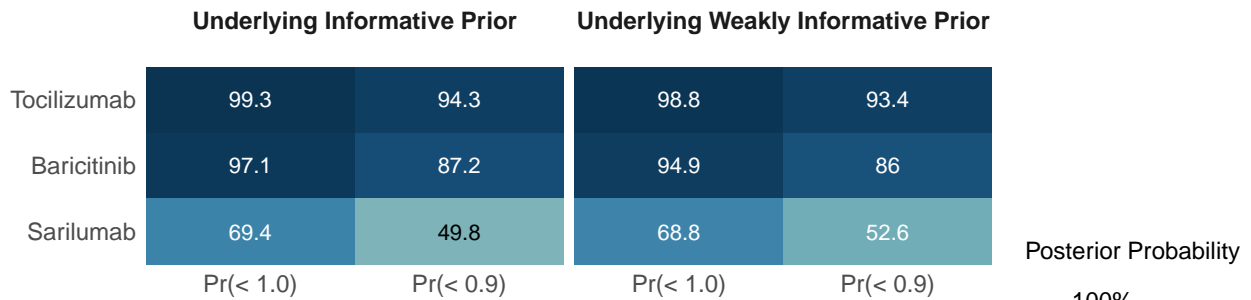


Distributions of proportions of risk of bias within each domain. Panel A: Tocilizumab studies; Panel B: Baricitinib; Panel C: Sarilumab.

Appendix Figure 3: Posterior Probabilities (Odds Ratio)

