NEJM Evidence

Published December 14, 2022
NEJM Evid 2022; 2 (1)
DOI: 10.1056/EVIDoa2200196

ORIGINAL ARTICLE

A Bayesian Interpretation of a Pediatric Cardiac Arrest Trial (THAPCA-OH)

Michael O. Harhay, Ph.D.,[1,2] Bryan S. Blette, Ph.D.,[1,2] Anders Granholm, M.D.,[3] Frank W. Moler, M.D., M.S.,[4] Fernando G. Zampieri, M.D., Ph.D.,[5,6] Ewan C. Goligher, M.D., Ph.D.,[7,8,9] Monique M. Gardner, M.D.,[10] Alexis A. Topjian, M.D., M.S.C.E.,[10] and Nadir Yehya, M.D., M.S.C.E.[10]

Abstract

BACKGROUND Pediatric out-of-hospital cardiac arrest results in high morbidity and mortality. Currently, there are no recommended therapies beyond supportive care. The THAPCA-OH (Therapeutic Hypothermia after Pediatric Cardiac Arrest Out-of-Hospital) trial compared hypothermia (33.0°C) with normothermia (36.8°C) in 295 children. Good neurobehavioral outcome and survival at 1 year were higher in the hypothermia group (20 vs. 12% and 38 vs. 29%, respectively). These differences did not meet the planned statistical threshold of $P<0.05$. To ensure that a potentially efficacious therapy is not prematurely discarded, we reassessed THAPCA-OH using a Bayesian statistical perspective.

METHODS We performed a Bayesian analysis, interpreting the trial in probabilistic terms (i.e., the probability that therapeutic hypothermia had any benefit, and overall absolute improvements greater than 2%, 5%, and 10% for 1-year neurobehavioral outcome and survival). Our primary analyses used noninformative priors, meaning that the analyses were based on the observed trial data without any information added by the priors. In the absence of pediatric trials to derive informative prior distributions, we used: (1) downweighted priors from adult trials; and (2) a previously published set of critical care priors that span benefit, equipoise, and harm.

RESULTS In the primary analyses, the probability of any benefit from hypothermia was 94% for both the neurobehavioral outcome and survival at 1 year. For both outcomes, the probability of benefit was >75% for all informative prior integrations with the THAPCA-OH results, except those with the most pessimistic priors.

CONCLUSIONS There is a high probability that hypothermia provides a modest benefit in neurobehavioral outcome and survival at 1 year. (ClinicalTrials.gov number, NCT00878644.)

The author affiliations are listed at the end of the article.

Dr. Harhay can be contacted at mharhay@pennmedicine.upenn.edu or at Perelman School of Medicine at the University of Pennsylvania, 304 Blockley Hall, 423 Guardian Dr., Philadelphia, PA 19104-6021.

Introduction

Pediatric out-of-hospital cardiac arrest (OHCA) has substantial morbidity and mortality.[1-4] Early trials of therapeutic hypothermia compared with usual care in adult OHCA showed improved outcomes.[5,6] Subsequent adult trials comparing therapeutic hypothermia (33°C) with therapeutic normothermia (37.5°C) or less hypothermia (36°C), such as the TTM (Target Temperature Management 33°C versus 36°C after Out-of-Hospital Cardiac Arrest) trial[7] and TTM2 (Targeted Hypothermia versus Targeted Normothermia after Out-of-Hospital Cardiac Arrest),[8] did not show benefit with hypothermia, whereas the HYPERION (Therapeutic Hypothermia after Cardiac Arrest in Nonshockable Rhythm) trial[9] (33° vs. 37°C) did show benefit. For pediatric OHCA, the THAPCA-OH (Therapeutic Hypothermia after Pediatric Cardiac Arrest Out-of-Hospital[10]) trial is, to the best of our knowledge, the sole randomized trial investigating the impact of therapeutic hypothermia on outcomes in children (ClinicalTrials.gov number, NCT00878644).

THAPCA-OH suggested that children treated with therapeutic hypothermia (33.0°C) may have a higher proportion of good neurobehavioral outcome, defined as a Vineland Adaptive Behavior Scales second edition (VABS-II) score (scores range from 20 to 160, with higher scores indicating better function) ≥70 (20 vs. 12%; P=0.14), and survival at 1 year (38 vs. 29%; P=0.13) than those treated with therapeutic normothermia (36.8°C). However, because THAPCA-OH used a frequentist statistical design powered to detect an absolute improvement of 20% in good neurobehavioral outcome,[10,11] it concluded that therapeutic hypothermia "did not confer a significant benefit in survival with a good functional outcome at 1 year."[10]

Approximately 7,000 children experience OHCA annually in the United States,[3] making recruitment of large sample sizes for trials difficult. A feasibility study from the Paediatric Intensive Care Audit Network of 33 United Kingdom and Republic of Ireland pediatric intensive care units concluded that OHCA interventional trials were infeasible on the basis of low prevalence, despite 50.5% mortality.[12] Until other trials are conducted, clinicians must leverage all available data for decision-making, even when those data are obtained from underpowered randomized controlled trials. Therefore, to aid in the interpretation and application of the THAPCA-OH results for additional research and guideline consideration, we reanalyzed THAPCA-OH using a Bayesian statistical framework to provide a probabilistic interpretation of the efficacy of therapeutic hypothermia in this population.

Methods

This was an unplanned Bayesian analysis of THAPCA-OH.[10] We adhered to the Reporting of Bayes Used in Clinical STudies guideline[13] and followed a standardized Bayesian framework for reanalyzing critical care trials proposed by Zampieri et al.[14] THAPCA-OH trial data were obtained from the National Heart, Lung, and Blood Institute Biologic Specimen and Data Repository Information Coordinating Center.[15]

THAPCA-OH

THAPCA-OH was a multicenter trial (n=295 randomly assigned patients from 38 children's hospitals) comparing the efficacy of therapeutic hypothermia (33.0°C) versus therapeutic normothermia (36.8°C) for 48 hours followed by slow rewarming and normothermia through 120 hours. Comatose children >48 hours and <18 years of age with OHCA requiring ≥2 minutes of cardiopulmonary resuscitation who needed mechanical ventilation after return of circulation were eligible. Major exclusion criteria included inability to randomly assign within 6 hours of return of circulation, Glasgow Coma Scale motor response of 5 or 6, decision to withhold aggressive treatment, or major trauma associated with the arrest. Eligible participants were randomly assigned in a 1:1 ratio using permuted blocks stratified according to clinical center and age (<2, ≥2 to <12, and ≥12 years).

OUTCOMES AND ANALYTIC SAMPLE

In this secondary analysis, we examined the two outcomes from the original trial for which there were reported treatment effect estimates: (1) the primary outcome of survival with a good neurobehavioral outcome at 1 year in evaluable patients (n=260; 25 were excluded for poor prearrest neurobehavioral status defined as a VABS-II score <70 and 10 for missing outcome data); and (2) survival at 1 year (n=287; eight with unknown survival status). Good neurobehavioral outcome was defined as an age-corrected standard score on the VABS-II ≥70 at 1 year of observation among those with prearrest VABS-II scores ≥70.[10,16]

BASICS OF A BAYESIAN REANALYSIS

Bayesian analyses compute the plausible distribution of values for the treatment effects (termed the posterior probability

distribution) given the observed trial results (termed the likelihood function) and, where relevant, prior beliefs about the plausible range of values for the treatment effect (termed priors or prior probability distributions).[17-21] The use of noninformative priors means that the posterior depends on the observed trial data only without any additional information added by the prior. Informative priors may be based on previous empirical data or derived to represent a prespecified range of prior optimism, skepticism, or pessimism about treatment effects. We used all three approaches (referred to hereafter as noninformative, evidence-based, and standardized) to specify priors for this study.

