


ORIGINAL



Dexamethasone 12 mg versus 6 mg for patients with COVID-19 and severe hypoxaemia: a pre-planned, secondary Bayesian analysis of the COVID STEROID 2 trial

Anders Granholm^{1,2*} , Marie Warrer Munch^{1,2}, Sheila Nainan Myatra³, Bharath Kumar Tirupakuzhi Vijayaraghavan^{4,5,6}, Maria Cronhjort⁷, Rebecka Rubenson Wahlin⁷, Stephan M. Jakob⁸, Luca Cioccarì⁸, Maj-Brit Nørregaard Kjær^{1,2}, Gitte Kingo Vesterlund^{1,2}, Tine Sylvest Meyhoff^{1,2}, Marie Helleberg⁹, Morten Hylander Møller^{1,2}, Thomas Benfield¹⁰, Balasubramanian Venkatesh¹¹, Naomi E. Hammond^{11,12}, Sharon Micallef¹¹, Abhinav Bassi⁶, Oommen John^{6,13}, Vivekanand Jha^{6,13,14}, Klaus Tjelle Kristiansen¹⁵, Charlotte Suppli Ulrik¹⁶, Vibeke Lind Jørgensen¹⁷, Margit Smitt¹⁸, Morten H. Bestle^{19,20}, Anne Sofie Andreasen²¹, Lone Musaeus Poulsen²², Bodil Steen Rasmussen^{2,23}, Anne Craveiro Brøchner^{2,24}, Thomas Strøm^{25,26}, Anders Møller²⁷, Mohd Saif Khan²⁸, Ajay Padmanaban⁴, Jigeeshu Vasishtha Divatia³, Sanjith Saseedharan²⁹, Kapil Borawake³⁰, Farhad Kapadia³¹, Subhal Dixit³², Rajesh Chawla³³, Urvi Shukla³⁴, Pravin Amin³⁵, Michelle S. Chew³⁶, Christian Aage Wamberg³⁷, Christian Gluud^{38,39}, Theis Lange⁴⁰ and Anders Perner^{1,2}

© 2021 Springer-Verlag GmbH Germany, part of Springer Nature

Abstract

Purpose: We compared dexamethasone 12 versus 6 mg daily for up to 10 days in patients with coronavirus disease 2019 (COVID-19) and severe hypoxaemia in the international, randomised, blinded COVID STEROID 2 trial. In the primary, conventional analyses, the predefined statistical significance thresholds were not reached. We conducted a pre-planned Bayesian analysis to facilitate probabilistic interpretation.

Methods: We analysed outcome data within 90 days in the intention-to-treat population (data available in 967 to 982 patients) using Bayesian models with various sensitivity analyses. Results are presented as median posterior probabilities with 95% credible intervals (CrIs) and probabilities of different effect sizes with 12 mg dexamethasone.

Results: The adjusted mean difference on days alive without life support at day 28 (primary outcome) was 1.3 days (95% CrI −0.3 to 2.9; 94.2% probability of benefit). Adjusted relative risks and probabilities of benefit on serious adverse reactions was 0.85 (0.63 to 1.16; 84.1%) and on mortality 0.87 (0.73 to 1.03; 94.8%) at day 28 and 0.88 (0.75 to 1.02; 95.1%) at day 90. Probabilities of benefit on days alive without life support and days alive out of hospital at day 90 were 85 and 95.7%, respectively. Results were largely consistent across sensitivity analyses, with relatively low probabilities of clinically important harm with 12 mg on all outcomes in all analyses.

*Correspondence: andersgran@gmail.com

¹ Department of Intensive Care, Rigshospitalet—Copenhagen University Hospital, Blegdamsvej 9, 2100 Copenhagen, Denmark

Full author information is available at the end of the article

Conclusion: We found high probabilities of benefit and low probabilities of clinically important harm with dexamethasone 12 mg versus 6 mg daily in patients with COVID-19 and severe hypoxaemia on all outcomes up to 90 days.

Keywords: COVID-19, Hypoxaemia, Critical illness, Corticosteroids, Bayesian analysis

Introduction

Coronavirus disease 2019 (COVID-19) may lead to critical illness and severe hypoxaemia [1]. As of September 2021, the ongoing COVID-19 pandemic has caused >4.5 million deaths worldwide [2]. Systemic corticosteroids decrease mortality in critically ill patients with COVID-19 [3], and dexamethasone 6 mg daily for up to 10 days is therefore recommended by the World Health Organization for patients with severe or critical COVID-19 [4]. Higher doses of systemic corticosteroids have been used in patients with COVID-19 and non-COVID-19 acute respiratory distress syndrome [3, 5, 6], and higher doses have been hypothesised to benefit patients with severe or critical COVID-19, although the balance between benefit and harm remains uncertain.

We conducted the COVID STEROID 2 trial to compare a higher (12 mg) versus the recommended dose (6 mg) of dexamethasone daily for up to 10 days in patients with COVID-19 and severe hypoxaemia [5, 7]. In the primary frequentist statistical analysis, the adjusted mean difference for days alive without life support up to 28 days after randomisation (primary outcome) was 1.3 days (95% confidence interval [CI]: 0–2.6, $p=0.07$) higher with 12 mg, and the adjusted relative risks (RR) for 28-day mortality and serious adverse reactions were 0.86 (99% CI: 0.68–1.08, $p=0.1$) and 0.83 (99% CI: 0.54–1.29, $p=0.27$) with 12 mg, respectively [7]. While the pre-defined thresholds for statistical significance were not reached in these analyses or in the analyses of outcomes registered at day 90, the results were more compatible with benefit with 12 mg, and consequently, a nuanced interpretation avoiding arbitrary dichotomisations is warranted [8, 9].

Here, we report the secondary, pre-planned Bayesian analyses of all outcomes registered up to day 90 in the COVID STEROID 2 trial to facilitate a more nuanced and probabilistic interpretation of the trial results [10].

Methods

We report the pre-planned secondary Bayesian analysis of the COVID STEROID 2 trial of all outcomes registered within 90 days. This study was conducted according to a published protocol and statistical analysis plan [10], with this manuscript prepared according to the Strengthening the Reporting of Observational Studies

Take-home message

In this Bayesian analysis of the COVID STEROID 2 trial, we found relatively high posterior probabilities of benefit with dexamethasone 12 mg versus 6 mg in patients with COVID-19 and severe hypoxaemia on all outcomes, including the days alive without life support and mortality at day 28 and 90. We found relatively low probabilities of clinically important harm with 12 mg dexamethasone for all outcomes.

in Epidemiology (STROBE) checklist (see Electronic Supplementary Material, ESM) and adhering to the Reporting of Bayes Used in clinical STudies (ROBUST) guideline [11, 12].