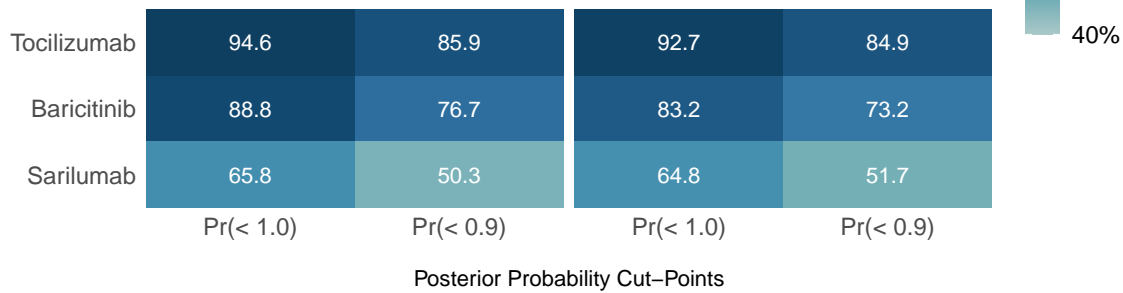
A

Average



B

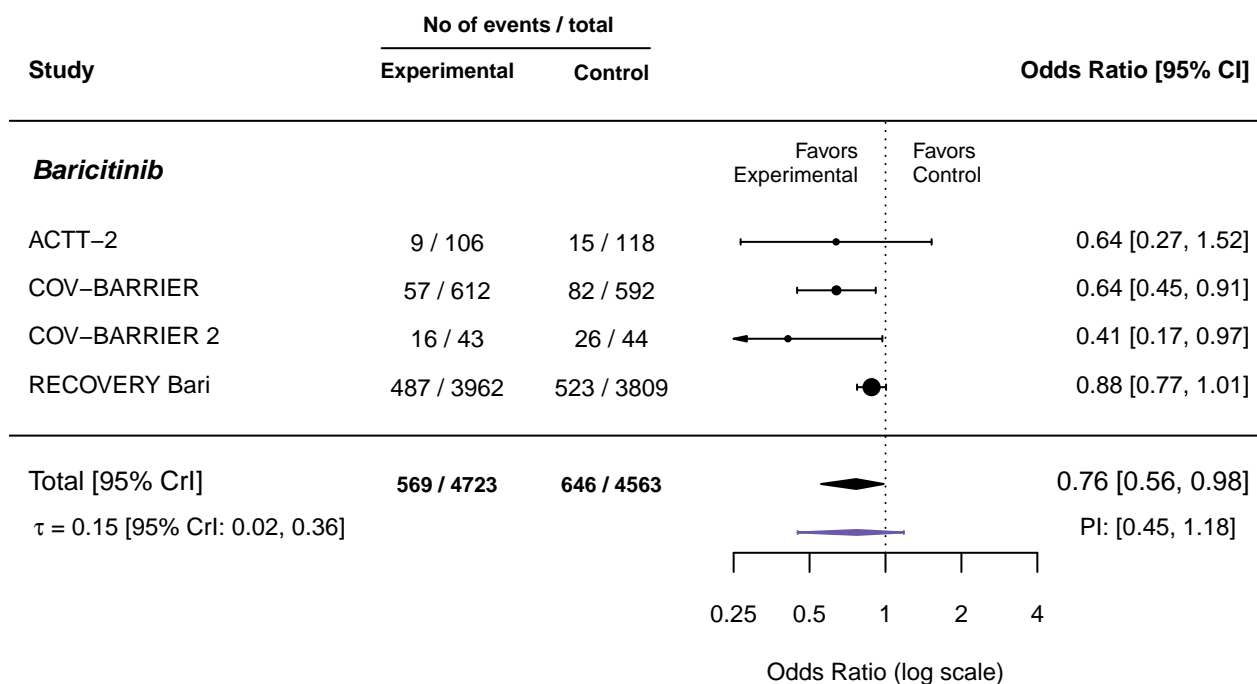
Predictive



Heatmaps with posterior probabilities of tocilizumab, baricitinib, and sarilumab meta-analyses' average (Panel A) and predictive effect (Panel B). “ $Pr(< 1.0)$ ” depicts probabilities of any benefit (< 1.0 odds ratio [OR]), and “ $Pr(< 0.9)$ ” depicts probabilities of clinically meaningful benefit (< 0.9 OR). The exact probability is shown in each cell, while darker colors correspond to higher probabilities.

Heatmaps on the left regard models assuming a more informative prior (Log-Normal[$-1.975, 0.67^2$]) for the between-study standard deviation, while heatmaps on the right regard models with weakly informative prior (Half-Normal[0.5^2]). All models assume a $Normal(0, 0.75^2)$ prior distribution for average effect (μ).

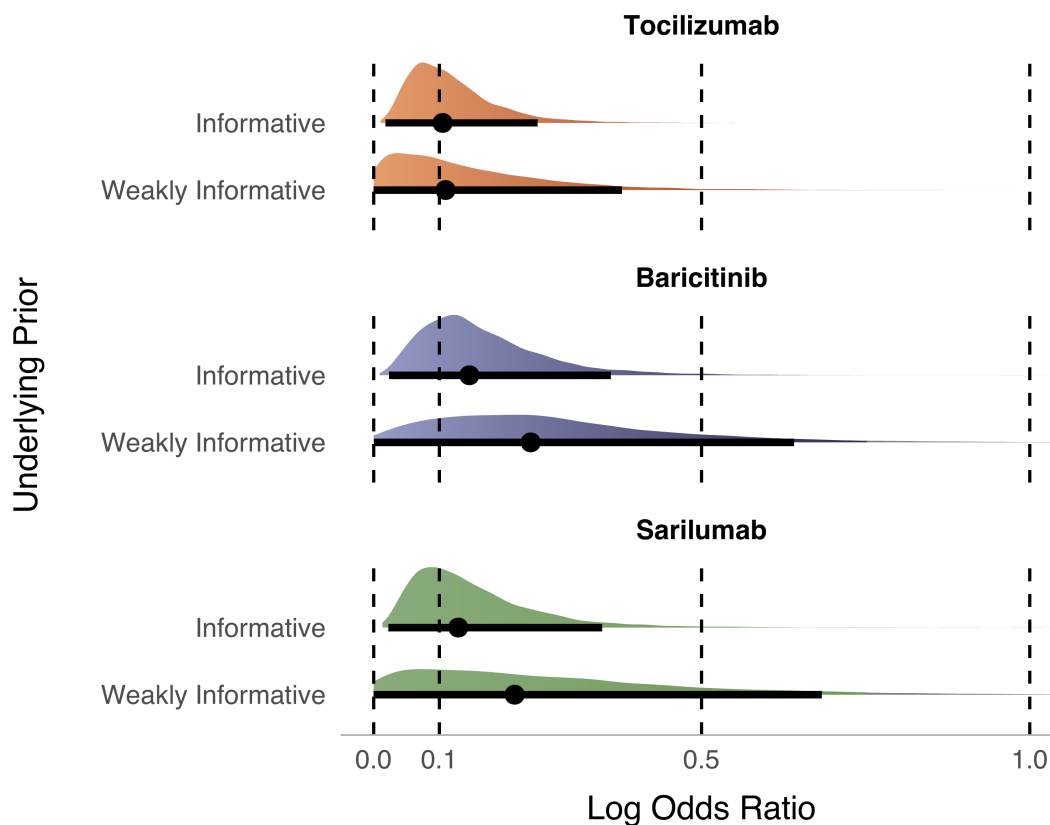
Appendix Figure 4: Baricitinib Sensitivity Meta-Analysis



Sensitivity analysis including all patients treated with corticosteroids, regardless of tocilizumab use, in RECOVERY Bari.