BAYESIAN PRIORS

Because THAPCA-OH is a completed trial, we could not prespecify our analysis and priors before trial results were unblinded and published. Thus, our primary analyses examined the results of the trial using noninformative priors. This allowed us to evaluate THAPCA-OH in terms of the probability of harm and benefit based on the observed trial data without any information added by the priors. Thus, the primary analysis makes no formal prior assumptions about the treatment effect. In the case of THAPCA-OH, there were few previous pediatric studies to guide the design of informative evidence-based priors, except retrospective studies[22,23] with significant heterogeneity and risk of bias. Available adult[5,6] and neonatal[24,25] cardiac arrest trials reflected distinct populations from pediatrics. Likewise, systematic reviews and meta-analyses conducted after THAPCA-OH have yielded inconsistent results.[26-28] Therefore, we took two additional approaches to determining priors for the current study.

First, we derived three evidence-based priors regarding the effect of therapeutic hypothermia based on the following: (1) a recent meta-analysis by Granfeldt et al.[29]; (2) a weighted average of the TTM and TTM2 trials[30]; and (3) the HYPERION trial,[9] which included the empirical data we considered most relevant to THAPCA-OH (Fig. 1A). These are referred to as Granfeldt, TTM, and Hyperion, respectively. Each of these was downweighted by 50% (equivalent to doubling the reported variance) because of clinical differences between adult and pediatric populations. For example, the Granfeldt meta-analysis reported a log relative risk of ln(1.2) with a variance of approximately 0.03, and thus the corresponding prior in our analysis had a mean of ln(1.2) and a variance of approximately 0.06. Only the primary outcome was examined by using evidence-based priors, as effect sizes for survival at 1 year were not reported uniformly in all studies considered.

We next examined an extensive continuum (n=9) of standardized priors previously constructed as part of a Bayesian framework specific to critical care trials (Fig. 1B and Table S1; the Supplementary Appendix is available with the full text of the article at evidence.nejm.org).[14] These standardized priors were designed to cover several potential scenarios, from potential harm (i.e., pessimistic priors) to true equipoise (i.e., neutral priors, including skeptical but neutral priors) to benefit (i.e., optimistic priors). As a result of a lack of pediatric trial data, for each of the beliefs, we examined three varying degrees of prior strength: strong, moderate, and weak.

STATISTICAL ANALYSIS

To align with the original analysis, we analyzed the same population of evaluable participants for both outcomes.[10] The results are presented as absolute and relative benefits, as opposed to absolute and relative risk, to align with the presentation in the original trial, with higher values reflecting benefit. Separate Bayesian regression models were fit for both the neurobehavioral outcome and survival outcome for each (evidence-based and standardized) prior distribution. Two types of models are used throughout this article: (1) evidence-based priors were defined on the log(relative risk) scale and corresponded to treatment effect coefficients in log-binomial models; and (2) standardized priors (in the original publication[14]) were defined on the log(odds ratio) scale and corresponded to treatment effect coefficients in logistic models (Fig. 1). Each model adjusted for the age group randomization stratification factor and marginalized across age groups by using g-computation following the procedure described by Granholm et al.[31] Default weakly informative priors were used for the age group parameter in each model. Regressions were run for eight Markov chains with 5,000 iterations, including 2,500 warm-up iterations. Model convergence was assessed as previously described,[31] with all diagnostics deemed adequate.

Posterior distributions for the absolute risk benefit and relative-risk benefit were sampled and plotted for each of the models. Posterior medians and percentile-based 95% credible intervals were calculated from each posterior distribution. To aid clinical interpretation, we also calculated the probability of a treatment benefit greater than various clinically meaningful thresholds of the absolute and relative scale for each prior. Probability of severe harm was assessed by using a cutoff of −0.05 for the absolute benefit parameter and 1/1.25 for the relative benefit parameter. All analyses were performed in R version 4.1.2 (R Foundation

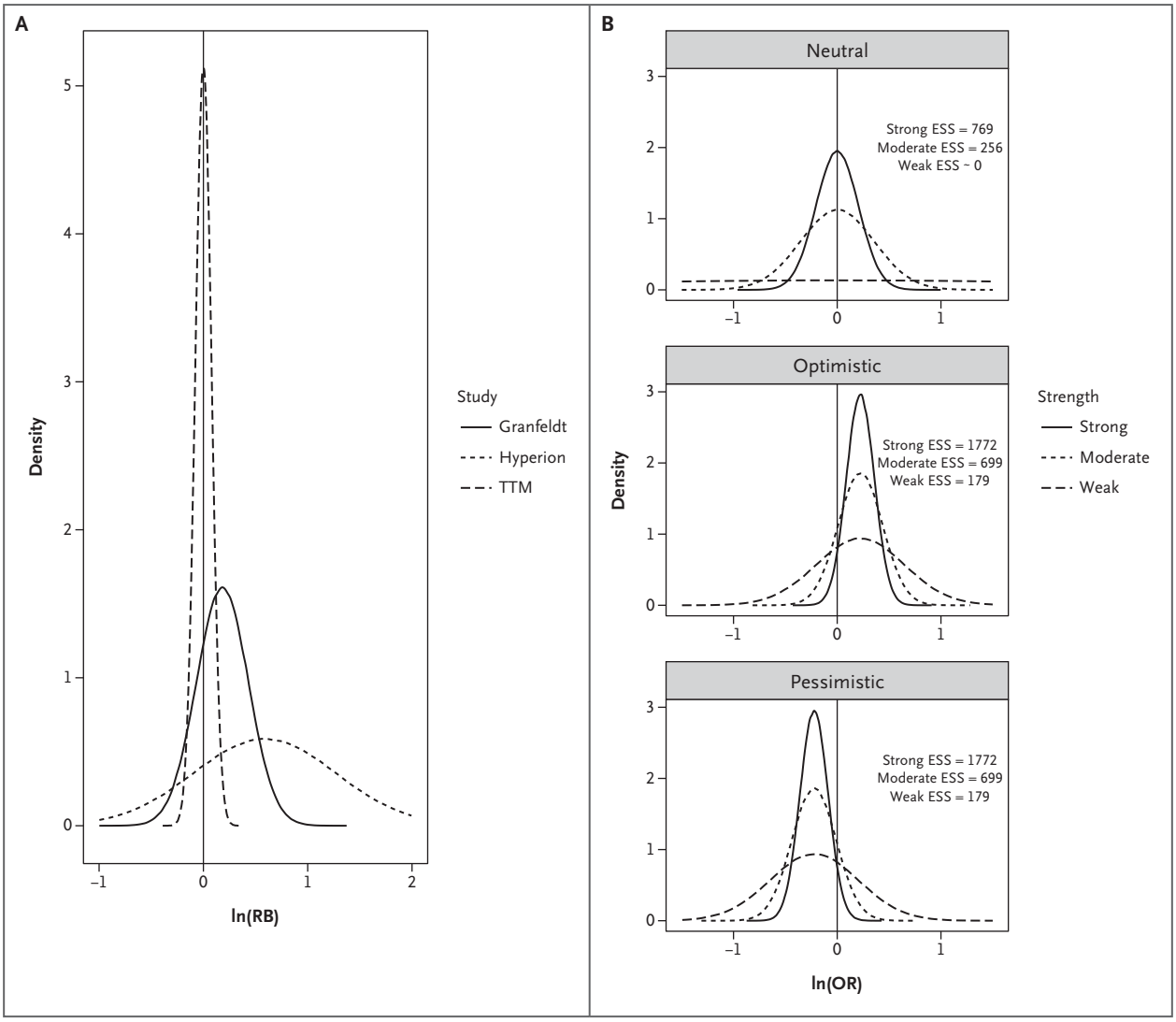

Figure 1. Priors Used for the Bayesian Reanalysis of the THAPCA-OH Trial.[10]