The COVID STEROID 2 trial

The COVID STEROID 2 trial was an investigator-initiated, international, centrally randomised, stratified (for site, use of invasive mechanical ventilation, and age below 70 years), parallel-group, blinded clinical trial. Adult patients hospitalised with COVID-19 and severe hypoxaemia (≥ 10 L oxygen supplementation/minute independent of delivery system, use of non-invasive ventilation or continuous positive airway pressure for hypoxaemia, or invasive mechanical ventilation) were included; exclusion criteria were primarily related to use of systemic corticosteroids for other indications than COVID-19 in doses >6 mg dexamethasone equivalents or previous use of systemic corticosteroids for COVID-19 for ≥ 5 days, invasive fungal infection or active tuberculosis, pregnancy or unobtainable consent (detailed in the ESM and elsewhere [5, 7, 10]).

Patients were randomised to dexamethasone 12 mg or 6 mg intravenously daily for maximum 10 days (depending on the number of consecutive days with corticosteroid treatment before randomisation; ESM). Randomisation took place between August 27, 2020 and May 20, 2021 at 31 sites in 26 hospitals in Denmark, India, Sweden, and Switzerland [7].

The trial was approved by the Ethics Committee of the Capital Region of Denmark with additional national/local approvals as required; additional details on approvals and consent procedures and the primary trial results are available elsewhere [5, 7, 10].

Outcomes

Primary outcome

Days alive without life support at day 28 (days alive without invasive mechanical ventilation, circulatory support, and kidney replacement therapy).

Secondary outcomes

1. One or more serious adverse reactions (new episodes of septic shock, invasive fungal infection, clinically important gastrointestinal bleeding, or anaphylactic reaction to intravenous dexamethasone) within 28 days of randomisation.
2. All-cause mortality at day 28.
3. All-cause mortality at day 90.
4. Days alive without life support at day 90.
5. Days alive and out of hospital at day 90.

Detailed definitions are provided in the ESM and elsewhere [5, 7, 10].

Statistical analyses

Analyses were conducted using R version 4.1.0 with the Tidyverse packages [13] and Stan (CmdStan version 2.26.1) [14] through the brms R package [15], with all analyses adjusted for the stratification variables (with sites with few patients in each country merged for these analyses). Technical details (including sampler settings and model diagnostics) are presented in the ESM and elsewhere [10].

Bayesian analyses and priors

Bayesian analyses start with prior probability distributions, representing prior beliefs, which are updated once data has been collected to posterior probability distributions that allow straightforward interpretation and calculation of direct probabilities [10, 16]. We estimated adjusted outcome data in each group from the joint posterior distributions and used these estimates to derive adjusted relative and absolute treatment effects. We planned to present conditional adjusted estimates using a reference patient in each group with all adjustment variables set to their most common value [10], but due to substantial differences between outcomes across trial sites, results would be difficult to interpret using this approach. Consequently, we primarily present average estimates and average treatment effects, secondarily supplemented with estimates calculated for three different representative reference patients (ESM). Posteriors were summarised by calculating direct probabilities of a number of different pre-defined effect sizes, and summarised with point estimates (posterior medians) and

percentile-based 95% credible intervals (CrIs) that represent the 95% most probable effects given the prior, the model, and the data [10, 16]. In addition, we graphically present full and cumulated posterior distributions (displaying the probabilities of all possible effect sizes) for the treatment effects.

Weakly informative priors including all plausible effect sizes and centred on no difference were used for all parameters in the primary Bayesian analyses; these priors primarily served to stabilise computations with minimal influence on the results. Sensitivity analyses were conducted using sceptic priors for the intervention effects; these priors are sceptic of large differences and shrink effect estimates towards no difference, as many interventions assessed in critically ill patients have shown small or uncertain effects [17]. We did not conduct sensitivity analyses using evidence-based priors due to the lack of available external data. Exact priors are specified in the ESM and elsewhere [10]; briefly, the primary priors corresponded to probability distributions for the adjusted incidence rate ratios (IRRs)/odds ratios for the intervention effect centred on 1.00 with 95% central probability mass between 0.14 and 7.10 in each (sub)model [10].

Analysis of the primary outcome

We expected substantial zero-inflation of the primary outcome, which we analysed using an adjusted hurdle-negative binomial model [10]. This two-part model consists of a logistic regression estimating the probability of exactly zero days and a zero-truncated negative binomial model [18] estimating the mean number of days in patients with >0 days. This model has conceptual similarities with the Kryger Jensen and Lange test used in the primary frequentist analyses of the trial [7, 19]. We estimated the adjusted mean number of days alive without life support in each group, the adjusted mean difference (MD) and adjusted IRR, and pre-defined a clinically important difference as an absolute $MD \geq 1$ day [10].

Analysis of the secondary outcomes

The binary secondary outcomes were analysed using adjusted logistic regression models, with results presented as adjusted probabilities in each group, adjusted relative risks (RRs) and adjusted risk differences (RDs). We pre-defined an absolute RD of ≥ 2 percentage points as clinically important for all binary outcomes [10]. The secondary count outcomes were analysed similarly to the primary outcome, with an absolute $MD \geq 1$ day pre-defined as clinically important [10].

Missing data handling

We planned to use multiple imputation if $\geq 5\%$ had missing data for variables included in an analysis [10]. As missingness was lower than this threshold for all analyses, we conducted complete case analyses, with additional best/worst-worst/best-case sensitivity analyses of the primary outcome (assuming all patients in the 12 mg group were alive without life support and that all patients in the 6 mg group were not alive without life support on all days without available data, and vice versa) [7].

Post hoc analyses

Our count outcome definitions used the actual number of days; however, non-survivors have commonly been assigned zero days in other trials [20–22], and thus we conducted additional post hoc analyses assigning non-survivors zero days to ease comparison. Moreover, we conducted best/worst-worst/best-case sensitivity analyses of the primary outcome to match the primary report; these were not detailed in the Bayesian analysis protocol [7, 10]. Last, while we expected substantial zero-inflation for all count outcomes, we expected limited inflation at the maximum values and expected the planned model to adequately fit the data [10]. Unexpectedly, 41.4% and 40.6% of patients had the maximum values of days alive without life support at day 28 and 90, respectively [7]. To account for the unexpected distributions, we conducted additional post hoc sensitivity analyses of both these outcomes using adjusted Bayesian linear models with weakly informative priors and a Bayesian bootstrap procedure using adjusted, conventional linear regressions, and derived all effect measures from the predicted, adjusted means in each group (ESM). This was not an issue for days alive and out of hospital at day 90 (only 1 patient had the maximum possible value of 89 days).

