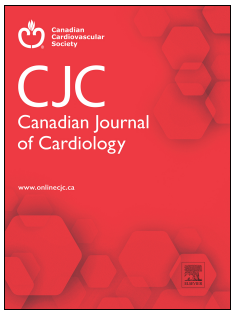
Journal Pre-proof



Bayesian Analytical Methods in Cardiovascular Clinical Trials: Why, When, and How

Samuel Heuts, MD, PhD, Michal J. Kawczynski, MD, Ahmed Sayed, MD, Sarah M. Urbut, MD, PhD, Arthur M. Albuquerque, MD, John M. Mandrola, MD, Sanjay Kaul, MD, Frank E. Harrell, Jr., PhD, Andrea Gabrio, PhD, James M. Brophy, MD, PhD

PII: S0828-282X(24)01130-9

DOI: https://doi.org/10.1016/j.cjca.2024.11.002

Reference: CJCA 5259

To appear in: *Canadian Journal of Cardiology*

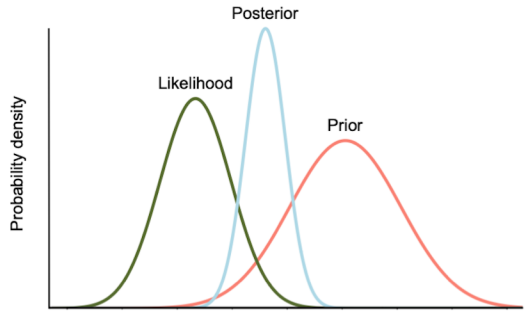
Received Date: 3 October 2024

Revised Date: 28 October 2024

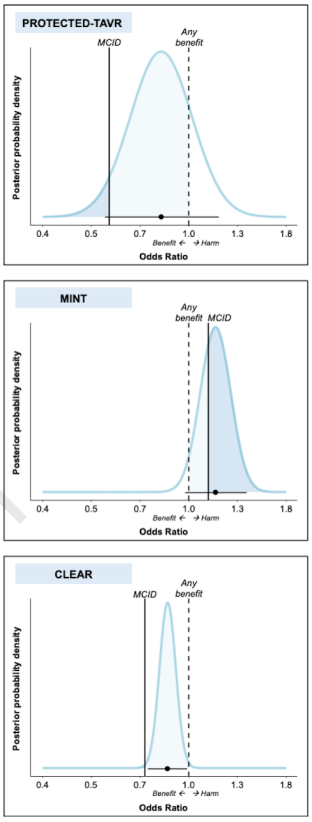
Accepted Date: 3 November 2024

Please cite this article as: Heuts S, Kawczynski MJ, Sayed A, Urbut SM, Albuquerque AM, Mandrola JM, Kaul S, Harrell Jr. FE, Gabrio A, Brophy JM, Bayesian Analytical Methods in Cardiovascular Clinical Trials: Why, When, and How, *Canadian Journal of Cardiology* (2024), doi: https://doi.org/10.1016/j.cjca.2024.11.002.

Bayesian Analytical Methods in Cardiovascular Clinical Trials:
Why, When, and How



1.	Identification of the appropriate research question
2.	Choose the desired risk estimate and define the MCID
3.	Select the designated statistical model
4.	Prespecify the priors and perform predictive checks: • *Less informative priors* • *Reference priors* • ***Literature-based priors***
5.	Obtain the posterior distribution of the desired effect conditional on the prior, the data and the assumed statistical model, using the appropriate statistical software
6.	Verify the correctness and robustness of results, including predictive checks, and sensitivity analyses using varying priors



Bayesian Analytical Methods in Cardiovascular Clinical Trials: Why, When, and How

Bayesian Methods in Clinical Trials

Samuel Heuts, MD, PhD[1,2,*], Michal J. Kawczynski, MD[1,2], Ahmed Sayed, MD[3], Sarah M. Urbut, MD, PhD[4], Arthur M. Albuquerque, MD[5], John M. Mandrola, MD[6], Sanjay Kaul, MD, Frank E. Harrell Jr., PhD[8], Andrea Gabrio, PhD[9,10], and James M. Brophy, MD, PhD[11]

[1] Department of Cardiothoracic Surgery, Maastricht University Medical Centre (MUMC+), Maastricht, the Netherlands

[2] Cardiovascular Research Institute Maastricht (CARIM), Maastricht University, Maastricht, the Netherlands