Evidence-based (Panel A) and standardized (Panel B) priors used for the Bayesian reanalysis of the THAPCA-OH Trial.[10] A corresponding summary of the standardized priors is provided in Table S1 and adapted from Zampieri et al.[14] Representations of the priors on the absolute risk difference scale are provided as dashed lines in Figures 3 and 4. Note that log(relative benefit, [RB]) and log(odds ratio, [OR])=0 are the same as RB and OR=1. ln(OR)=1 is the same as OR≅2.7. Granfeldt refers to the recent meta-analysis by Granfeldt et al.,[29] and Hyperion refers to the HYPERION (Therapeutic Hypothermia after Cardiac Arrest in Nonshockable Rhythm) trial.[9] ESS denotes effective sample size; THAPCA-OH, Therapeutic Hypothermia after Pediatric Cardiac Arrest Out-of-Hospital; and TTM, Targeted Temperature Management trials.[30]

for Statistical Computing), and all Bayesian regressions were fit using Stan version 2.21.0 with the R interface to stan and the brms R package.[32] All R code used in the analysis is publicly available.[33]

Results

For the primary outcome, as in the original trial, 260 trial participants were evaluated and had a VABS-II score ≥70 before OHCA. The frequency of the composite outcome of survival with a good neurobehavioral outcome at 1 year was 20% in the therapeutic hypothermia group and 12% in the therapeutic normothermia group, resulting in a frequentist relative benefit of 1.54 (95% confidence interval, 0.86 to 2.76; P=0.14) or an absolute improvement of 7.3 percentage points (95% confidence interval, −1.5% to 16.1%) as previously reported.[10] Under a noninformative prior, the posterior probability of any improvement in survival with a good neurobehavioral outcome was 94%, the posterior

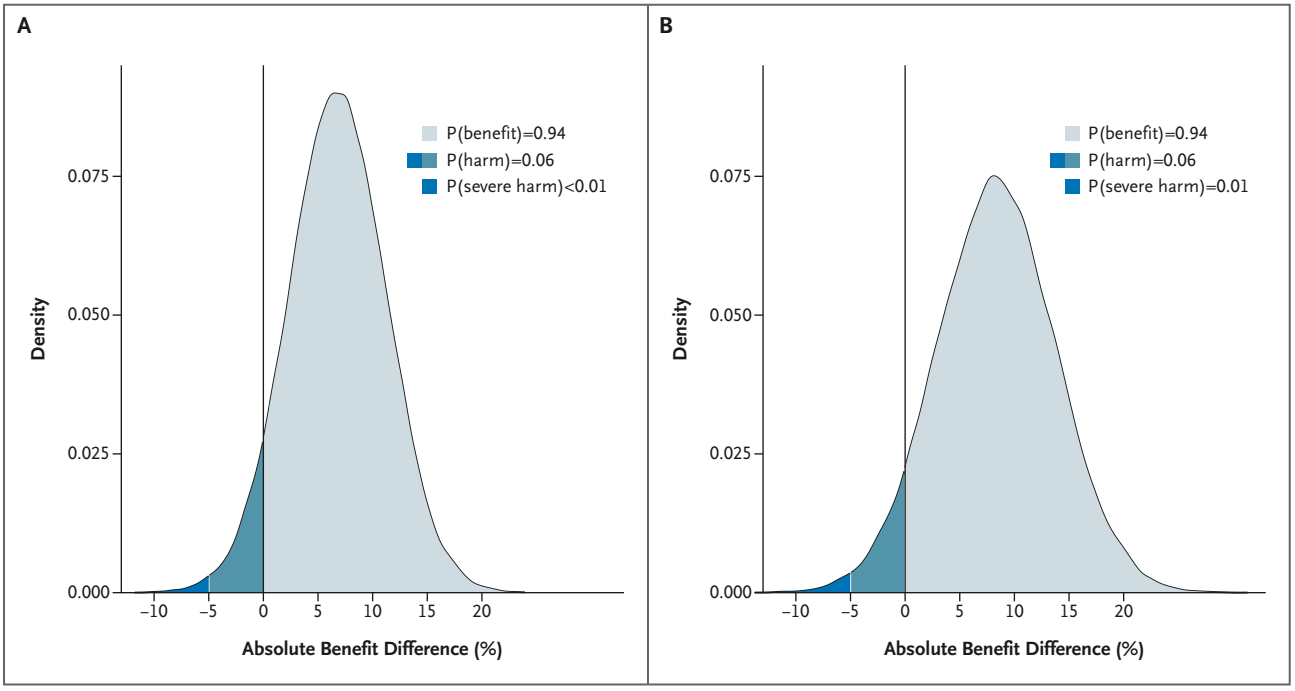

Figure 2. Posterior Probability Distributions from a Bayesian Analysis on the Absolute Benefit (i.e., Risk) Difference Scale Using a Noninformative Prior.

Bayesian analysis of the THAPCA-OH (Therapeutic Hypothermia after Pediatric Cardiac Arrest Out-of-Hospital) trial.[10] Panel A depicts the Bayesian posterior probability distribution for the primary composite outcome of good neurobehavioral outcome at 1 year. Panel B depicts the posterior probability distribution for the secondary outcome of 1-year survival. The legend indicates the shading of the posterior probability distribution indicating different trial interpretations. For example, the light blue shading to the right of a null value of 0 on the x-axis indicates that 94% of the probability distribution supports a trial interpretation of an absolute benefit difference >0. The two darker blue shades indicates the region of the posterior probability distribution congruent with an interpretation of harm. Probability of severe harm (≤1% for both outcomes) of the therapeutic hypothermia treatment was assessed using a cutoff of −0.05 for the absolute benefit parameter. Figure S1 presents these results on the relative benefit scale. P denotes probability.

probability of any harm was 6% (Fig. 2A), and the posterior median absolute benefit was 6.8 percentage points (95% credible interval, −1.9 to 15.4), similar to the frequentist estimate provided in the original article (Table 1). The posterior probability of a relative benefit >1.1 in survival with good neurobehavioral outcome was 88%. The posterior probability of improved survival at 1 year was also 94% (Fig. 2B). Results of the noninformative prior analysis on the relative scale are presented in Figure S1.

We next examined how THAPCA-OH would be interpreted across our three evidence-based priors and a range of theoretical standardized priors (Fig. 1 and Table S1). Our Bayesian analyses using evidence-based priors (Fig. 3, Fig. S2, and Table 1) yielded largely the same conclusions as those from the standardized framework (Fig. 4, Fig. S3, and Table 1). In particular, the analysis showed strong evidence of benefit under the Granfeldt- and Hyperion-based priors, which were similar to optimistic standardized priors (Fig. 4), with posterior probabilities of any benefit >90% and ≥85% for a relative benefit of >1.1. Under the TTM prior, benefit (66%) was more likely than harm (34%), but large benefits were unlikely, mimicking the conclusions from the strong neutral prior analysis.

