Supplementary Material

Arthur M. Albuquerque, Lucas Tramujas, Lorenzo R. Sewanan, Donald R. Williams, James M. Brophy

Table of Contents

# eMethods

## Bayesian meta-analysis

Our meta-regression random-effect model is defined as:

where is the observed mean log odds ratio of tocilizumab versus control and is the known sampling variance in study . Because this is a random-effect model, each study has its own distribution, where represents its mean effect. All s are drawn from normal distribution where the mean effect is and the variance , which represents the between-study heterogeneity. is predicted by a no-intercept linear regression, where each subgroup (simple oxygen only; noninvasive ventilation; invasive mechanical ventilation) has its own parameter , representing the mean effect of each respective subgroup.

In this case, we are able to assess tocilizumab’s effect in each subgroup while assuming a common between-study heterogeneity:

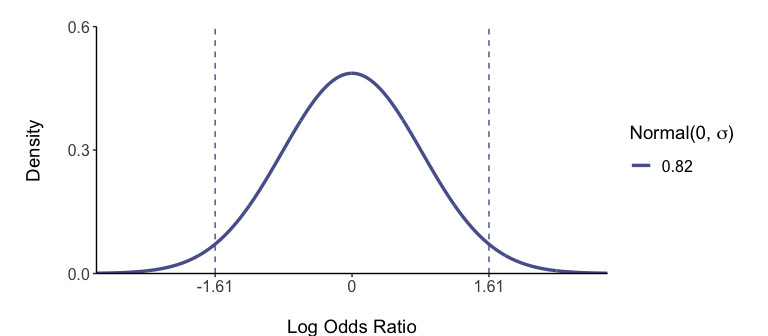
* Simple oxygen only =
* Noninvasive ventilation =
* Invasive mechanical ventilation =

### Weakly informative priors

Because we applied the Bayesian framework, we assigned a prior distribution for each parameter. In our main model, we implemented priors that cover plausible values for all parameters, assigning limited density to impossible values, and thus employed little influence in the results (hereafter, known as weakly informative priors). These are our weakly informative priors:

Now, we will explain the rationale underlying these distributions.

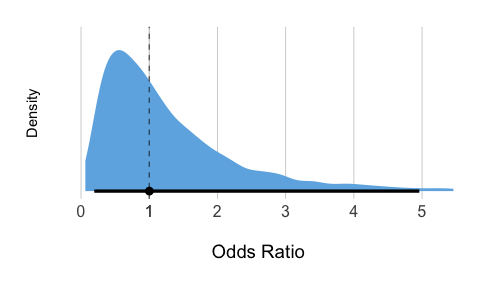
We find highly unlikely that a pharmacological treatment, such as tocilizumab, will yield a 80% odds reduction in 28-days all-cause mortality regardless of the subgroup of patients, as suggested by empirical evidence. Thus, for , we set a prior distribution of in the log odds ratio scale.



To facilitate the understanding of these distributions, here are the corresponding 95% quantile intervals in the linear scale.

| Log scale | | Linear scale | |
| --- | --- | --- | --- |
| Mean | SD | Mean | 95% CI |
| 0 | 0.82 | 1 | [0.2, 5] |

Another way to assess the plausibility of the aforementioned priors is to perform a prior predictive check, which can be visualized below:



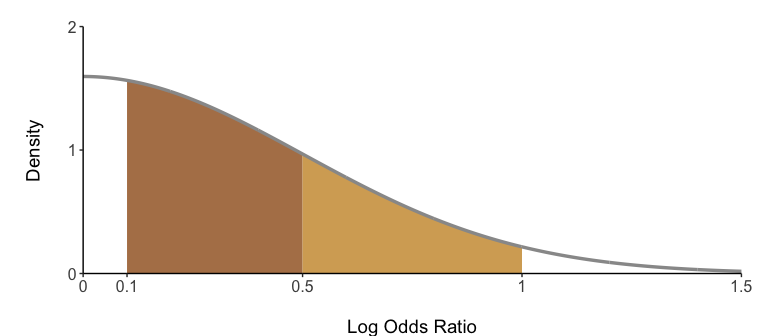
Point estimate depicts the median and interval bar depicts the 95% quantile interval.