Forest plot of Bayesian random-effect meta-analysis of baricitinib versus control. Black diamond represents the median and 95% credible interval of posterior overall effect (μ). Purple diamonds represents the 95% prediction interval on the posterior predictive distribution. The median and 95% credible interval of the between-study standard deviation parameter (τ) are displayed on the left bottom corner. Abbreviations: RE, random-effect; CrI, credible interval; PI, prediction interval.

Appendix Figure 5: Posterior Between-Study Standard Deviations



Between study standard deviation (τ) marginal posterior distributions upon different underlying prior distributions (informative or weakly informative) for τ . This parameter is an estimate for between-study heterogeneity in random-effect meta-analyses. Point estimates depict the median, while intervals bars 95% CrIs. Dashed lines represent different heterogeneity ranges: from 0 to 0.1 log odds ratio, “Low heterogeneity”; from 0.1 to 0.5, “Reasonable”; from 0.5 to 1.0, “Fairly High” (per Spiegelhalter et al., (3)).

For τ , the “Informative” prior regards a Log-Normal($-1.975, 0.67^2$) distribution, while “Weakly Informative” regards a Half-Normal(0.5) distribution. All models assume a $Normal(0, 0.75^2)$ prior distribution for average effect (μ).

Abbreviation: CrI, credible interval

Appendix Table 2: Posterior Probabilities of Heterogeneity

	Posterior Probability within Heterogeneity Ranges, %		
	Low	Reasonable	Fairly High
Tocilizumab			
Informative	50.0	49.9	0.1
Weakly Informative	51.1	48.2	0.7
Baricitinib			
Informative	20.6	77.9	1.5
Weakly Informative	11.2	73.6	14.5
Sarilumab			
Informative	34.2	65.0	0.8
Weakly Informative	23.7	62.3	13.5

Low heterogeneity range: between 0 and 0.1 log odds ratio; Reasonable range: between 0.1 and 0.5; Fairly high: between 0.5 and 1.0.

Appendix Table 3: Meta-Analyses Summary Results

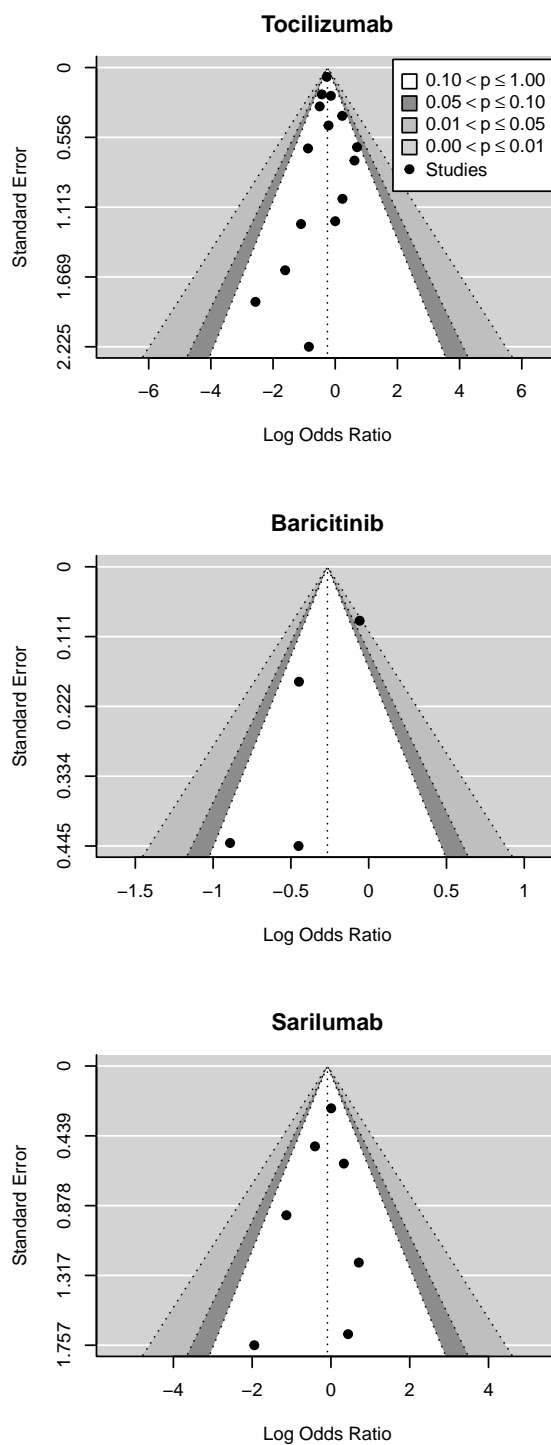
Treatment / Prior	Posterior Results, Median and 95% CrI		
	Average Effect ¹	Between-study SD ²	Predictive Effect ¹
Tocilizumab			
Informative	0.78 [0.65, 0.94]	0.1 [0.02, 0.24]	0.78 [0.58, 1.09]
Weakly Informative	0.78 [0.64, 0.95]	0.1 [0, 0.34]	0.77 [0.53, 1.15]
Baricitinib			
Informative	0.78 [0.56, 1.03]	0.16 [0.03, 0.38]	0.78 [0.44, 1.26]
Weakly Informative	0.75 [0.46, 1.07]	0.27 [0, 0.69]	0.76 [0.31, 1.79]
Sarilumab			
Informative	0.9 [0.59, 1.41]	0.13 [0.02, 0.35]	0.9 [0.51, 1.56]
Weakly Informative	0.89 [0.51, 1.41]	0.22 [0, 0.68]	0.89 [0.36, 1.99]