Results

In total, 982 patients were included in the intention-to-treat population, with 497 assigned to 12 mg dexamethasone and 485 assigned to 6 mg dexamethasone. Detailed data on baseline characteristics, protocol adherence, treatment durations, and individual components of the outcomes have been presented elsewhere [7]; data for the adjustment variables and descriptive data for the outcomes analysed here (including missingness) are presented in Tables 1 and S1 (ESM).

Primary outcome

For days alive without life support at day 28, the adjusted MD was 1.3 days (95% CrI -0.3 to 2.9), corresponding to an adjusted IRR of 1.08 (95% CrI 0.98–1.18), and probabilities of any benefit and clinically important benefit with 12 mg dexamethasone of 94.2

and 63.9%, respectively (Table 2). Full posterior distributions and probabilities of all possible effect sizes are presented in Fig. 1. In the sensitivity analysis using a sceptic prior (Table 2), the probabilities of any benefit and clinically important benefit with 12 mg dexamethasone were 94.9 and 61.8%, respectively. Similar results were found in the post hoc sensitivity analyses (Table S2, ESM), with probabilities of benefit with 12 mg dexamethasone of $>90\%$ in all analyses for the average treatment effects. The direction of conditional effects was similar with largely consistent relative effects across all reference patients and somewhat larger variation in absolute effects; probabilities of benefit with 12 mg dexamethasone varied more due to increased uncertainty for one reference patient (Table S3, ESM). The probabilities of clinically important harm were 0.3 and 0.2% in the primary analyses and the analysis using a sceptic prior, respectively, and $\leq 4.5\%$ in all analyses for the average effect and for all reference patients (Table 2, Tables S2–S3, ESM).

Secondary outcomes

Results for the effects of 12 mg versus 6 mg dexamethasone on all secondary outcomes are presented in Table 2, Figs. 2–3, Table S4 and Figs. S1–S3 (ESM). For serious adverse reactions at day 28, the adjusted RR was 0.85 (95% CrI 0.63–1.16), with probabilities of any benefit, clinically important harm, and no clinically important difference of 84.1, 2.1, and 49.5%, respectively. For mortality at day 28, the adjusted RR was 0.87 (95% CrI 0.73–1.03), with probabilities of any benefit, clinically important benefit, and clinically important harm of 94.8, 80.7, and 0.9%, respectively. For mortality at day 90, the adjusted RR was 0.88 (95% CrI 0.75–1.02), with similar probabilities of different effect sizes. For days alive without life support at day 90, the uncertainty was somewhat larger with an adjusted MD of 3.6 days (95% CrI -3.1 to 10.2) and probabilities of any benefit and clinically important harm of 85.0 and 9.2%, respectively. For days alive and out of hospital at day 90, the adjusted MD was 3.9 days (95% CrI -0.6 to 8.4), with probabilities of any benefit, clinically important benefit, and clinically important harm of 95.7, 89.7, and 1.5%, respectively.

The relative effect estimates were largely consistent with sceptic priors, in the post hoc analyses of days alive without life support at day 90, and for conditional effects in different reference patients, with variations in absolute effects and larger uncertainty for some outcome/reference patient combinations. Probabilities of clinically important harm were very or relatively low for all analyses of all secondary outcomes on average and for the different reference patients, with somewhat larger uncertainty for the day 90 count outcomes.

Table 1 Data on baseline variables, stratification variables, and outcomes

Variable	Dexamethasone 12 mg (n = 497)	Dexamethasone 6 mg (n = 485)
Baseline data/stratification variables		
Country		
Denmark	251 (50.5%)	234 (48.2%)
India ^a	182 (36.6%)	187 (38.6%)
Sweden	40 (8%)	39 (8%)
Switzerland	24 (4.8%)	25 (5.2%)
Age, years	65 (56–74)	64 (54–72)
Age below 70 years	330 (66.4%)	318 (65.6%)
Invasive mechanical ventilation at baseline	107 (21.5%)	99 (20.4%)
Descriptive outcome data^b		
Days alive without life support at day 28	22 (6–28)	20.5 (4–28)
Mortality at day 28	133 (27.1%)	155 (32.3%)
One or more serious adverse reactions at day 28	56 (11.3%)	65 (13.4%)
Mortality at day 90	157 (32%)	180 (37.7%)
Days alive without life support at day 90	84 (9.2–90)	80 (6–90)
Days alive and out of hospital at day 90	61.5 (0–78)	48 (0–76)

Categorical variables are presented as numbers (percentages) and numerical variables are presented as medians (interquartile ranges). Additional baseline and outcome data are presented in the primary trial publication [7], and data for the sensitivity analyses (best/worst-worst/best case of the primary outcome and assignment of 0 days to non-survivors for all outcomes) are presented in Table S1 in the electronic supplementary material (ESM)

^a Individual sites are not listed, but the site enrolling most patients was in India

^b Data for days alive without life support and mortality at day 28 were missing in 11 patients (6 patients in the dexamethasone 12 mg group and 5 patients in the dexamethasone 6 mg group). Data for serious adverse reactions at day 28 were analysed for all available days only, and thus there is no missing data for this outcome. Data for days mortality and days alive and out of hospital at day 90 were missing in 14 patients (7 in each group). Data for days alive without life support at day 90 were missing in 15 patients (7 patients in the dexamethasone 12 mg group and 8 patients in the dexamethasone 6 mg group). No data were missing for the other variables presented in this table

Discussion

In this pre-planned secondary Bayesian analysis of the COVID STEROID 2 trial, we found high posterior probabilities of benefit with a higher daily dose (12 mg) of dexamethasone than the currently recommended dose (6 mg) on days alive without life support at day 28 (94.2%) and all secondary outcomes (94–96% for mortality at day 28 and day 90 and days alive and out of hospital at day 90; 84–85% for serious adverse reactions at day 28 and days alive without life support at day 90). Similarly, we found relatively low probabilities of clinically important harm for all outcomes, with probabilities $\leq 2.1\%$ for the average effects on serious adverse reactions at day 28 (a composite of septic shock, clinically important gastrointestinal bleeding, anaphylaxis, and invasive fungal infections), mortality at day 28 and 90, and days alive and out of hospital at day 90. Results were largely consistent across different planned and post hoc sensitivity analyses conducted and largely similar for different reference patients, although uncertainty was somewhat larger in some reference patients and absolute effects varied due to different control group counts and event rates.