As expected, the distribution approximately ranges from 0.2 to 5.0.

Lastly, we will now discuss the weakly informative prior distribution for . Because we wanted to perform unconditional inferences beyond the included studies, we fitted a random-effect meta-analysis. In this model, one assumes there is within-study heterogeneity (represented by , the known sampling variance in study ) and the between-study heterogeneity (represented by ).

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, “fairly high” between 0.5 and 1.0, and “fairly extreme” for values larger than 1.0 log odds ratio).[7,13] We added a category for low heterogeneity (between 0 and 0.1).

The distribution yields plausible probabilities in each of these ranges.



Here are the corresponding probabilities within each of the aforementioned heterogeneity ranges:

| Heterogeneity Range | | | |
| --- | --- | --- | --- |
| Low | Reasonable | Fairly High | Fairly Extreme |
| 16% | 52% | 27% | 5% |

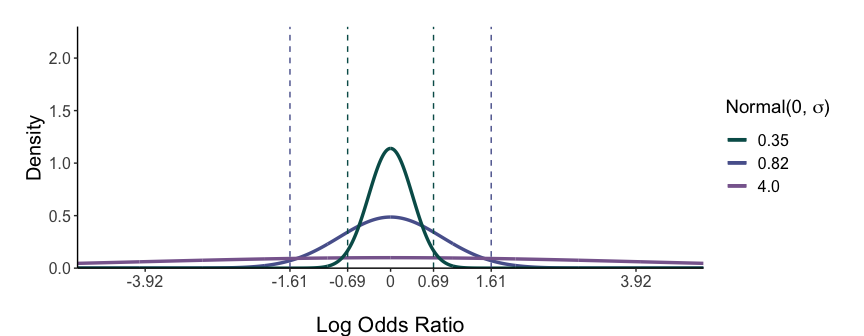
### Alternative priors

To check whether the choice of weakly informative priors meaningfully impacted our results or our conclusions, we also fitted models using vague or informative priors.

Vague priors:

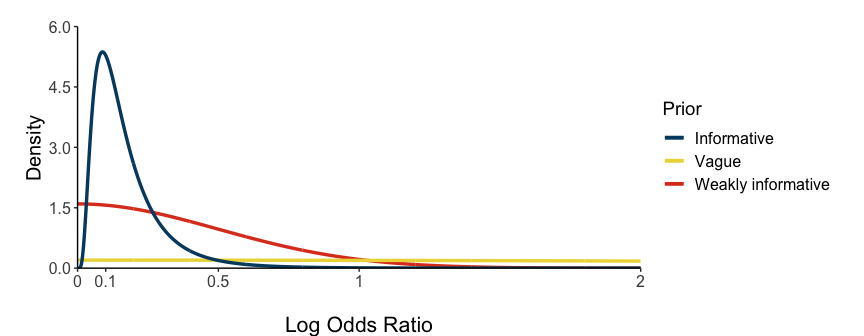
Informative priors:

Here are graphical representations of these normal distributions (along with the weakly informative mentioned before):



| Log scale | | Linear scale | |
| --- | --- | --- | --- |
| Mean | SD | Mean | 95% CI |
| 0 | 0.35 | 1 | [0.5, 2] |
| 0 | 0.82 | 1 | [0.2, 5] |
| 0 | 4.00 | 1 | [0, 2540.2] |