[3] Faculty of Medicine, Ain Shams University, Cairo, Egypt

[4] Division of Cardiology, Department of Medicine, Massachusetts General Hospital, Harvard Medical School, Boston, Massachusetts, USA

[5] Universidade Federal do Rio de Janeiro, Rio de Janeiro, Brazil

[6] Baptist Health Louisville, Louisville, Kentucky, USA

[7] Department of Cardiology, Cedars-Sinai Medical Centre, Los Angeles, California, USA

[8] Department of Biostatistics, Vanderbilt University School of Medicine, Nashville, Tennessee, USA

[9] Department of Methodology and Statistics, Maastricht University, Maastricht, the Netherlands

[10] Care and Public Health Institute (CAPHRI), Maastricht University, Maastricht, the Netherlands

[11] McGill University Health Centre, Centre for Health Outcome Research (CORE), Montreal, Quebec, Canada

Word count abstract: 250

Word count text: 10273

Number of references: 60

Conflicts of interest: none declared

Funding: none received

Acknowledgements: none

*Correspondence to:

Samuel Heuts, MD, PhD

Department of Cardiothoracic Surgery

Maastricht University Medical Centre (MUMC+)

P. Debyelaan 25

6229HX, Maastricht, the Netherlands

e: sam.heuts@mumc.nl

t: +31 443 387 50 70

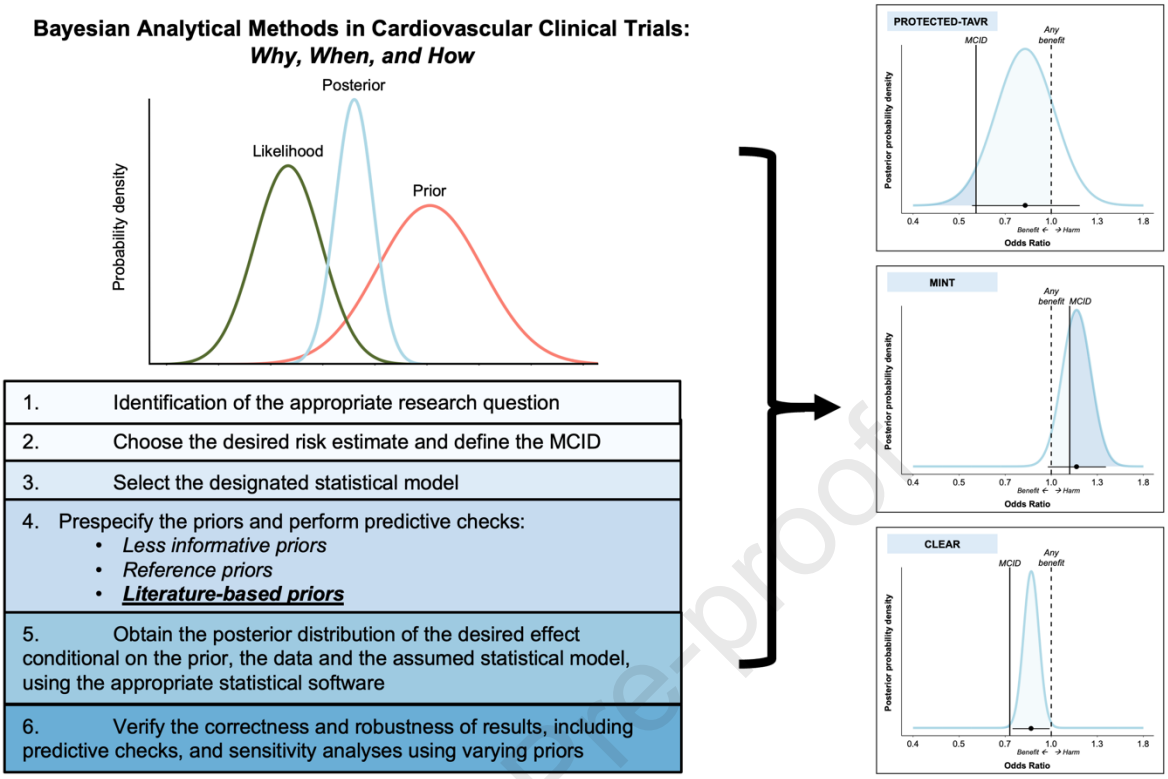
ABSTRACT

The Bayesian analytical framework is clinically intuitive, characterized by the incorporation of previous evidence into the analysis, and allowing an estimation treatment effects and their associated uncertainties. The application of Bayesian statistical inference is not new to the cardiovascular field, as illustrated by various recent randomized trials that applied a primary Bayesian analysis. Given the guideline-shaping character of trials, a thorough understanding of the concepts and technical details of Bayesian statistical methodology is of utmost importance to the modern practicing cardiovascular physician. Therefore, this Review aims to present a step-by-step guide to interpreting and performing a Bayesian (re-)analysis of cardiovascular clinical trials, while highlighting the main advantages of Bayesian inference for the clinical reader. After an introduction of the concepts of frequentist and Bayesian statistical inference and reasons to apply Bayesian methods, key steps for performing a Bayesian analysis are presented, including: the verification of the clinical appropriateness of the research question, the quality and completeness of the trial design, as well as the adequate elicitation of the prior (i.e., ones belief towards a certain treatment before the current evidence becomes available), identification of the likelihood, and their combination into a posterior distribution. Examination of this posterior distribution offers the possibility of not only determining the probability of treatment superiority, but also the probability of exceeding any chosen minimal clinically important difference. Multiple priors should be transparently prespecified, limiting post-hoc manipulations. Using this guide, three cardiovascular randomized controlled trials are re-analysed, demonstrating the clarity and versatility of Bayesian inference.

KEYWORDS

Bayesian statistics; Bayesian inference; frequentism, randomized controlled trials; clinical trials; statistics

73 GRAPHICAL ABSTRACT



74

75

76

INTRODUCTION