While the primary frequentist analyses of the COVID STEROID 2 trial did not reach the pre-defined thresholds for statistical significance [7], absence of statistical significance is not evidence of absence of an effect [23]. The results from the primary analyses were mostly compatible with benefits with 12 mg dexamethasone, and this is further supported by these pre-planned Bayesian analyses. Arguments have been made to entirely abandon the concept of statistical significance [9], and to embrace uncertainty and interpret results in a more nuanced manner, regardless of the analytical approach used [24]. Similarly, it has been argued that while strong evidence should be provided before implementing new, costly, and potentially burdensome interventions, practice changes for well-known treatments already in widespread use may require less firm evidence [8]. This is arguably the case for two different doses of a well-known and inexpensive drug like dexamethasone when the evidence for all outcomes seems to favour one dose with relatively high probabilities. Importantly, longer-term outcomes from the COVID STEROID 2 trial will be reported after 180 days follow-up [5]. Furthermore,

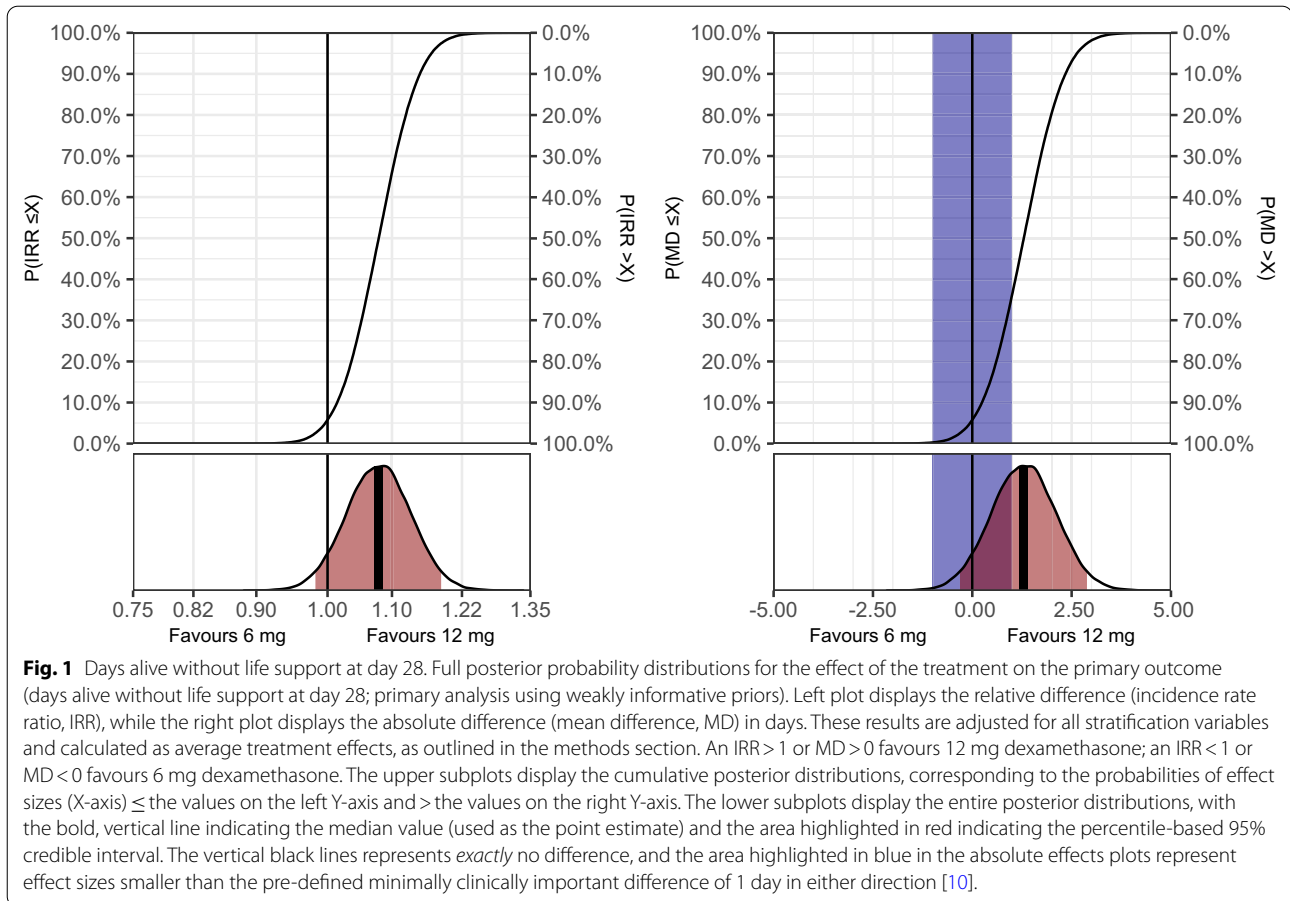
Table 2 Average treatment effect estimates and probabilities of effects

Outcome	Effect estimates				Probability of effects with 12 mg dexamethasone				
	Dexamethasone 12 mg	Dexamethasone 6 mg	Relative difference	Absolute difference	Any benefit	Any harm	Clinically important benefit	Clinically important harm	No clinically important difference
Primary analyses using weakly informative priors									
Days alive without life support at day 28	Mean: 17.8 (16.7–19) days	Mean: 16.5 (15.4–17.6) days	IRR: 1.08 (0.98–1.18)	MD: 1.3 (–0.3 to 2.9) days	94.2%	5.8%	63.9%	0.3%	35.9%
Serious adverse reactions at day 28	Prob.: 11.4% (8.9–14.1%)	Prob.: 13.3% (10.7–16.2%)	RR: 0.85 (0.63–1.16)	RD: –1.9% (–5.7 to 1.9%)	84.1%	15.9%	48.4%	2.1%	49.5%
Mortality at day 28	Prob.: 27.5% (24–31.2%)	Prob.: 31.8% (28.1–35.6%)	RR: 0.87 (0.73–1.03)	RD: –4.3% (–9.4 to 0.9%)	94.8%	5.2%	80.7%	0.9%	18.5%
Mortality at day 90	Prob.: 32.5% (28.8–36.3%)	Prob.: 37.1% (33.2–41.0%)	RR: 0.88 (0.75–1.02)	RD: –4.6% (–10 to 0.9%)	95.1%	4.9%	82.3%	0.8%	16.9%
Days alive without life support at day 90	Mean: 59.3 (54.6–64.2) days	Mean: 55.7 (51.1–60.6) days	IRR: 1.06 (0.95–1.2)	MD: 3.6 (–3.1 to 10.2) days	85%	15%	77.2%	9.2%	13.6%
Days alive and out of hospital at day 90	Mean: 44.1 (40.9–47.3) days	Mean: 40.2 (37–43.5) days	IRR: 1.10 (0.99–1.22)	MD: 3.9 (–0.6 to 8.4) days	95.7%	4.3%	89.7%	1.5%	8.8%
Pre-specified sensitivity analyses using sceptic priors									
Days alive without life support at day 28	Mean: 17.8 (16.7–18.9) days	Mean: 16.5 (15.5–17.6) days	IRR: 1.07 (0.99–1.17)	MD: 1.2 (–0.2 to 2.7) days	94.9%	5.1%	61.8%	0.2%	38%
Serious adverse reactions at day 28	Prob.: 12% (9.9–14.3%)	Prob.: 12.7% (10.5–15.1%)	RR: 0.95 (0.79–1.14)	RD: –0.7% (–3 to 1.6%)	72.3%	27.7%	13.2%	1%	85.8%
Mortality at day 28	Prob.: 28.6% (25.5% to 31.8%)	Prob.: 30.7% (27.6% to 34%)	RR: 0.93 (0.82–1.05)	RD: –2.1% (–5.7 to 1.5%)	87.2%	12.8%	52.3%	1.2%	46.4%
Mortality at day 90	Prob.: 33.6% (30.3–37%)	Prob.: 36% (32.6–39.4%)	RR: 0.93 (0.83–1.05)	RD: –2.4% (–6.3 to 1.5%)	88.3%	11.7%	57.3%	1.4%	41.3%
Days alive without life support at day 90	Mean: 59 (54.6–63.7) days	Mean: 56 (51.6–60.7) days	IRR: 1.05 (0.95–1.17)	MD: 3 (–3 to 9.1) days	83.6%	16.4%	74.4%	9.7%	16%
Days alive and out of hospital at day 90	Mean: 43.1 (40.3–46.1) days	Mean: 41.1 (38.2–44.1) days	IRR: 1.05 (0.96–1.14)	MD: 2 (–1.6 to 5.7) days	86.2%	13.8%	71.1%	5.1%	23.8%