Each standardized prior and resulting posterior probability distribution are depicted in Figure 4 and Figure S3. It is noteworthy that the standardized optimistic prior mean is lower than the observed log(odds ratio) from THAPCA-OH; that is, even the standardized optimistic priors are pulling the posterior probability distributions toward the null. Table 1 shows that only the moderate pessimistic and strong pessimistic priors result in trial interpretations of a low probability of any benefit (47% and 16% probabilities, respectively). All other priors suggest probabilities upward of 75% of any benefit. Higher benefits become increasingly limited to the weak neutral prior and the optimistic priors once increasing to an absolute benefit of 2% or relative benefit of >1.1.

Table 1. Probability of Absolute and Relative Treatment Effects Estimated by Using Bayesian Analysis According to Varying Prior Beliefs.*

Prior Belief and Strength	Posterior Median ABD‡	Posterior Probability That the True ABD Is Greater Than the Specified Threshold†					Posterior Median RB‡	Posterior Probability That the True RB Is Greater Than the Specified Threshold†				
		>0%	>2%	>5%	>10%	>20%§		>1	>1.1	>1.25	>1.5	>2
Noninformative												
Minimal	6.8 (−1.9 to 15.4)	94	86	66	23	0	1.55 (0.89 to 2.84)	94	88	77	54	20
Evidence-based priors (Fig. 1A)												
Granfeldt	4.4 (−1.2 to 10.3)	94	80	42	3	0	1.33 (0.93 to 1.94)	94	85	63	26	2
TTM	0.5 (−1.9 to 2.9)	66	10	0	0	0	1.03 (0.89 to 1.19)	66	19	0	0	0
Hyperion	7.0 (−0.9 to 14.7)	96	90	69	22	0	1.59 (0.95 to 2.76)	96	92	82	58	20
Standardized priors (Fig. 1B)												
Optimistic												
Weak	5.2 (−1.6 to 12.1)	93	82	52	8	0	1.39 (0.90 to 2.18)	93	85	68	37	6
Moderate	4.0 (−0.7 to 8.8)	95	80	33	1	0	1.28 (0.96 to 1.73)	95	85	57	15	0
Strong	3.4 (0.2 to 6.8)	98	80	17	0	0	1.24 (1.01 to 1.52)	98	88	47	3	0
Neutral												
Weak	6.6 (−2.1 to 15.3)	93	85	65	22	0	1.54 (0.88 to 2.80)	94	88	76	54	19
Moderate	3.5 (−3.0 to 9.8)	86	67	32	2	0	1.24 (0.84 to 1.87)	86	73	49	18	1
Strong	1.8 (−2.8 to 6.3)	78	46	8	0	0	1.11 (0.84 to 1.48)	78	54	21	2	0
Pessimistic												
Weak	2.9 (−4.0 to 9.8)	80	60	27	2	0	1.20 (0.79 to 1.87)	80	66	43	16	1
Moderate	−0.2 (−5.0 to 4.5)	47	18	2	0	0	0.99 (0.74 to 1.32)	47	24	6	0	0
Strong	−1.6 (−5.0 to 1.6)	16	2	0	0	0	0.90 (0.74 to 1.10)	16	3	0	0	0

* The priors in the first column are displayed in Figure 1 and defined in the Methods and Table S1. Granfeldt refers to the recent meta-analysis by Granfeldt et al,[29] and Hyperion refers to the HYPERION (Therapeutic Hypothermia after Cardiac Arrest in Nonshockable Rhythm) trial.[9] ABD denotes absolute benefit difference; RB, relative benefit; and TTM, Targeted Temperature Management trials.[30]

† Probability of absolute and relative treatment effects estimated by using Bayesian analysis according to varying prior beliefs about good neurobehavioral outcome benefit at 1 year from hypothermia versus normothermia treatment for pediatric out-of-hospital cardiac arrest in the THAPCA-OH (Therapeutic Hypothermia after Pediatric Cardiac Arrest Out-of-Hospital) trial.[10]

‡ Data are shown as percent or percent (95% credible interval).

§ The absolute improvement of 20% was displayed to align with the targeted value used to power the original trial.

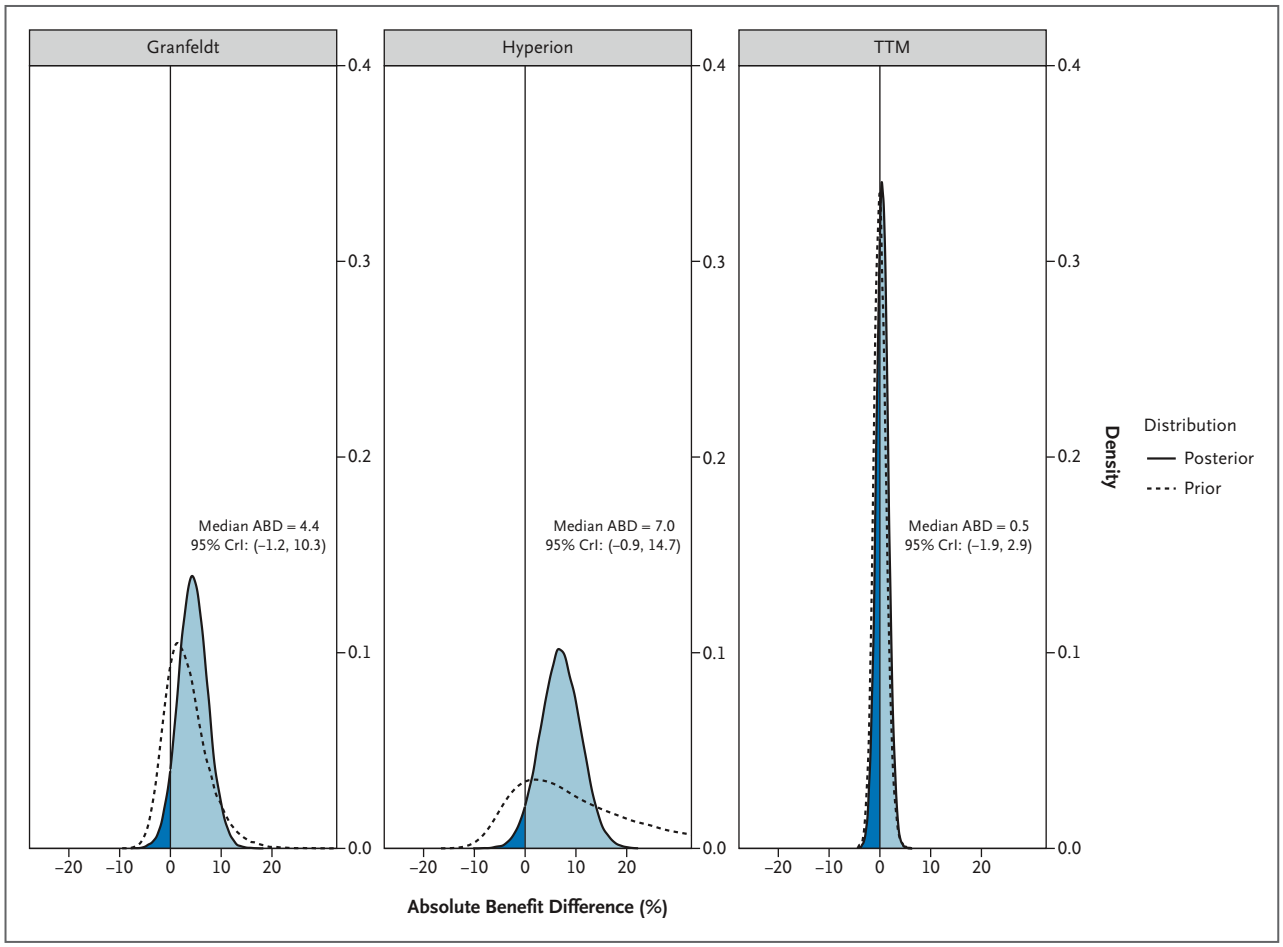

Figure 3. Bayesian Reanalysis Using Evidence-Based Priors.