Well-designed and executed randomized controlled trials (RCTs) are at the core of evidence-based medicine [1]. The most commonly applied method for statistical inference in trials, the frequentist framework, relies on familiar concepts such as null hypothesis significance testing (NHST) and p-values [2]. Still, the frequentist approach presents several cognitive difficulties [1-4]. Some of these difficulties may be more intuitively addressed through the application of Bayesian statistical inference. After an introduction to the basic concepts of frequentist and Bayesian statistical inference, and reasons and venues to apply the Bayesian methods, we present a *how-to-do-it* guide using intuitive examples of three contemporary cardiovascular clinical trials. Consequently, this Review on *Methods in Cardiovascular Research and Practice* aims to present the advantages (and limitations) of Bayesian statistical inference and provide the tools to allow interested readers to independently perform a Bayesian (re)analysis of an RCT.

THE FREQUENTIST STATISTICAL FRAMEWORK

Clinicians may not be aware of the existence and implications of different statistical inferential frameworks. Historically, RCTs have been analysed under a *frequentist* statistical framework, which incorporates null hypothesis significance testing (NHST) and p-values. Although the originators of these two concepts deemed them irreconcilable [5], this combination has paradoxically become the cornerstone of statistical inference regarding trial data.

The NHST concept is an example of *deductive* inference, in which one starts with a hypothesis (typically the null hypothesis of no effect, denoted as H_0), and tests whether observations are consistent with that hypothesis. This can be mathematically denoted as *P (data | H).* In this paradigm, the hypothesis is therefore considered 'known', and the data or observations are considered 'variable'. Clinicians often erroneously believe the p-value denotes the probability that the alternative hypothesis (denoted as H_1) is true, when in fact, the p-value is the probability of observing the data, or more extreme data, under the assumption that H_0 is true, in an infite number of future hypothetical trials under similar circumstances. P-value misconceptions have long been appreciated and have been elegantly summarized by American Statistical Association publications [3, 4]. Moreover, the frequentist framework applies a virtually universal level of statistical significance, known as the α-level (usually <0.05). As such, the focus is on statistical significance of the null hypothesis, rather than on effect size estimation. Notably, the p-value is fully dependent on the treatment effect size and, particularly, the sample size. Consequently, a negligibly small effect may reach statistical significance with an infinitely large sample size. While confidence intervals (CIs, usually 95%) are seen as an improvement, they suffer from the same limitations as they only provide a sampling distribution for theoretically repeated experiments under the null hypothesis of no effect. While providing long-term assurance that the true effect will be in the interval, this provides little assistance to

any particular study at the present time, as the true effect is either in the interval, or not. A further explanation of these frequentist statistical terms is summarized and provided in the glossary in **Table 1**. Furthermore, **Table 2** presents the main features, advantages, and downsides of the frequentist statistical approach.