Here are graphical representations of distributions for the between-study standard deviation () (along with the weakly informative mentioned before):



|  | | Heterogeneity Range | | | |
| --- | --- | --- | --- | --- | --- |
| Prior | Distribution | Low | Reasonable | Fairly High | Fairly Extreme |
| Weakly informative | Half-Normal(0.5) | 16% | 52% | 27% | 5% |
| Vague | Half-Normal(4.0) | 2% | 8% | 10% | 80% |
| Informative | Log-Normal(-1.975, 0.67) | 31% | 66% | 3% | 0% |

## Predictive analysis to confirm tocilizumab’s association with mortality benefit

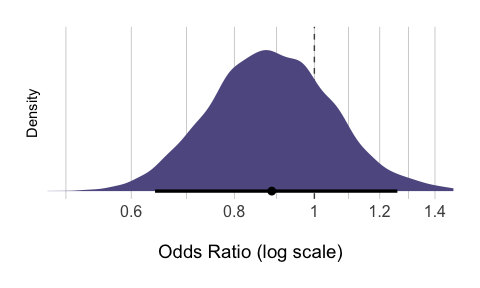
Lastly, we will update our current evidence (as modeled in our main meta-regression model) with generated randomized clinical trials (RCTs) of different sample sizes comparing tocilizumab to control on patients on invasive mechanical ventilation.

We will use the estimated marginal posterior mean and standard error on this subgroup to create a prior distribution. Then, we will use normal conjugate analyses to update this prior with new data (likelihood), and form an updated posterior distribution.

### Prior

As described before, we fitted a Bayesian meta-regression model, from which we estimated marginal posterior distributions on different subgroups.

Here, the subgroup of interest is the invasive mechanical ventilation:



Marginal posterior distribution of the invasive mechanical ventilation subgroup. The interval bar depicts the mean and 95% quantile interval.

In the linear scale, the mean of this marginal posterior distribution is 0.89. Because we will use normal conjugate analysis, it is of greater interest to evaluate this distribution on the log scale, which is approximately normally distributed. In this case, the mean is -0.12 and the standard error is 0.17.

### Likelihood

We will create six different RCTs and update the prior distribution mentioned above six separate times.

Assuming the prior is normally distributed

and so is the data (likelihood).

The mean and variance of the posterior distributions can be estimated by the following formulas:

* Mean
* Variance

In summary, we are able to update a normally distributed prior distribution (shown in the Figure above) with normally distributed data to generate a normally distributed posterior distribution. Based on the posterior’s mean and variance, we will generate 100,000 random samples (seed number of 123).

As mentioned before, we will create six different RCTs to separately generate six different posterior distributions (from the same prior distribution). Therefore, we now have to decide the mean and standard deviation of the likelihood.

All RCTs will have a mean of -0.26 (log scale). This value is the equal to 0.77 in the linear scale, which was chosen based on WHO’s meta-analysis ([page 14 in Supplement 2](doi:10.1001/jama.2021.11330)). This is the mean odds ratio of tocilizumab vs. control in patients using corticosteroids (overall results).

Here are our assumptions to calculate the standard deviation of each RCTs:

1. We assume equal allocation in both treatment arms (tocilizumab and control).
2. Adapting from the suggestions in the [GRADE guidelines](http://dx.doi.org/10.1016/j.jclinepi.2012.01.012), we found a striking discrepancy between the control mortality risk in this data (52%) in comparison to another [previously published meta-analysis](https://jamanetwork.com/journals/jama/fullarticle/2770279?utm_campaign=articlePDF&utm_medium=articlePDFlink&utm_source=articlePDF&utm_content=jama.2020.17023) (34% in patients on IVM and using corticosteroids). Thus, we have decided to use 43% (arithmetic mean between 34 and 52) as our reference risk in the IVM subgroup..
3. The mortality risk in the tocilizumab was calculated using the following formula (Box 1 in [Doi et al., 2020](https://doi.org/10.1016/j.jclinepi.2020.08.019)):

where is the mortality risk in the tocilizumab group, is the mortality risk in the control group and is the odds ratio mentioned above. Thus, the tocilizumab risk is equal to 0.37.