Reanalysis of the THAPCA-OH (Therapeutic Hypothermia after Pediatric Cardiac Arrest Out-of-Hospital) trial[10] using the evidence-based priors (dashed line) described in the Methods section. The solid line and shaded distribution represent the updated trial interpretation (i.e., posterior probability distribution of the absolute benefit [i.e., risk difference] of the primary outcome of good neurobehavioral outcome in THAPCA-OH when incorporated with each prior). Light blue indicates the region of the updated posterior probability distribution congruent with an interpretation of benefit (i.e., absolute benefit difference [ABD] >0). Dark blue indicates the region of the updated posterior probability distribution congruent with an interpretation of harm (i.e., ABD <0). Figure S2 presents these results on the relative benefit scale. Granfeldt refers to the recent meta-analysis by Granfeldt et al,[29] and Hyperion refers to the HYPERION (Therapeutic Hypothermia after Cardiac Arrest in Nonshockable Rhythm) trial.[9] CrI denotes credible interval; and TTM, Targeted Temperature Management trials.[30]

Stated differently, there is a probability of some benefit regardless of the selected prior. Results for the Bayesian analysis for 1-year survival had nearly identical numerical values and interpretation (Tables S2 and S3 and Figs. S4 and S5).

Discussion

This reexamination of THAPCA-OH through a Bayesian perspective showed that, in children with OHCA, treatment with therapeutic hypothermia has a high probability of good neurobehavioral outcome and survival at 1 year relative to therapeutic normothermia. This result was consistent across several informative priors. Although the 95% credible interval for the estimate of absolute improvement with a noninformative prior included zero (−1.9 to 15.4), from a Bayesian standpoint, this estimate shows a high probability of benefit from treatment (specifically, 94%). In contrast, when frequentist confidence intervals include zero, this merely signifies that the null hypothesis cannot be rejected, and no additional inference about treatment benefit is permitted. In Bayesian analyses, the

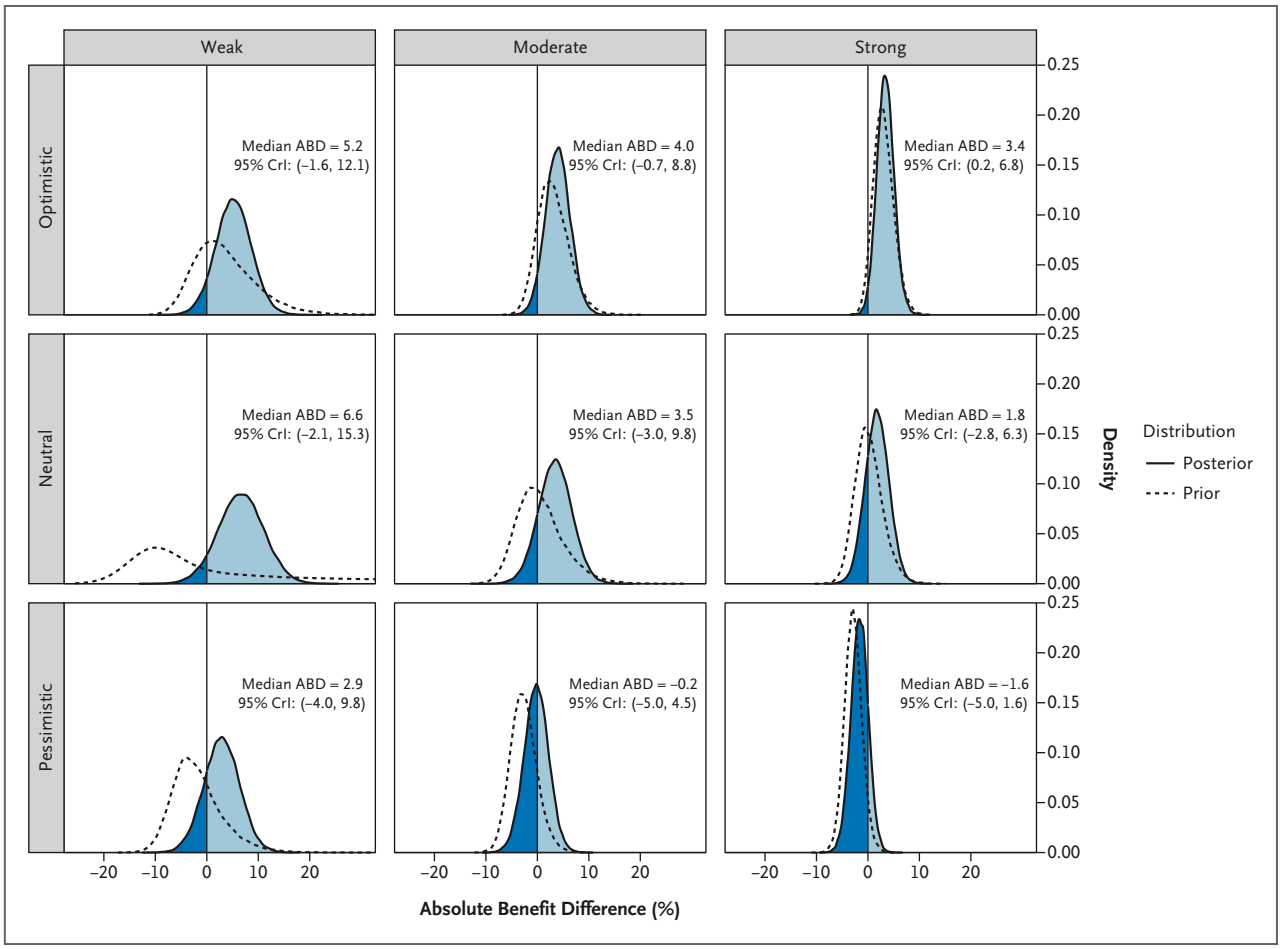

Figure 4. Bayesian Reanalysis Using Standardized Critical Care Priors.