¹Odds Ratio

²Log Odds Ratio

Informative Between-study SD Prior: LN(-1.975, 0.67); Weakly Informative Prior: HN(0.5);
LN(mu, SD) = Log-Normal(mean, SD); HN(SD) = Half-Normal(SD)

Appendix Figure 6: Funnel Plots

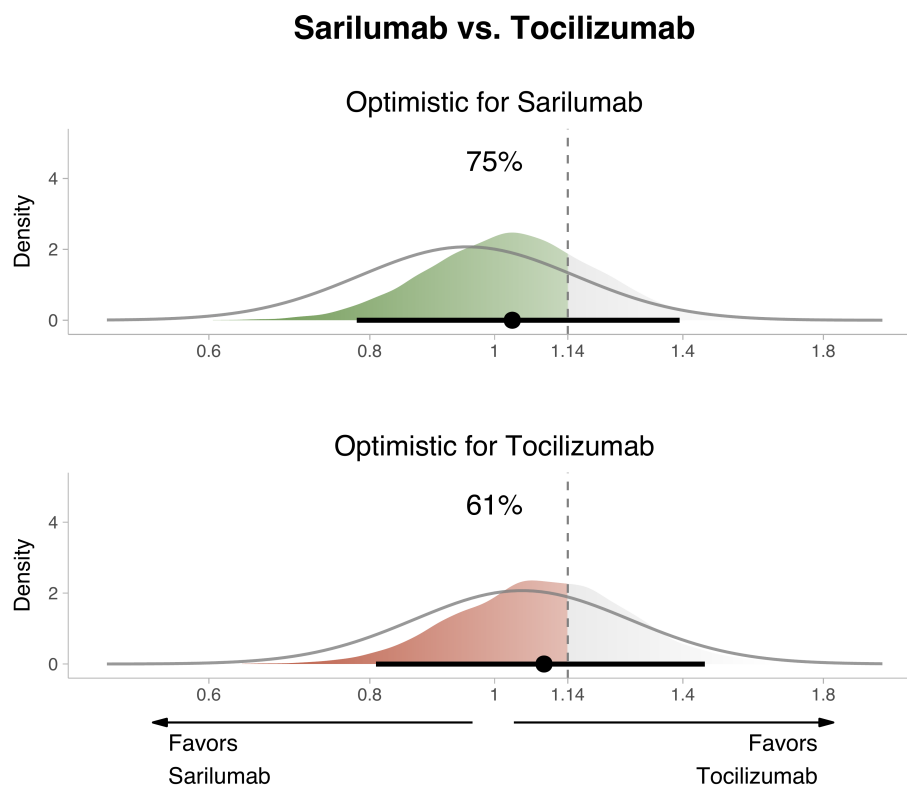


Contour-enhanced funnel plots for small study effect assessment.