Thus, we are simulating RCTs with mean OR equal to 0.77, control risk mortality of 0.43, and tocilizumab risk of 0.37.

Based on these values, we can estimate the standard deviation () with the following [formula](https://doi.org/10.1016/j.jclinepi.2008.07.006):

where , , and are number of events and follow this 2x2 table:

| Event | Tocilizumab | Control |
| --- | --- | --- |
| Death | a | c |
| No death | b | d |

As similarly shown in the supplementary material of [Higgins and Spiegelhalter, 2002](https://academic.oup.com/ije/article/31/1/96/655931?login=true), we can estimate these values as:

where and are the sample sizes in the tocilizumab and control arms, respectively. As mentioned above, we assume equal allocation in both treatment arms, thus .

Finally, we can estimate the (log scale) based on the 6 different sample sizes mentioned above:

| Sample size in each treatment arm | Total sample size | SD |
| --- | --- | --- |
| 100 | 200 | 0.29 |
| 250 | 500 | 0.18 |
| 500 | 1,000 | 0.13 |
| 750 | 1,500 | 0.11 |
| 1,000 | 2,000 | 0.09 |
| 2,000 | 4,000 | 0.06 |

## Deriving risk difference from odds ratio

We used the odds ratio as our primary estimand, following the suggestions by [Doi et al., 2020](https://doi.org/10.1016/j.jclinepi.2020.08.019) and [Doi et al., 2021](https://doi.org/10.1016/j.jclinepi.2021.08.003).

Moreover, we derived the risk in the tocilizumab group using the following formula (Box 1 in [Doi et al., 2020](https://doi.org/10.1016/j.jclinepi.2020.08.019)):

where is the mortality risk in the tocilizumab group, is the mortality risk in the control group and is the odds ratio.

We then calculated the risk difference (RD) with the following formula, which was also the procedure suggested by [Doi et al., 2020](https://doi.org/10.1016/j.jclinepi.2020.08.019) (page 4):

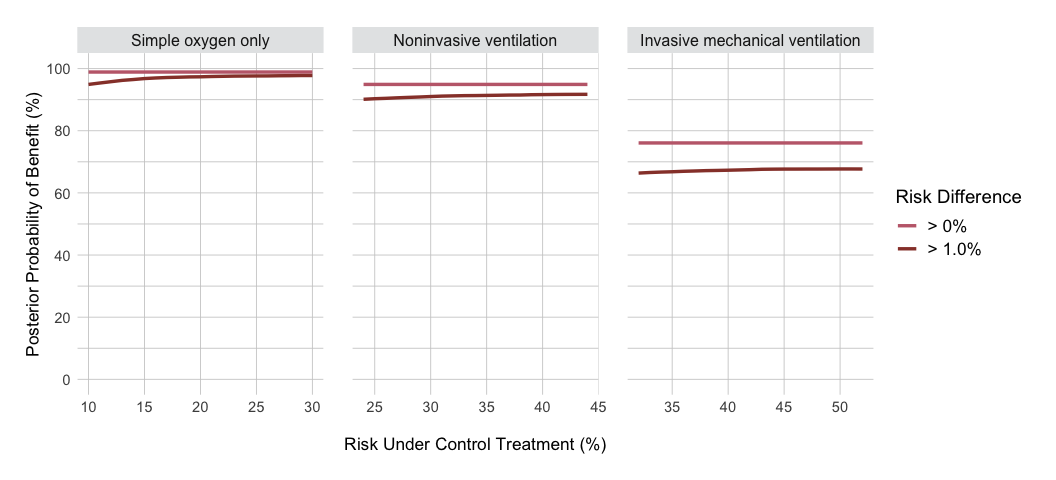
Adapting from the suggestions in the [GRADE guidelines](http://dx.doi.org/10.1016/j.jclinepi.2012.01.012), we assumed different mortality risks in each subgroup. For the simple oxygen only and noninvasive ventilation subgroups, we used the average mortality risk in each subgroup based on the data included in this reanalysis of a [previously published meta-analysis](10.1001/jama.2021.11330). In contrast, regarding the invasive mechanical ventilation (IVM) subgroup, we found a striking discrepancy between the control mortality risk in this data (52%) in comparison to another [previously published meta-analysis](https://jamanetwork.com/journals/jama/fullarticle/2770279?utm_campaign=articlePDF&utm_medium=articlePDFlink&utm_source=articlePDF&utm_content=jama.2020.17023) (34% in patients on IVM and using corticosteroids). Thus, we have decided to use 43% (arithmetic mean between 34 and 52) as our reference risk in the IVM subgroup. Recognizing the potential variability of the subgroup baseline risks, we estimated the risk differences with twenty different plausible baseline risks for each subgroup (spanning +- 10% change from the reference risks mentioned above).