Reanalysis of the THAPCA-OH (Therapeutic Hypothermia after Pediatric Cardiac Arrest Out-of-Hospital) trial[10] using the standardized critical care priors (dashed line) from Zampieri et al.[14] and summarized in Table S1. The solid line and shaded distribution represent the updated trial interpretation (i.e., posterior probability distribution of the absolute benefit [i.e., risk difference] of the primary outcome of good neurobehavioral outcome in THAPCA-OH when incorporated with each prior). Light blue indicates the region of the updated posterior probability distribution congruent with an interpretation of benefit (i.e., absolute benefit difference [ABD] >0). Dark blue indicates the region of the updated posterior probability distribution congruent with an interpretation of harm (i.e., ABD <0). Figure S3 presents these results on the relative benefit scale. CrI denotes credible interval.

posterior probabilities complement the credible intervals by indicating how much of the probability density lies above or below zero (or any clinically relevant threshold value) and thus clarify the strength of the evidence in support of treatment benefit or harm. With the noninformative prior, the probability of any benefit from hypothermia was 94% for both neurobehavioral outcome and 1-year survival. Our interpretation of these numbers is that, although harm from therapeutic hypothermia cannot be ruled out, the risk of harm, given these data, is low (6%). Decisions about treatment based on these probabilities must incorporate additional considerations about the intervention, such as other risks, costs, prognosis, and patient or family preferences. By directly addressing the primary question raised by patients and clinicians — what are the probabilities of benefit and harm? — Bayesian posterior probabilities maximize the information available from trials for clinical decision-making.

Indeed, the possibility of a Bayesian reanalysis for THAPCA-OH was suggested in the first letters responding to its publication,[34] an intuitive recognition that describing therapeutic hypothermia as "not confer[ing] a significant benefit in survival with a good functional outcome at one

year" may have been an unfortunate statistical conclusion required by the traditional frequentist framework when, in fact, a potentially important clinical benefit was not ruled out. The original publication noted this possibility in the discussion,[10] and such a conclusion is supported by our Bayesian analysis and a previous analysis by Tasker and Akhondi-Asl.[35] THAPCA-OH is not unique in this regard and is one of several recently published critical care trials that has stirred debate about trial interpretation and led to concern that a promising therapy could be discarded. Other examples include ANDROMEDA-SHOCK,[36] EOLIA (ECMO to Rescue Lung Injury in Severe ARDS),[37] and the COVID STEROID 2 (Higher vs. Lower Doses of Dexamethasone in Patients With COVID-19 and Severe Hypoxia) trial,[38] all of which identified a sizeable improvement in outcomes that were not statistically significant using a frequentist approach. Although any missed intervention for an illness is a loss to society and a failure of medical research, the public health impact of pediatric OHCA alongside limited randomized trials, as well as the difficulty or inability to conduct new trials, make the failure to fully consider the clinical utility of therapeutic hypothermia especially salient.

There are currently no therapies to improve outcomes from pediatric OHCA beyond supportive care. Differences in the etiology, arrest characteristics, and outcomes of pediatric OHCA[3] make extrapolation of adult data problematic, as the relative risks and benefits of a particular therapy, including therapeutic hypothermia, cannot be assumed to be parallel to those of adults. Trial outcomes and rehabilitation trajectories likely differ between children and adults as the pediatric brain is still developing at the time of cardiac arrest, thus making the need to fully analyze and interpret the limited pediatric data more compelling. Although THAPCA-OH was a relatively small trial, the benefit of therapeutic hypothermia was consistent with earlier adult trials,[5,6] as well as the more contemporary adult HYPERION trial, which included initial presenting rhythms that were more consistent with pediatric OHCA.[9] Our posterior probabilities favored benefit of therapeutic hypothermia (at least 75% probability of any benefit and >50% probability of a relative benefit of >1.1) in all except the most strongly pessimistic priors. Of importance, the pessimistic priors provided here are highly unlikely on the basis of the existing literature. In other words, therapeutic hypothermia was favored over the entire range of plausible priors informing pediatric OHCA, as no adult or pediatric data support the substantial harm (relative risk, 0.8 for good neurobehavioral outcome) reflected by the pessimistic standardized prior. The pessimistic priors, therefore, are more appropriately considered a sensitivity analysis reflecting an extremely unlikely scenario.

In recent years, there has been an increase in Bayesian trial reanalysis in critical care,[14,17,31,39-41] and the number of trials designed and analyzed primarily under a Bayesian framework is rising.[17,42,43] Undoubtedly, Bayesian analyses can be a helpful tool by which to augment the interpretation of critical care trials.[18,19,39-41,44,45] This is both because the interpretation of results from a trial conducted using frequentist methods can be rigid (i.e., binary) and, especially in critical care, trial results are sensitive to improbable assumptions made during the design phase. As critical care outcomes improve over time, historic epidemiologic and outcome data become outdated, affecting the reasonableness of the inputs for power calculations. Furthermore, overly optimistic expectations about the effect of treatments can influence trial design and interpretation.[17,46] Of note, our reanalysis showed low probability, even under optimistic priors, of a beneficial effect of therapeutic hypothermia as large as the postulated 20% absolute improvement. It is widely appreciated that such estimates are speculative. Thus, we believe there are compelling arguments for the continued use of Bayesian methods to enhance both the design and analysis of trials to improve the ability of trials to resolve challenging questions in clinical practice.

There is also a broader debate in the medical community about interpreting large, but not statistically significant, effects in trials that stems from confusion and frustration regarding the meaning — and thus interpretation — of P values and null hypothesis testing using the frequentist framework.[14,47] Many superiority trials such as THAPCA-OH are designed assuming a null hypothesis of no treatment effect. P values then provide evidence of how (in)compatible trial results are with that assumption, with the common binary interpretation that $P \geq 0.05$ is compatible with the null and $P < 0.05$ is incompatible. Furthermore, P values provide limited information regarding the size or clinical significance of the measured treatment effect, which is the underlying information of interest. As a result, large effects that do not cross this binary threshold often emerge as highly controversial. In this specific instance, on the basis of the THAPCA-OH trial, the International Liaison Committee on Resuscitation recommends a better frequentist THAPCA-OH interpretation that there is "inconclusive evidence to support or refute the use of induced hypothermia (32°C to 34°C) compared with active control of temperature at normothermia (36°C to 37.5°C; or an alternative temperature) for

children who achieve return of spontaneous circulation but remain comatose after OHCA or in-hospital [cardiac arrest]."[48] These interpretations show the limitation of using null hypothesis testing and the misconception that trials which fail to reach arbitrary P value thresholds are negative. Such trials may be more accurately termed indeterminate,[49] but this precision in language and interpretation is rare in the abstracts of most trial manuscripts. Consequently, researchers, readers, and editors often equate trials with $P>0.05$ as providing no evidence regarding the true treatment effect or, worse, suggesting that it provides evidence of intervention inefficacy.[50] As we report with THAPCA-OH, which found at least a 75% probability of any benefit with therapeutic hypothermia under all plausible priors, it is possible for a trial with $P>0.05$ to largely support a conclusion of efficacy.

There are limitations to the current analyses. Foremost, ideally, all analyses should be prespecified — along with the priors that will be used for Bayesian analyses — before the trial results are unblinded. When the opportunity for prespecification is not available, standardization of priors then becomes a way to minimize potential post hoc bias, as they encapsulate a wide range of harm and benefit scenarios; however, as occurred in our analyses with the pessimistic priors, these scenarios are still hypothetical and may not perfectly capture reality. The additional use of evidence-based priors helps mitigate this issue. Relatedly, the sample size of the THAPCA-OH trial was small. Yarnell et al.[41] examined the relationship between Bayesian and frequentist results in critical care trials. Because frequentist designs generally use a null hypothesis of no effect, such as in THAPCA-OH, as opposed to an effect less than a minimal clinically important difference, even large frequentist trials with highly significant P values can yield low posterior probabilities of clinically meaningful benefit if the treatment effects are very small; however, they also showed that decreases in sample size progressively increased the susceptibility of trial results to varying priors. In the current study, we can see that our small sample size led to variable interpretations across the different standardized priors; therefore, we suggest the strongest emphasis be placed on the primary analysis that used a noninformative prior. Second, our reanalysis focused on the overall, or average, treatment effect and did not examine heterogeneity of treatment effects. Third, a prior preplanned analysis[51] pooled THAPCA-OH and THAPCA-IH (Therapeutic Hypothermia after In-Hospital Cardiac Arrest in Children), which focused on in-hospital cardiac arrest; however, we focused only on OHCA, and thus our reanalysis should not be extrapolated to in-hospital cardiac arrest.