WHAT IS BAYESIAN STATISTICAL METHODOLOGY?

The Bayesian statistical framework is based on Bayes' Theorem and involves specific components, namely; the 'prior', 'likelihood', and 'posterior' (**Table 1**). In Bayes Theorem, the posterior is directly proportional to the product of the prior and the likelihood [6]. Many clinicians are well aware of the utility of Bayes Theorem in the interpretation of diagnostic tests, in which an a-priori probability of a disease (the revered 'clinical context') is updated by the outcome of a test, resulting in an a-posteriori probability of the disease. Such reasoning was previously elegantly described by Diamond and Forrester in their example of obstructive coronary artery disease [7]. Bayesian statistical inference of RCTs occurs in an analogous manner, where the likelihood (which summarizes the current experimental data) updates our prior beliefs to form a posterior probability distribution (**Figure 1A**).

There are several benefits to the Bayesian approach. First, the updating of prior information with current evidence is exemplary for clinical reasoning. Second, as the result of Bayesian statistical machinary, the following posterior has a probability distribution, which can be used to estimate the probability of treatment effect thresholds, including clinically relevant ones (**Figure 1B**). The Bayesian statistical methodology is therefore an example of inverse probability, or *inductive* inference, where the probability of the hypothesis or unknown treatment effect is estimated, conditional on the observed data (mathematically denoted as *P [H | data]*). The posterior probability distribution can be summarized by a mean/median treatment effect, and a 95% credible interval (CrI, or the highest posterior density [HPD] interval, please see **Table 1** and *Technical aspects*). In contrast to the frequentist confidence interval, the Bayesian 95% credible interval *is* the interval for which there is a 95% probability that it contains the true treatment effect, conditional on the current and prior data.

Bayesian analyses can therefore provide important additional insights and potentially address some of the cognitive interpretative difficulties arising with frequentist analyses [8]. Specifically, a Bayesian approach provides the information that clinicians are generally searching; an estimate of the desired treatment effect and its associated uncertainty. Furthermore, it avoids the common misconceptions of the frequentist approach including an over-dependency on the null hypothesis (nullism), the reduction of clinical trial results to a dichotomous positive or negative conclusion (based on the significance level α and the p-value in frequentism), facilitating the distinction between clinical relevance and statistical significance. An ideal scenario for the application of the Bayesian statistical framework is an expected small sample size, for example in trials assessing interventions for relatively rare diseases, or, when minor – though clinically relevant – treatment effects are foreseen and meaningful prior evidence exists that can contribute to the estimation of the posterior probability of the treatment effect [9]. **Table 2** summarizes the main features and (dis)advantages, of the Bayesian statistical methodology.

The application of Bayesian statistical inference is not new to the cardiovascular field (illustrated by the use of the Bayesian methodology in the PROTECT-AF [10], SURTAVI [11-13], EVOLUT-LR [14], PERSIST-AVR [15], ORBITA-2 [16, 17], ORBITA-COSMIC [18], and ORBITA-STAR trials [19], amongst others). Given the important guideline-shaping character of these trials, we believe a thorough understanding of statistical inference, and especially the less well-known concepts and technical details of Bayesian statistical methodology, are of utmost importance to the modern cardiovascular physician. Finally, adequate interpretation of complex trial data by virtue of the Bayesian approach (i.e. by re-analyses) can also make such trial data more digestible for patients, facilitating the shared decision-making process.

Previous reviews have introduced and summarized the concepts of the Bayesian statistical methodology [20-23], but actual guides for clinicians (i.e. on a basic level) to performing

171 such analyses are scarce [24], particular in the cardiovascular arena. Therefore, we aim to provide
172 the invested reader with the technical details and guidance to conduct a Bayesian analysis of a
173 cardiovascular trial, in addition to the explanation of the basic concepts of Bayesian statistical
174 inference. An explanation of the Bayesian terminology is provided in the glossary in **Table 1**,
175 and their technical details are discussed in more detail throughout this Review.
176
177

TECHNICAL ASPECTS IN THE APPLICACTION OF BAYESIAN METHODS

Contemporary software packages.