Analyses conducted in all patients with available outcome data (Table 1). All analyses were adjusted for the stratification variables, and effect sizes are presented as average treatment effects as outlined in the methods section, summarised using median posterior values as point estimates and percentile-based 95% credible intervals (CrIs). Data from additional post hoc sensitivity analyses and results estimated for reference patients are presented in Tables S2–S7 in the ESM

Any benefit is the probability of a MD > 0 days (IRR > 1) or a RD < 0 percentage points (RR < 1); any harm is the probability of a MD < 0 days (IRR < 1) or a RD > 0 percentage points (RR > 1); no clinically important difference is the probability of an absolute MD < 1 days or an absolute RD < 2 percentage points; clinically important benefit/harm are probabilities of effect sizes larger than no clinically important difference in either direction. All definitions of clinically important effect sizes were pre-specified in the protocol [10]

ESM electronic supplementary material, IRR incidence rate ratio (> 1 favours 12 mg); MD mean difference (> 0 favours 12 mg), prob. Probability, RD risk difference in percentage points (< 0 favours 12 mg), RR relative risk (< 1 favours 12 mg)

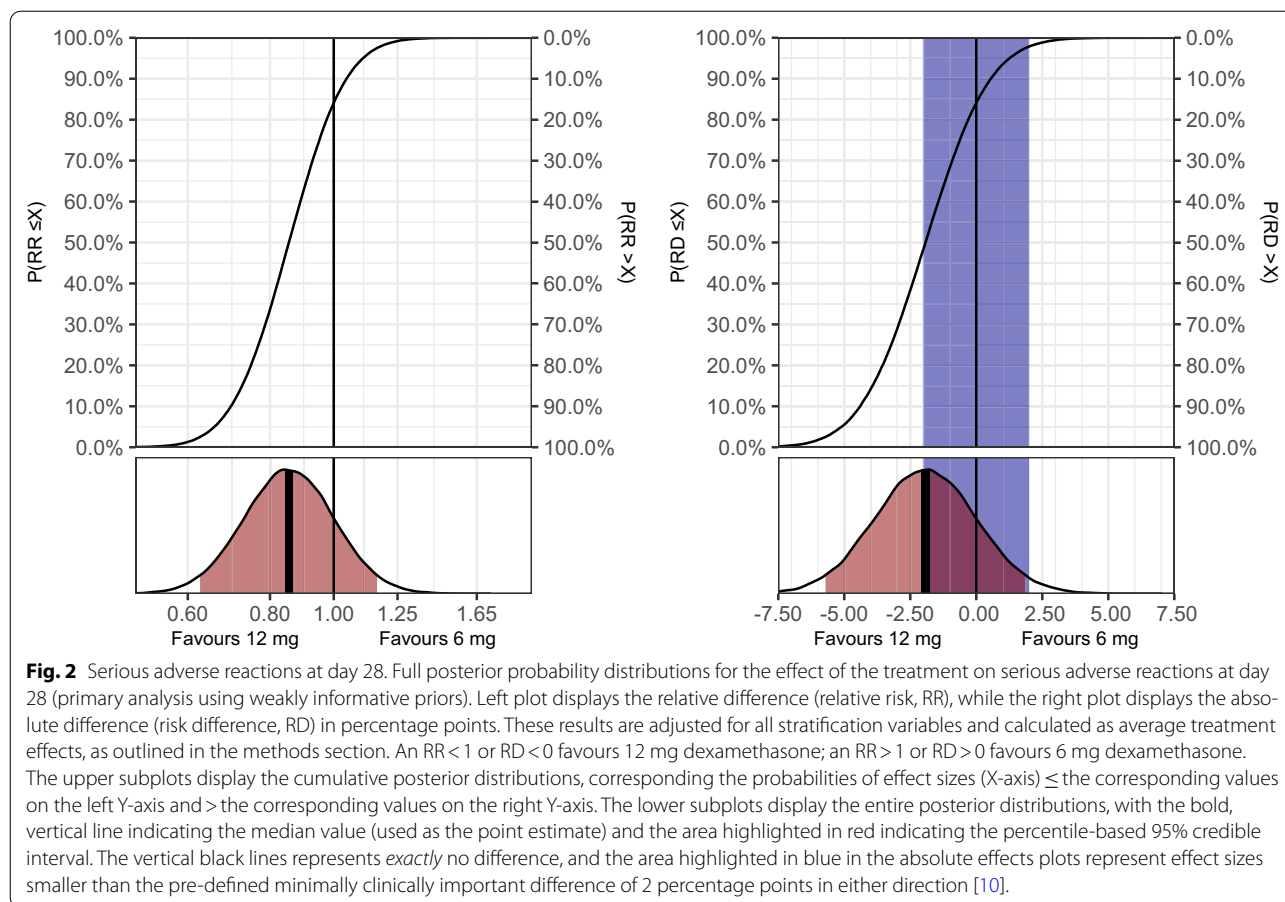


several other randomised clinical trials are currently comparing different doses of dexamethasone; results from these trials and the COVID STEROID 2 trial will be pooled in a prospective meta-analysis [25], and thus, the overall evidence base is expected to improve in the coming months. Until then, considering 12 mg instead of 6 mg dexamethasone daily in patients with COVID-19 and severe hypoxaemia seems reasonable given the current evidence, although potential interaction with other treatments (e.g., interleukin-6 receptor antagonists) remains unresolved.