In conclusion, a post hoc secondary analysis of THAPCA-OH in a Bayesian framework using noninformative, evidence-based, and standardized priors was most consistent with an interpretation that children who remained comatose after OHCA and received therapeutic hypothermia had higher probabilities of a good neurobehavioral outcome and survival at 1 year than those who received therapeutic normothermia, supporting the use of therapeutic hypothermia after pediatric OHCA.

Disclosures

Author disclosures and other supplementary materials are available at evidence.nejm.org.

Author Affiliations

[1] Clinical Trials Methods and Outcomes Lab, Palliative and Advanced Illness Research Center, Perelman School of Medicine at the University of Pennsylvania, Philadelphia

[2] Department of Biostatistics, Epidemiology, and Informatics, Perelman School of Medicine at the University of Pennsylvania, Philadelphia

[3] Department of Intensive Care 4131, Copenhagen University Hospital, Rigshospitalet, Copenhagen

[4] Department of Pediatrics, University of Michigan Medical School, Ann Arbor, MI

[5] Academic Research Organization, Hospital Albert Einstein, São Paulo

[6] Department of Critical Care Medicine, Faculty of Medicine and Dentistry, University of Alberta, AB, Canada

[7] Interdepartmental Division of Critical Care Medicine, University of Toronto, Toronto

[8] Department of Medicine, Division of Respirology, University Health Network, Toronto

[9] Toronto General Hospital Research Institute, Toronto

[10] Department of Anesthesiology and Critical Care Medicine, Children's Hospital of Philadelphia, Philadelphia

References

1. Michelson KA, Hudgins JD, Monuteaux MC, Bachur RG, Finkelstein JA. Cardiac arrest survival in pediatric and general emergency departments. Pediatrics 2018;141:e20172741. DOI: 10.1542/peds.2017-2741.
2. Jayaram N, McNally B, Tang F, Chan PS. Survival after out-of-hospital cardiac arrest in children. J Am Heart Assoc 2015;4:e002122. DOI: 10.1161/JAHA.115.002122.
3. Tsao CW, Aday AW, Almarzooq ZI, et al. Heart disease and stroke statistics — 2022 update: a report from the American Heart Association. Circulation 2022;145:e153-e639. DOI: 10.1161/CIR.0000000000001052.
4. Meert KL, Telford R, Holubkov R, et al. Pediatric out-of-hospital cardiac arrest characteristics and their association with survival and neurobehavioral outcome. Pediatr Crit Care Med 2016;17:e543-e550. DOI: 10.1097/PCC.0000000000000969.
5. The Hypothermia after Cardiac Arrest Study Group. Mild therapeutic hypothermia to improve the neurologic outcome after cardiac arrest. N Engl J Med 2002;346:549-556. DOI: 10.1056/NEJMoa012689.

6. Bernard SA, Gray TW, Buist MD, et al. Treatment of comatose survivors of out-of-hospital cardiac arrest with induced hypothermia. N Engl J Med 2002;346:557-563. DOI: 10.1056/NEJMoa003289.

7. Nielsen N, Wetterslev J, Cronberg T, et al. Targeted temperature management at 33°C versus 36°C after cardiac arrest. N Engl J Med 2013;369:2197-2206. DOI: 10.1056/NEJMoa1310519.

8. Dankiewicz J, Cronberg T, Lilja G, et al. Hypothermia versus normothermia after out-of-hospital cardiac arrest. N Engl J Med 2021; 384:2283-2294. DOI: 10.1056/NEJMoa2100591.

9. Lascarrou J-B, Merdji H, Le Gouge A, et al. Targeted temperature management for cardiac arrest with nonshockable rhythm. N Engl J Med 2019;381:2327-2337. DOI: 10.1056/NEJMoa1906661.

10. Moler FW, Silverstein FS, Holubkov R, et al. Therapeutic hypothermia after out-of-hospital cardiac arrest in children. N Engl J Med 2015;372:1898-1908. DOI: 10.1056/NEJMoa1411480.

11. Holubkov R, Clark AE, Moler FW, et al. Efficacy outcome selection in the therapeutic hypothermia after pediatric cardiac arrest trials. Pediatr Crit Care Med 2015;16:1-10. DOI: 10.1097/PCC.0000000000000272.

12. Scholefield BR, Gao F, Duncan HP, et al. Observational study of children admitted to United Kingdom and Republic of Ireland paediatric intensive care units after out-of-hospital cardiac arrest. Resuscitation 2015;97:122-128. DOI: 10.1016/j.resuscitation.2015.07.011.

13. Sung L, Hayden J, Greenberg ML, Koren G, Feldman BM, Tomlinson GA. Seven items were identified for inclusion when reporting a Bayesian analysis of a clinical study. J Clin Epidemiol 2005;58:261-268. DOI: 10.1016/j.jclinepi.2004.08.010.

14. Zampieri FG, Casey JD, Shankar-Hari M, Harrell FE Jr, 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 2021;203:543-552. DOI: 10.1164/rccm.202006-2381CP.

15. Biologic Specimen and Data Repository Information Coordinating Center. National Heart, Lung, and Blood Institute. December 3, 2021 (https://biolincc.nhlbi.nih.gov/home/).

16. Sparrow SS, Cicchetti D, Balla DA. Vineland-II: Vineland Adaptive Behavior Scales: Survey Forms Manual. Bloomington, MN: Pearson, 2005. DOI: 10.1037/t15164-000.

17. Harhay MO, Casey JD, Clement M, et al. Contemporary strategies to improve clinical trial design for critical care research: insights from the First Critical Care Clinical Trialists Workshop. Intensive Care Med 2020;46:930-942. DOI: 10.1007/s00134-020-05934-6.

18. Bittl JA, He Y. Bayesian analysis: a practical approach to interpret clinical trials and create clinical practice guidelines. Circ Cardiovasc Qual Outcomes 2017;10:e003563. DOI: 10.1161/CIRCOUTCOMES.117.003563.

19. Wijeysundera DN, Austin PC, Hux JE, Beattie WS, Laupacis A. Bayesian statistical inference enhances the interpretation of contemporary randomized controlled trials. J Clin Epidemiol 2009;62: 13-21.e5. DOI: 10.1016/j.jclinepi.2008.07.006.

20. Gelman A, Carlin JB, Stern HS, Dunson DB, Vehtari A, Rubin DB. Bayesian data analysis. Boca Raton, FL: CRC Press, 2013. DOI: 10.1201/b16018.

21. Lewis RJ, Angus DC. Time for clinicians to embrace their inner Bayesian? Reanalysis of results of a clinical trial of extracorporeal membrane oxygenation. JAMA 2018;320:2208-2210. DOI: 10.1001/jama.2018.16916.