With the introduction of several packages implementing Bayesian methods into the *R* ecosystem, the Bayesian approach has increasingly been applied for the analysis of cardiovascular clinical trial data. Several non-*R* based programs also exist (i.e., Python, Julia, JASP, SPSS, STATA, MATLAB), and an overview of the most popular applications is provided in **Table 3**. These applications can be divided into programs with flexible properties (i.e., a high degree of code customization from the researcher's perspective), and programs with a readily-usable – though fixed – interface (i.e., low degree of code customization). Since Bayesian analyses mostly rely on random sampling from the posterior distribution, the setting of the pseudorandom number generator is necessary (set.seed() in *R*) to assure reproducibility. In addition, the number of samples, chains, and diagnostics of convergence need to be specified (see below).

Considerations for the choice of risk of estimate

An in-depth discussion on the selection of the ideal risk estimate is beyond the scope of this Review. Natural logarithmic (log) transformations of the hazard ratio (HR), odds ratio (OR), and relative risk (RR) are common risk estimates with convenient statistical properties. However, absolute effect measures are generally simpler to interpret than ratios, and can more intuitively be used to determine the minimal clinically important difference (MCID). Consequently, the absolute risk difference (ARD) can be a convenient risk estimate as well, although not well suited for time to event analyses. Finally, it must be noted that RRs/ARDs are not portable, in contrast to ORs, due to their dependency on baseline risks[25].

To ensure clarity in the current Review, effect measures are defined such that a log OR or ARD of <0 will imply benefit, while a log OR or ARD of >0 will imply harm (**Figure 1**, **Table 4**), unless otherwise stated.

The minimal clinically important difference (MCID)

Posterior probabilities can be calculated for *any* desired effect size, based on the posterior probability distribution. Rather than simply reporting a risk difference/ratio being greater or less than 0/1, it is generally of more interest to estimate the probability of a clinically relevant treatment effect. The MCID is the smallest treatment effect that is relevant to patients and/or the health care system, and should ideally be based on empirical data. Still, the MCID may differ between cultures, countries, hospitals, physicians, and patients, and suffers from a lack of standardization and consensus [22, 26, 27].

In cardiovascular clinical trials, mortality is frequently (part of) the primary endpoint, and one additional life saved per 100 patients (ARD -1%, number needed to treat 100), has occasionally been considered an acceptable MCID [8].

Determine the prior

Table 4 summarizes a tabular and graphical overview of the eligible priors. We assume priors to follow a (log) normal distribution, to facilitate the reader's intuitive interpretation.

Priors assuming no effect, with large uncertainty

I. *Non-informative* – also referred to as improper, diffuse or flat – priors are represented by uniform distributions and contain no information regarding the prior effect of the intervention under investigation. By not exerting any influence, they allow the posterior to be dominated by the study data (the

likelihood). Of note, the use of non-informative priors is *not* recommended, as even a prior with an exceptionally large variance can be informative on the equal distribution of effects (clinically plausible or not) when trial data is exceedingly sparse.

II. *Weakly informative, or vague,* priors are centred around a '0' mean effect, with a wide distribution that captures *clinically plausible values*, assigning limited density to virtually impossible values (i.e., OR > 50). These priors thereby exert a negligible influence on trial results, as they do not consider all (clinically implausible) treatment effects equally likely. Such a weakly informative prior provides a valid starting point for an objective re-analysis of a clinical trial.

These priors are specified in terms of a mean (μ), and a standard deviation (σ) reflecting the distribution, location, and degree of (un)certainty of the assumed normal distribution. Priors with a normal distribution are consequently denoted as N[μ,σ]. A weakly informative prior could therefore be formulated as N[0,2] on the log OR scale (corresponding to an OR of 1.00 with a 95% CrI of 0.02-50.00).

Reference priors

It is generally recommended to use different standardized priors in a systematic manner to evaluate the robustness of the results, irrespective of prior beliefs [22]. Consequently, reference priors can be used, which are objectively defined to represent the beliefs of physicians ranging from a sceptical, enthusiastic, or pessimistic prior belief towards a treatment. These prior distributions can be defined as follows:

I. *Sceptical priors* centre around no effect with a high degree of *certainty*, and with only a small probability of a clinically relevant treatment effect, determined by

the choice of a small σ. As proposed in **Table 4**, the MCID, and the posterior probability thereof, may be used to formulate the distribution of the sceptical prior in a standardized manner [28].