Strengths and limitations

This study has several strengths. First, the general strengths of the COVID STEROID 2 trial, including the sample size, blinding, international recruitment, and high inclusion rate, also apply to this secondary analysis [7]. Second, this study was pre-planned, and the statistical analysis plan was published prior to enrolment of the last patient and before any data were analysed [10]. Third, results were consistent across the different sensitivity analyses conducted.

This study comes with limitations, too. First, the distributions of days alive without life support after 28 and 90 days were different than expected, and thus the planned model did not fit the data as well as expected. However, this did not appear to influence the conclusions, as similar results were found in the post hoc analyses using different approaches. Second, we deviated from the pre-specified analysis plan by primarily presenting average treatment effects instead of conditional treatment effects. This was considered necessary due the low event rates (primarily for serious adverse reactions, with overall few events/centre) and possible differences in distributions across sites, which made the results difficult to interpret for the planned reference patient approach and posed a risk of underestimating serious adverse reactions in either group for the full trial. Of note, the underlying statistical models were unchanged, and conditional effects for the planned reference patients and additional other reference patients were also presented. Overall, the results were largely consistent, especially for the relative effect measures. The absolute effect measures (and thus the interpretations of clinical importance in different

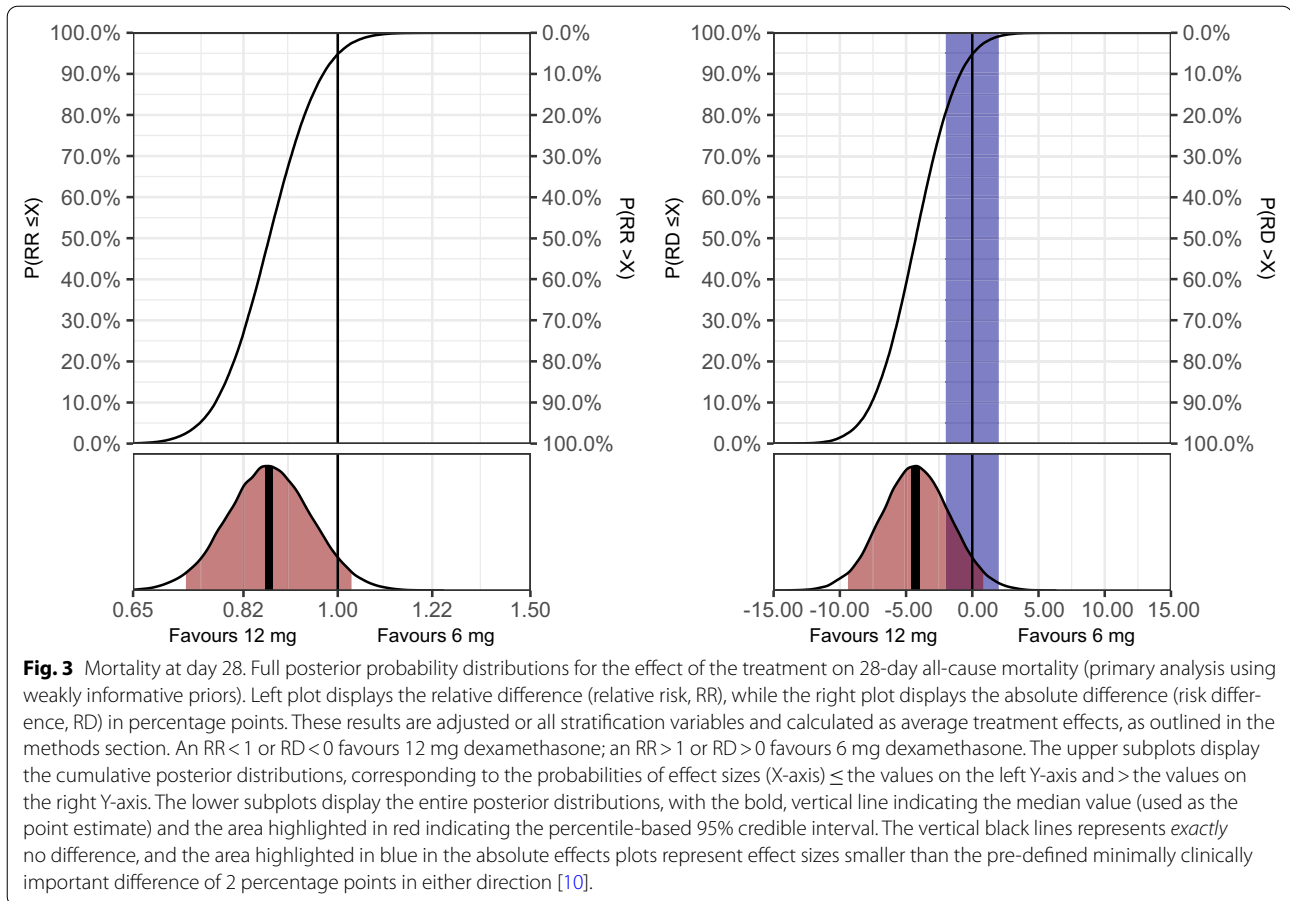


reference patients) may have varied, as is always the case when relative effects are consistent and baseline risk varies [26]. Third, general limitations of the COVID STEROID 2 trial, including changes in the standard of care during the trial (e.g., possibly increased use of interleukin-6-receptor antagonists [27], which may affect the effects of different doses of systemic corticosteroids) and no data on baseline inflammatory status also applies to this secondary analysis [7]. Finally, although pre-defined [10], our definitions of clinically important effect sizes may be challenged. For all count outcomes, we considered an absolute difference of at least 1 day to be clinically important, as shorter periods will have small implications for patients and for health care capacities. For mortality and serious adverse reactions, we pragmatically considered absolute risk differences of at least 2 percentage

points as important, as even relatively small absolute differences may have important implications due to the large number of patients affected by the pandemic. However, we recognize that other reasonable thresholds could have been chosen instead.

Conclusion

In conclusion, we found high probabilities of benefit and relatively low probabilities of clinically important harm with dexamethasone 12 versus 6 mg daily for up to 10 days in patients with COVID-19 and severe hypoxaemia on all outcomes assessed within 90 days of randomisation.



Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1007/s00134-021-06573-1>.