Appendix Table 4: Certainty of Evidence (GRADE)

Drug	Risk of Bias	Inconsistency	Imprecision	Indirectness	Publication Bias	Judgment
Tocilizumab	We did not downgrade as the most weighted studies were judged as low risk of bias.	Certainty was downgraded by 1 level due to heterogeneity, as the prediction interval suggests that future studies could estimate different results leading to opposite clinical decisions (benefit as large as 44% reduction, harm as large as 10% increase).	We did not downgrade due to imprecision, although by a close call, despite reasonably large credible intervals (35% reduction to 6% reduction in mortality) as the upper bound of the CrI suggests clinically relevant benefit.	We did not find compelling evidence to downgrade due to indirectness.	We did not find compelling evidence to downgrade due to publication bias.	MODERATE QUALITY OF EVIDENCE, downgraded by 1 level due to inconsistency
Baricitinib	We did not downgrade as most studies were of reasonably low risk of bias or minor concerns, unlikely to have materially biased the finals results with clinically meaningful magnitudes.	We did not downgrade due to heterogeneity as tau was not judged as a matter of concern (tau = 0.16) and the prediction interval is consistent with the estimated credible interval (future studies would not change the conclusions of the current range of possible effects).	Certainty was downgraded by 1 level due to imprecision, as the credible interval includes benefit but also includes null or trivially harmful effects.	We did not find compelling evidence to downgrade due to indirectness.	We did not find compelling evidence to downgrade due to publication bias.	MODERATE QUALITY OF EVIDENCE, downgraded by 1 level due to imprecision
Sarilumab	We did not downgrade as the most weighted studies were judged as low risk of bias.	We did not downgrade due to heterogeneity as tau was not judged as a matter of concern (tau = 0.13) and the prediction interval is consistent with the estimated credible interval (future studies would not change the conclusions of the current range of possible effects).	Certainty was downgraded by 2 levels due to serious imprecision (the credible interval includes both clinically relevant benefits and harms).	We did not find compelling evidence to downgrade due to indirectness.	We did not find compelling evidence to downgrade due to publication bias.	LOW QUALITY OF EVIDENCE, downgraded by 2 levels due to serious imprecision.

Appendix Figure 7: Sarilumab vs. Tocilizumab Sensitivity Analysis



Post-hoc sensitivity analysis on sarilumab vs. tocilizumab meta-regression model (ratio of odds ratios). Color filled curves represent marginal posterior distributions. Color filled areas represent the posterior probability of noninferiority ($\text{Pr} < 1.14$), as the percentages on top of each panel. Interval bars depict the posterior median and 95% credible intervals. Solid gray lines represent underlying prior distributions. Each belief is labeled on top of each figure.

Underlying prior distributions: “Optimistic for Sarilumab,” $\text{Normal}(-0.049, 0.193^2)$; “Optimistic for Tocilizumab,” $\text{Normal}(0.049, 0.193^2)$

Appendix Table 5: Sarilumab vs. Tocilizumab Sensitivity Analysis

Belief	ROR (95% CrI)	Probability of Noninferiority, % (1)	Probability of Superiority, % (2)
Optimistic for Sarilumab	1.03 (0.76, 1.34)	75	41
Optimistic for Tocilizumab	1.09 (0.81, 1.45)	60	28

Note:

Abbreviations: ROR, ratio of odds ratios; CrI, credible interval

¹ Probability of Noninferiority, %

² Probability of Superiority, %

Underlying prior distributions: “Optimistic for Sarilumab,” $\text{Normal}(-0.049, 0.193^2)$; “Optimistic for Tocilizumab,” $\text{Normal}(0.049, 0.193^2)$

Appendix Table 6: Baricitinib vs. Tocilizumab Sensitivity Analysis

Belief	ROR (95% CrI)	Probability of Noninferiority, % (1)	Probability of Superiority, % (2)
Vague	0.99 (0.7, 1.29)	85	53
Skeptical	1 (0.73, 1.27)	85	51
Optimistic for Baricitinib	0.98 (0.69, 1.25)	87	56
Optimistic for Tocilizumab	1.01 (0.72, 1.27)	83	48

Note:

Abbreviations: ROR, ratio of odds ratios; CrI, credible interval

¹ Probability of Noninferiority, %

² Probability of Superiority, %

Underlying prior distributions: “Skeptical,” $\text{Normal}(0, 0.354^2)$; “Optimistic for Baricitinib,” $\text{Normal}(-0.105, 0.416^2)$; “Optimistic for Tocilizumab,” $\text{Normal}(0.105, 0.416^2)$; “Vague,” $\text{Normal}(0, 4^2)$;