II. *Enthusiastic priors* are typically chosen to centre around a clinically relevant effect of benefit (μ = MCID) with a smaller, though still non-negligible, probability of no effect or harm. The width of the distribution is determined by σ and can reflect any desired probability of harm [29]. Analogous to the formulation of σ for the sceptical prior, the MCID can be used to determine the distribution of the enthusiastic prior (**Table 4**).

III. *Pessimistic priors* can be constructed in an analogous manner to enthusiastic priors (**Table 4**), but they centre around a clinically relevant harmful effect (μ = reversed MCID) instead, with a smaller, though still non-negligible, probability of benefit. Again, σ can be chosen any desired residual probability of benefit.

In **Table 4**, we propose standardized settings for such reference priors, based on the probability of the MCID. Still, one can be flexible with these settings, as they may change based on the treatment or patient population under investigation. Furthermore, the MCID for a certain treatment may not always be evident.

Literature-based priors

Following clinical reasoning, it can be considered counterintuitive to *not* combine prior evidence with the current evidence available at the analytical stage. This sequential approach mirrors the human learning process. Consequently, the use of prior RCT data is a foundational step in the application of Bayesian analysis to clinical trials. When determining priors based on published literature, critical appraisal skills are essential to determine risk of bias and overall

study quality, preferably by means of a standardized approach, such as Cochrane's risk of bias (RoB 2.0) tool [30].

Given the key role played by prior distributions, guidelines suggest that priors need to be formulated clearly and transparently [24, 31-33]. This transparency should include prior predictive checks, graphical representations and verification of their correctness from data input errors [34]. The use of multiple priors is strongly encouraged to assess the robustness of the final conclusions.

Cautions for the choice of priors and their careful interpretation

The choice for priors must be predefined (not post-hoc) and, when possible, substantiated by empirical evidence of the highest quality. For example, randomized evidence should ideally be analysed in the light of priors derived from previous randomized evidence, to avoid the introduction of additional bias associated with non-randomized study designs. Inappropriate selection of priors (i.e., not corresponding to the current study's patient population, treatment, or trial design) can lead to deceptive and confusing results, as the posterior is directly proportional to prior and the likelihood.

Furthermore, if the data on which the likelihood is based is sparse, the prior can have an overwhelming effect on the posterior distribution. As stated, the use of multiple priors is advocated to confirm the robustness of results. If the posterior remains insensitive to the use of adequately formulated enthusiastic and pessimistic priors, one can be quite certain of the observed treatment effect. In contrast, when the posterior changes markedly with introduction of various justified priors, results are far less reliable, and call for more robust evidence to be obtained from future trials. In that light, regulatory agencies such as the food and drug administration (FDA) recommend 'to identify as many sources of good prior information as possible' when performing a trial with a Bayesian design [33].

302

303 Determine the likelihood

304 The importance of the likelihood in Bayesian inference, and the assumptions regarding
305 the choice of the statistical models and the selection of the variables, are frequently overlooked
306 when the focus is uniquely on the presumed *prior* subjectivity. Nevertheless, the data model
307 used to create a likelihood in a Bayesian analysis is equally important to the model used in a
308 frequentist analysis and consequently subject to the same concerns. It is also imperative to
309 present the statistical data-generating models transparently, in addition to performing
310 robustness checks for prior choices.

311

312 Produce the posterior distribution

313 The Bayesian framework estimates the entire distribution of the model parameters,
314 summarized in a posterior mean/median and 95% credible interval (CrI) [34]. The use of the HPD
315 interval has been advocated, particularly for a posterior that does not follow a normal
316 distribution. The HPD interval is the shortest interval that contributes most to the posterior
317 density at a certain threshold (for example 95%) under the posterior density function (PDF). In
318 case of a normal distribution, the HPD interval equals the 95%CrI. For reasons of practicality,
319 95%CrIs will be used throughout this Review and its (re-)analyses.

320 Because of its complexity and multidimensionality, the posterior distribution and its
321 summary statistics often cannot be calculated analytically. Although this was historically an
322 important reason for favouring frequentism, the introduction of the MCMC sampling
323 techniques has greatly facilitated the numerical approximation of the posterior distribution
324 through simulations and optimization algorithms [35]. The computational aspects of the MCMC
325 techniques (such as the number of chains, iterations, warm-up phase, and methods for the
326 testing of model convergence) are well described in Bayesian guideline reporting publications

and handbooks [24, 34, 35]. Similarly to prior predictive checking, though more strict, posterior predictive checking can be performed to assess whether there seems to be any strong discrepancy between the data and the posterior results, which may in turn indicate a problem in the model selection.