| Subgroup | Control Risk |
| --- | --- |
| Simple oxygen only | 20 +- 10% |
| Noninvasive ventilation | 34 +- 10% |
| Invasive mechanical ventilation | 42 +- 10% |

# eFigures and eTables

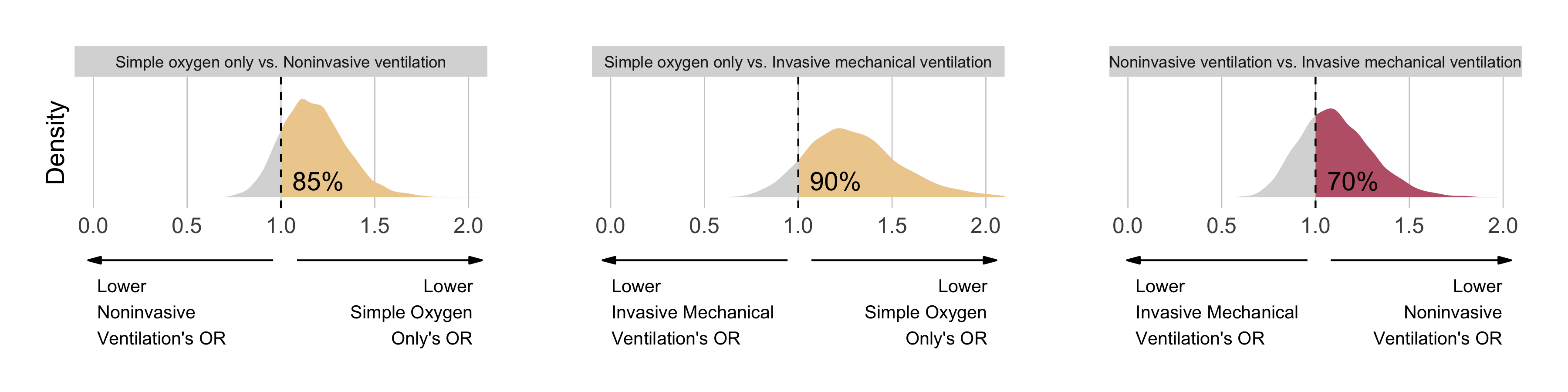
## eTable 1

## eFigure 1



Posterior probabilities of benefit per subgroup in the risk difference scale assuming weakly informative priors. Each line represents the posterior probability of benefit for a specific cutoff, such as risk difference greater than 0% or 1%, across plausible ranges of mortality risk under control treatment. Underlying weakly informative priors are N(0, 0.82) for the mean effect, N(0, 1.5) for the coefficients, and HN(0.5) for the between-study standard deviation. N(mu, sigma) = Normal(mean, standard deviation); HN(sigma) = Half-Normal(standard deviation)