Author details

¹ Department of Intensive Care, Rigshospitalet—Copenhagen University Hospital, Blegdamsvej 9, 2100 Copenhagen, Denmark. ² Collaboration for Research in Intensive Care (CRIC), Copenhagen, Denmark. ³ Department of Anaesthesia, Critical Care and Pain, Tata Memorial Hospital, Homi Bhabha National Institute, Mumbai, India. ⁴ Department of Critical Care, Apollo Hospitals, Chennai, India. ⁵ Chennai Critical Care Consultants, Chennai, India. ⁶ The George Institute for Global Health, New Delhi, India. ⁷ Department of Clinical Science and Education, Södersjukhuset, Karolinska Institutet, Stockholm, Sweden. ⁸ Department of Intensive Care Medicine, Inselspital, Bern University Hospital, University of Bern, Bern, Switzerland. ⁹ Department of Infectious Diseases, Rigshospitalet, University of Copenhagen, Copenhagen, Denmark. ¹⁰ Center of Research & Disruption of Infectious Diseases, Department of Infectious Diseases, Copenhagen University Hospital—Amager and Hvidovre, Hvidovre, Denmark. ¹¹ The George Institute for Global Health, University of New South Wales, Sydney, Australia. ¹² Malcolm Fisher Department of Intensive Care, Royal North Shore Hospital, St Leonards, NSW, Australia. ¹³ Prasanna School of Public Health, Manipal Academy of Higher Education, Manipal, India. ¹⁴ School of Public Health, Imperial College London, London, UK. ¹⁵ Department of Anaesthesia and Intensive Care, Hvidovre Hospital, University of Copenhagen, Hvidovre, Denmark. ¹⁶ Department of Respiratory Medicine, Hvidovre Hospital, University of Copenhagen, Copenhagen, Denmark. ¹⁷ Department of Thoracic Anaesthesiology, Rigshospitalet, University of Copenhagen, Copenhagen, Denmark. ¹⁸ Department of Neuroanaesthesiology, Rigshospitalet, University of Copenhagen, Copenhagen, Denmark.

¹⁹ Department of Anaesthesia and Intensive Care, Copenhagen University Hospital – North Zealand, Hillerød, Denmark. ²⁰ Department of Clinical Medicine, University of Copenhagen, Copenhagen, Denmark. ²¹ Department of Anaesthesia and Intensive Care, Herlev Hospital, University of Copenhagen, Herlev, Denmark. ²² Department of Anaesthesiology, Zealand University Hospital, Køge, Denmark. ²³ Department of Anaesthesia and Intensive Care, Aalborg University Hospital, Aalborg, Denmark. ²⁴ Department of Anaesthesia and Intensive Care, Kolding Hospital, University Hospital of Southern Denmark, Kolding, Denmark. ²⁵ Department of Anaesthesia and Critical Care Medicine, Odense University Hospital, Odense C, Denmark. ²⁶ Department of Anaesthesia and Critical Care Medicine, Hospital Sønderjylland, University Hospital of Southern, Odense, Denmark. ²⁷ Department of Anaesthesia and Intensive Care, Næstved-Slagelse-Ringsted Hospital, Slagelse, Denmark. ²⁸ Department of Critical Care Medicine, Rajendra Institute of Medical Sciences, Ranchi, India. ²⁹ Department of Intensive Care, SL Raheja Hospital, Mumbai, Maharashtra, India. ³⁰ Department of Intensive Care, Vishwaraj Hospital, Pune, India. ³¹ Section of Critical Care, Department of Medicine, Hinduja Hospital, Mahim, Mumbai, India. ³² Department of Critical Care Medicine, Sanjeevan Hospital, Pune, Maharashtra, India. ³³ Department of Respiratory and Critical Care Medicine, Indraprastha Apollo Hospital, New Delhi, India. ³⁴ Intensive Care Unit and Emergency Services, Symbiosis University Hospital and Research Centre, Lavale, Pune, India. ³⁵ Department of Critical Care Medicine, Bombay Hospital Institute of Medical Sciences, Mumbai, India. ³⁶ Department of Anesthesiology and Intensive Care, Biomedical and Clinical Sciences, Linköping University, Linköping, Sweden. ³⁷ Department of Anaesthesia and Intensive Care, Bispebjerg Hospital, Copenhagen, Denmark. ³⁸ Centre for Clinical Intervention Research, Copenhagen Trial Unit, Capital Region of Denmark, Copenhagen University Hospital – Rigshospitalet, Copenhagen, Denmark. ³⁹ Department of Regional Health Research, Faculty of Health Sciences, University of Southern

Denmark, Odense, Denmark. ⁴⁰ Department of Public Health, Section of Biostatistics, University of Copenhagen, Copenhagen, Denmark.

Acknowledgements

The authors thank everyone involved in the COVID STEROID 2 trial (patients and relatives, research staff, clinical staff, investigators, funding sources, and regulatory authorities), and everyone involved in the first COVID STEROID trial [28], which served as the foundation for the COVID STEROID 2 trial.

Authors' contributions

AG conducted all analyses presented in this manuscript and wrote the first draft, which was critically revised by all authors. All authors contributed to the design and/or conduct of the COVID STEROID 2 trial, including the present secondary study. Detailed author contributions for the complete trial are presented in the primary trial report [7].

Funding

The COVID STEROID 2 trial was funded by the Novo Nordisk Foundation (grant number 0062998) and supported by Rigshospitalet's Research Council (E-22703-06). None of the funders had any influence on trial design, conduct, analysis, or decision to report. AG's salary is paid by a grant from Sygeforsikringen "danmark", which had no influence on this trial.

Availability of data and material

Fully de-identified individual participant data will be shared after approval by the trial Management Committee after reasonable request to the trial Management Committee.

Code availability

Analysis code is available from the corresponding author upon reasonable request.

Declarations

Conflicts of interest

AG, MWM, MBNK, GKV, TSM, MHM and AP are affiliated with the Department of Intensive Care at Rigshospitalet, University of Copenhagen, which has received grants from the Novo Nordisk Foundation during the conduct of the trial; and grants from Pfizer, Fresenius Kabi, The Novo Nordisk Foundation, and Sygeforsikringen "danmark" outside the submitted work. SMJ and LC are affiliated with Inselspital, Bern University Hospital, University of Bern, Bern, Switzerland, which has received grants from Edwards Lifesciences Services GmbH, Phagenesis Limited, and Nestlé; all outside the submitted work. TB has received grants from the Novo Nordisk Foundation, Simonsen Foundation, Lundbeck Foundation, Kai Hansen Foundation, and Erik and Susanna Olesen's Charitable Fund; grants and personal fees from GSK, Pfizer, and Gilead; personal fees from Boehringer Ingelheim, MSD, and Pentabase ApS; all outside the submitted work. BV and NEH have received grants from Baxter outside the submitted work. VJ has received grants and personal fees from Baxter Healthcare; personal fees from Astra Zeneca, Visterra, Chinook, and NephroPlus; all outside the submitted work. CSU has received grants and personal fees from GSK, Sanofi Genzyme, Astra Zeneca, Teva, OrionPharma, Boehringer-Ingelheim, Actelion and Chiesi; grants, personal fees and non-financial support from Novartis; all outside the submitted work. JVD has received personal fees (paid to his institution) from Edwards India outside the submitted work. PA has received personal fees from CIPLA, Dr. Reddy's Laboratories, Abbott Nutrition and Sanofi; all outside the submitted work. The remaining authors declared no known conflicts of interest.