22. Fink EL, Clark RS, Kochanek PM, Bell MJ, Watson RS. A tertiary care center's experience with therapeutic hypothermia after pediatric cardiac arrest. Pediatr Crit Care Med 2010;11:66-74. DOI: 10.1097/PCC.0b013e3181c58237.

23. Chang I, Kwak YH, Shin SD, et al. Therapeutic hypothermia and outcomes in paediatric out-of-hospital cardiac arrest: a nationwide observational study. Resuscitation 2016;105:8-15. DOI: 10.1016/j.resuscitation.2016.04.021.

24. Gluckman PD, Wyatt JS, Azzopardi D, et al. Selective head cooling with mild systemic hypothermia after neonatal encephalopathy: multicentre randomised trial. Lancet 2005;365:663-670. DOI: 10.1016/S0140-6736(05)17946-X.

25. Shankaran S, Laptook AR, Ehrenkranz RA, et al. Whole-body hypothermia for neonates with hypoxic–ischemic encephalopathy. N Engl J Med 2005;353:1574-1584. DOI: 10.1056/NEJMcps050929.

26. Schenone AL, Cohen A, Patarroyo G, et al. Therapeutic hypothermia after cardiac arrest: a systematic review/meta-analysis exploring the impact of expanded criteria and targeted temperature. Resuscitation 2016;108:102-110. DOI: 10.1016/j.resuscitation.2016.07.238.

27. Shrestha DB, Sedhai YR, Budhathoki P, et al. Hypothermia versus normothermia after out-of-hospital cardiac arrest: a systematic review and meta-analysis of randomized controlled trials. Ann Med Surg (Lond) 2022;74:103327. DOI: 10.1016/j.amsu.2022.103327.

28. Kim JH, Nagy Á, Putzu A, et al. Therapeutic hypothermia in critically ill patients: a systematic review and meta-analysis of high quality randomized trials. Crit Care Med 2020;48:1047-1054. DOI: 10.1097/CCM.0000000000004364.

29. Granfeldt A, Holmberg MJ, Nolan JP, Soar J, Andersen LW. Targeted temperature management in adult cardiac arrest: systematic review and meta-analysis. Resuscitation 2021;167:160-172. DOI: 10.1016/j.resuscitation.2021.08.040.

30. Holgersson J, Meyer MAS, Dankiewicz J, et al. Hypothermic versus normothermic temperature control after cardiac arrest. NEJM Evid. 2022;1(11). DOI: 10.1056/EVIDoa2200137.

31. Granholm A, Munch MW, Myatra SN, et al. 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. Intensive Care Med 2022;48:45-55. DOI: 10.1007/s00134-021-06573-1.

32. Bürkner P-C. brms: an R package for Bayesian multilevel models using stan. J Stat Softw 2017;80:1-28. DOI: 10.18637/jss.v080.i01.

33. harhay-lab/THAPCA-OH. GitHub. November 17, 2022 (https://github.com/harhay-lab/THAPCA-OH).

34. Geva A, Tasker RC, Randolph AG. Therapeutic hypothermia in children. N Engl J Med 2015;373:979-980. DOI: 10.1056/NEJMc1508693.

35. Tasker RC, Akhondi-Asl A. Targeted temperature management after cardiac arrest due to drowning: "Frequentist" and "Bayesian" decision making. Pediatr Crit Care Med 2016;17:789-791. DOI: 10.1097/PCC.0000000000000799.

36. Hernández G, Ospina-Tascón GA, Damiani LP, et al. Effect of a resuscitation strategy targeting peripheral perfusion status vs serum lactate levels on 28-day mortality among patients with septic shock: the ANDROMEDA-SHOCK randomized clinical trial. JAMA 2019;321:654-664. DOI: 10.1001/jama.2019.0071.

37. Combes A, Hajage D, Capellier G, et al. Extracorporeal membrane oxygenation for severe acute respiratory distress syndrome. N Engl J Med 2018;378:1965-1975. DOI: 10.1056/NEJMoa1800385.

38. COVID STEROID 2 Trial Group. 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 2021;326:1807-1817. DOI: 10.1001/jama.2021.18295.

39. Goligher EC, Tomlinson G, Hajage D, et al. Extracorporeal membrane oxygenation for severe acute respiratory distress syndrome and posterior probability of mortality benefit in a post hoc Bayesian analysis of a randomized clinical trial. JAMA 2018;320:2251-2259. DOI: 10.1001/jama.2018.14276.

40. Zampieri FG, Damiani LP, Bakker J, et al. Effects of a resuscitation strategy targeting peripheral perfusion status vs serum lactate levels among patients with septic shock: a Bayesian reanalysis of the ANDROMEDA-SHOCK trial. Am J Respir Crit Care Med 2020;201:423-429. DOI: 10.1164/rccm.201905-0968OC.

41. Yarnell CJ, Abrams D, Baldwin MR, et al. Clinical trials in critical care: can a Bayesian approach enhance clinical and scientific decision making? Lancet Respir Med 2021;9:207-216. DOI: 10.1016/S2213-2600(20)30471-9.

42. Angus DC, Berry S, Lewis RJ, et al. The REMAP-CAP (Randomized Embedded Multifactorial Adaptive Platform for Community-acquired Pneumonia) study. Rationale and design. Ann Am Thorac Soc 2020;17:879-891. DOI: 10.1513/AnnalsATS.202003-192SD.

43. The REMAP-CAP, ACTIV-4a, and ATTACC Investigators. Therapeutic anticoagulation with heparin in critically ill patients with Covid-19. N Engl J Med 2021;385:777-789. DOI: 10.1056/NEJMoa2103417.

44. Granholm A, Marker S, Krag M, et al. Heterogeneity of treatment effect of prophylactic pantoprazole in adult ICU patients: a post hoc analysis of the SUP-ICU trial. Intensive Care Med 2020;46:717-726. DOI: 10.1007/s00134-019-05903-8.

45. Brophy JM. Bayesian interpretation of the EXCEL trial and other randomized clinical trials of left main coronary artery revascularization. JAMA Intern Med 2020;180:986-992. DOI: 10.1001/jamainternmed.2020.1647.

46. Harhay MO, Wagner J, Ratcliffe SJ, et al. Outcomes and statistical power in adult critical care randomized trials. Am J Respir Crit Care Med 2014;189:1469-1478. DOI: 10.1164/rccm.201401-0056CP.

47. Neyman J. Outline of a theory of statistical estimation based on the classical theory of probability. Philos Trans R Soc Lond A 1937;236:333-380. DOI: 10.1098/rsta.1937.0005.

48. ILCOR Pediatric Life Support (PLS) Task Force. Post-arrest temperature management in children. Statement on post cardiac arrest temperature management in children. International Liaison Committee on Resuscitation. November 2021 (https://ilcor.org/news/post-arrest-temperature-management-in-children).

49. Sackett DL. Superiority trials, noninferiority trials, and prisoners of the 2-sided null hypothesis. ACP J Club 2004;140:A11. DOI: 10.7326/ACPJC-2004-140-2-A11.

50. Altman DG, Bland JM. Absence of evidence is not evidence of absence. BMJ 1995;311:485. DOI: 10.1136/bmj.311.7003.485.

51. Scholefield BR, Silverstein FS, Telford R, et al. Therapeutic hypothermia after paediatric cardiac arrest: pooled randomized controlled trials. Resuscitation 2018;133:101-107. DOI: 10.1016/j.resuscitation.2018.09.011.