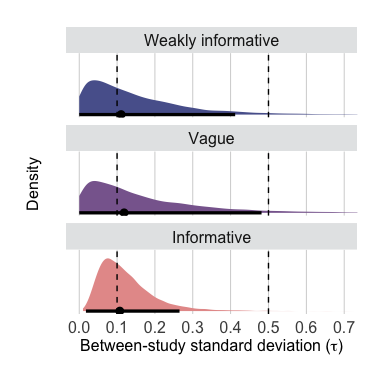
## eFigure 2



Posterior distributions for comparisons of effect sizes between subgroups while assuming weakly informative priors. Each distribution represents the ratio of odds ratios of two subgroups. On top of each distribution, there is a percentage representing the posterior probability of a ratio of odds ratios greater than 1.0. Arrows on the bottom represent - in that comparison - which subgroup benefited to a greater extent tocilizumab’s effect on mortality reduction. For example, assuming our model, there was a 85% probability that tocilizumab reduces mortality to a greater extent in the simple oxygen subgroup in comparison to noninvasive ventilation. Underlying weakly informative priors are N(0, 0.82) for the mean effect, N(0, 1.5) for the coefficients, and HN(0.5) for the between-study standard deviation. N(mu, sigma) = Normal(mean, standard deviation); HN(sigma) = Half-Normal(standard deviation).

## eTable 2

## eFigure 3



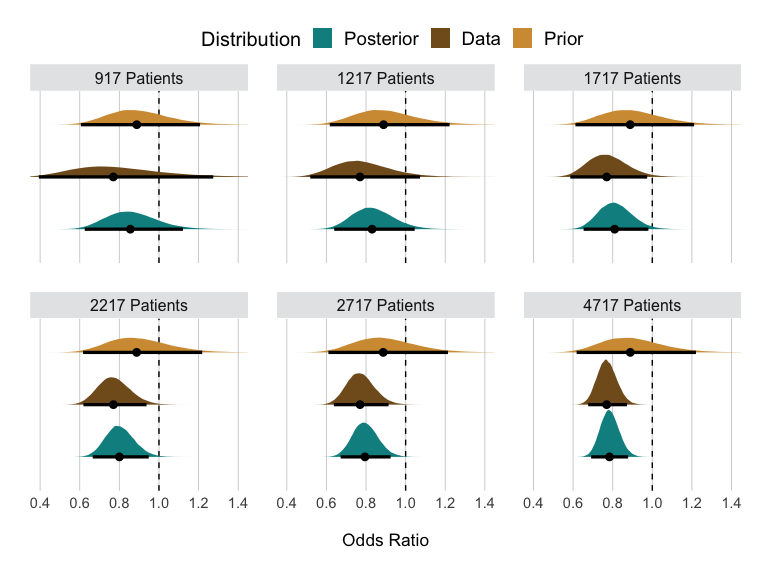
Posterior distributions (log scale) of the between study standard deviation (tau) upon different underlying prior distributions (weakly informative, vague or informative). Tau is a proxy for the between-study heterogeneity in random-effect meta-analyses. Weakly informative priors: Intercept N(0, 0.82); Coefficients N(0, 1.5); Between-study standard deviation HN(0.5) / Vague priors: Intercept N(0, 4); Coefficients N(0, 4); Between-study standard deviation HN(4) / Informative priors: Intercept N(0, 0.35); Coefficients N(0, 0.2); Between-study standard deviation LN(-1.975, 0.67). N(mu, sigma) = Normal(mean, standard deviation); HN(sigma) = Half-Normal(standard deviation); LN(mu, sigma) = Log-Normal(mean, standard deviation).