Ethics approval

Approved by the Ethics Committee of the Capital Region of Denmark with additional national/local approvals as required; additional details on approvals are available elsewhere [5, 7, 10].

Consent to participate

Consent was obtained from patients or legal surrogates according to applicable laws and regulations. Enrolment according to an emergency procedure (i.e., consent from a doctor independent by the trial followed by consent from relatives/patients) was allowed at many sites; additional details are available elsewhere [5, 7, 10].

Consent for publication

All authors approved the manuscript for publication.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Received: 27 September 2021 Accepted: 29 October 2021

Published online: 10 November 2021

References

- Haase N, Plovsing R, Christensen S et al (2021) Characteristics, interventions and longer-term outcomes of COVID-19 ICU patients in Denmark – a nationwide, observational study. *Acta Anaesthesiol Scand* 65:68–75
- World Health Organization (2021) Coronavirus disease (COVID-19) Pandemic. <https://www.who.int/emergencies/diseases/novel-coronavirus-2019>. Accessed 27 Sep 2021.
- The WHO Rapid Evidence Appraisal for COVID-19 Therapies (REACT) Working Group (2020) Association between administration of systemic corticosteroids and mortality among critically ill patients with COVID-19: a meta-analysis. *JAMA* 324:1330–1341
- Rochwerf B, Agarwal A, Siemieniuk RA et al (2020) A living WHO guideline on drugs for covid-19. *BMJ* 370:m3379
- Munch MW, Granholm A, Myatra SN et al (2021) Higher vs. lower doses of dexamethasone in patients with COVID-19 and severe hypoxia (COVID STEROID 2) trial: protocol and statistical analysis plan. *Acta Anaesthesiol Scand* 65:834–845
- Villar J, Ferrando C, Martínez D et al (2020) Dexamethasone treatment for the acute respiratory distress syndrome: a multicentre, randomised controlled trial. *Lancet Respir Med* 8:267–276
- The COVID STEROID 2 Trial Group (2021) Effect of 12 mg vs 6 mg of dexamethasone on the number of days alive without life support in adults with COVID-19 and severe hypoxemia: the COVID STEROID 2 randomized trial. *JAMA*. <https://doi.org/10.1001/jama.2021.18295>
- Young PJ, Nickson CP, Perner A (2020) When should clinicians act on non-statistically significant results from clinical trials? *JAMA* 323:2256–2257
- Amrhein V, Greenland S, McShane B (2019) Scientists rise up against statistical significance. *Nature* 567:305–307
- Granholm A, Munch MW, Myatra SN et al (2021) Higher vs. lower doses of dexamethasone in patients with COVID-19 and severe hypoxia (COVID STEROID 2) trial: protocol for a secondary Bayesian analysis. *Acta Anaesthesiol Scand* 65:702–710
- von Elm E, Altman DG, Egger M, Pocock SJ, Gøtzsche PC, Vandenbroucke JP (2008) The Strengthening of Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. *J Clin Epidemiol* 61:344–349
- Sung L, Hayden J, Greenberg ML, Koren G, Feldman BM, Tomlinson GA (2005) Seven items were identified for inclusion when reporting a Bayesian analysis of a clinical study. *J Clin Epidemiol* 58:261–268
- Wickham H, Averick M, Bryan J et al (2019) Welcome to the Tidyverse. *J Open Source Softw* 4:1686
- Carpenter B, Guo J, Hoffman MD et al (2017) Stan: a probabilistic programming language. *J Stat Softw*. <https://doi.org/10.18637/jss.v076.i01>
- Bürkner P-C (2017) brms: an R Package for Bayesian Multilevel models using stan. *J Stat Softw*. <https://doi.org/10.18637/jss.v080.i01>
- Bendtsen M (2018) A gentle introduction to the comparison between null hypothesis testing and Bayesian analysis: reanalysis of two randomized controlled trials. *J Med Internet Res* 20:e10873
- Santacruz CA, Pereira AJ, Celis E, Vincent JL (2019) Which multicenter randomized controlled trials in critical care medicine have shown reduced mortality? A systematic review. *Crit Care Med* 47:1680–1691
- Zeileis A, Kleiber C, Jackman S (2008) Regression models for count data in R. *J Stat Softw*. <https://doi.org/10.18637/jss.v027.i08>
- Jensen AK, Lange T (2019) A novel high-power test for continuous outcomes truncated by death. *ArXiv* 1910.12267 [stat.ME].
- Contentin L, Ehrmann S, Giraudeau B (2014) Heterogeneity in the definition of mechanical ventilation duration and ventilator-free days. *Am J Respir Crit Care Med* 189:998–1002

-
21. Yehya N, Harhay MO, Curley MAQ, Schoenfeld DA, Reeder RW (2019) Reappraisal of ventilator-free days in critical care research. *Am J Respir Crit Care Med* 200:828–836
 22. Writing Committee for the REMAP-CAP Investigators (2020) Effect of hydrocortisone on mortality and organ support in patients with severe COVID-19: the REMAP-CAP COVID-19 corticosteroid domain randomized clinical trial. *JAMA* 324:1317–1329
 23. Altman DG, Bland JM (1995) Statistics notes: absence of evidence is not evidence of absence. *BMJ* 311:485
 24. Sterne JAC, Smith GD (2001) Sifting the evidence—what’s wrong with significance tests? *BMJ* 322:226–231
 25. Møller MH, Munch MW, Granholm A, et al (2021) Higher vs. standard doses of dexamethasone in patients with COVID-19 and hypoxia: a prospective meta-analysis. OSF Registries. <https://doi.org/10.17605/OSF.IO/FR5SV>.
 26. Kent DM, van Klaveren D, Paulus JK et al (2020) The Predictive Approaches to Treatment effect Heterogeneity (PATH) statement: explanation and elaboration. *Ann Intern Med* 172:W1-25
 27. The REMAP-CAP Investigators (2021) Interleukin-6 receptor antagonists in critically ill patients with Covid-19. *N Engl J Med* 384:1491–1502
 28. Munch MW, Meyhoff TS, Helleberg M et al (2021) Low-dose hydrocortisone in patients with COVID-19 and severe hypoxia: the COVID STEROID randomised, placebo-controlled trial. *Acta Anaesthesiol Scand*. <https://doi.org/10.1111/aas.13941>