## eTable 3

## eTable 4

## eTable 5

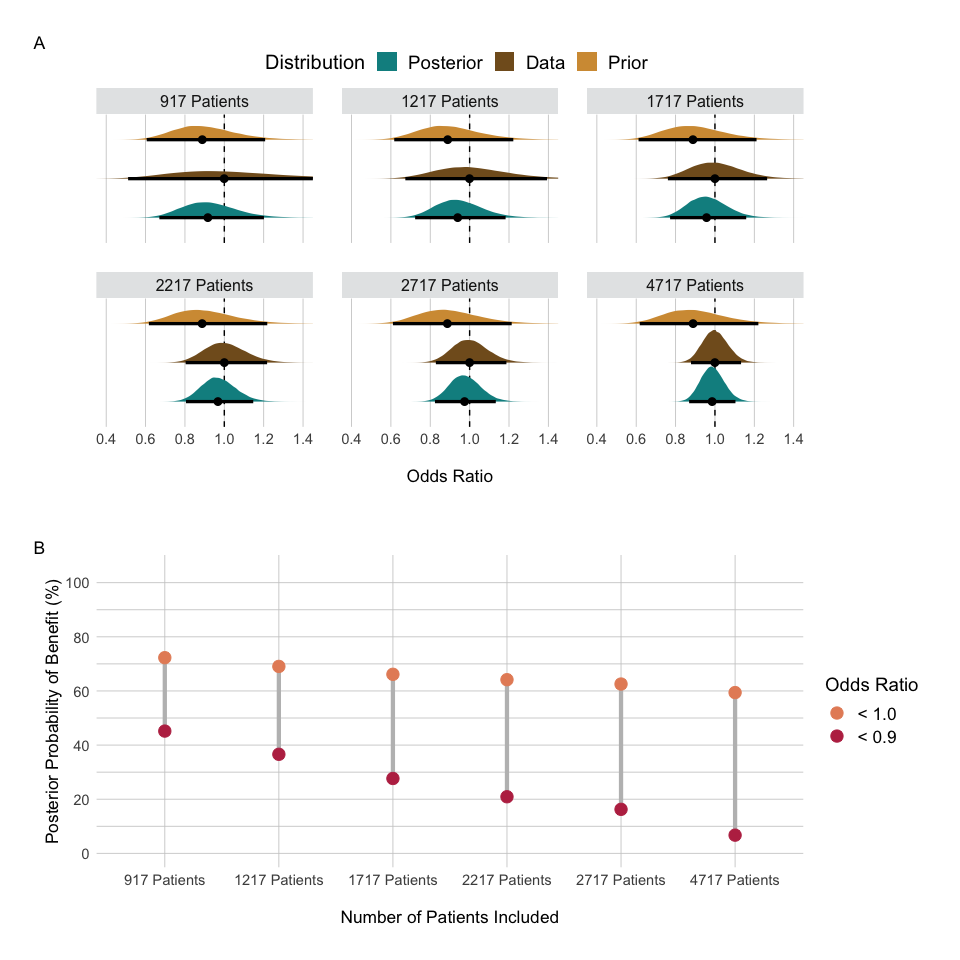
## eFigure 4



Results from the normal conjugate analyses updating current evidence on invasive mechanical ventilation (used as the Prior) with generated RCTs (used as Data, eTable 5). These analyses yield posterior distributions. Each panel represents a different model, in which the prior distribution is centered at 0.77 odds ratio. The label on top of each panel depict the number of total patients on invasive mechanical ventilation included in each respective model (current plus generated patients). Point estimates depict the median and interval bars represent the 95% credible intervals for both prior, data (likelihood) and posterior distributions.

## eTable 6

## eFigure 5



Results from the normal conjugate analyses updating current evidence on invasive mechanical ventilation (used as the Prior) with generated RCTs (used as Data, eTable 6). These analyses yield posterior distributions as depicted on Panel A. In contrast to the results shown in eFigure4, the genereated RCTs in these analyses are centered at 1.0 odds ratio. Panel A: Each panel represents a different model. The label on top of each panel depict the number of total patients on invasive mechanical ventilation included in each respective model (current plus generated patients). Point estimates depict the median and interval bars represent the 95% credible intervals for both prior, data (likelihood) and posterior distributions. Panel B shows the posterior probability of benefit for different thresholds (OR < 1.0 and < 0.9).