Estimate the posterior probability

The posterior probability of any treatment effect of interest can be calculated from the area under the curve (AUC) of the PDF adjacent to that treatment effect. For example, the posterior probability of *any* difference between groups lies to the left of 0% ARD (or log OR 0) in the hypothetical simulated trial of **Figure 1B**, while the AUC to the left of -1% ARD is the posterior probability of the MCID in **Figure 1B**.

One can also estimate the posterior probability of an effect that lies between two effect sizes, for example the region between -1MCID and +1MCID. This area under the PDF is referred to as the 'region of practical equivalence' (ROPE) [24].

Translating treatment effects

When the probabilities of treatment effects are presented, conversions between relative to absolute measures may be helpful. If an MCID of -1.0% ARD is established, this can be evaluated and converted to a (log) OR. The risk of the control group for the outcome can be applied (in absolute percentages), from which the MCID in ARD can be used. In turn, this leads to a new (log) OR of the MCID. For example, consider a trial where an MCID of -1.0% ARD has been established, and a control group risk of 10.0% for the outcome was observed. We are then interested in the probability that the treated group has a 9.0% or less absolute risk for the outcome of interest (in relative measures an OR of 0.89, or log OR -0.12). The posterior

probability of this effect can then be estimated by the AUC to the left of log OR -0.12 under the posterior probability density function.

Bayesian reporting guidelines

Similar to reporting guidelines for clinical trials, observational studies, and diagnostic studies, guidelines for the reporting of a Bayesian analysis are available [24, 31-33]. Using these guidelines, a checklist can be followed, to ensure the adherence and reporting of the key requirements.

EXAMPLES OF THE APPLICATION OF BAYESIAN METHODS TO EXISTING CLINICAL TRIALS

Selection of trials for re-analysis

To demonstrate the application of Bayesian analyses, three contemporary trials are re-analysed using the approach proposed in this how-to-do-it guide (The *Stroke PROTECTion With SEntinel During Transcatheter Aortic Valve Replacement* [PROTECTED TAVR], the *Myocardial Ischemia and Transfusion* [MINT] trial, and the *Cholesterol Lowering via Bempedoic Acid, an ACL-inhibiting Regimen* [CLEAR] [36-38]). Of note, these trials were elicited based on their conclusions, which illustrate some of the limitations associated with frequentist statistical analyses. The statistical analysis plans of these trials prespecified the use of frequentist statistics and as such, the presented Bayesian re-analyses should not be seen as replacing - but rather complementing - the original trial analyses. Consequently, herein we aim to present the interested reader with hands-on examples and coding for the Bayesian approach. To supplement this information, we have also created a dedicated website in which the approach is outlined step-by-step (https://github.com/samuelheuts/Bayes_in_RCTs, with separate pages for the PROTECTED TAVR, MINT, and CLEAR trials). An overview of the trials and a risk of bias assessment can be found in Supplemental Table S1 and Supplemental Table S1, while **Tables 5 and 6** present the considerations for the re-analyses and summarize their results.

Confusing conditional probabilities: The PROTECTED TAVR trial

PROTECTED TAVR trial was a multicentre RCT, randomizing transfemoral TAVR patients to filter-based cerebral embolic protection (CEP) or TAVR without CEP [36]. The primary endpoint was all stroke, which occurred in 2.3% of the patients in the CEP arm vs 2.9% of patients in the control arm (-0.6%, 95%CI -1.7; 0.5%, p=0.30).

According to the authors' conclusions 'the use of CEP did not have a significant effect on the incidence of periprocedural stroke, but on the basis of the 95% CI, the results may not rule out a benefit of CEP during TAVR'. This statement is making an inference about *the probability of the alternative hypothesis* (CEP being protective) being true given the observed data, denoted as as *P (H | data)*. Yet under a frequentist framework, a parameter or hypothesis being investigated is considered to be fixed but unknown, and fixed quantities cannot be attached to probabilistic statements. Instead, the data are assumed to be variable, given an hypothesis, mathematically expressed as *P (data | H)*. The error of conflating these two conditional probabilities has bedeviled people across multiple disciplines and is often referred to as *the Prosecutor's fallacy* or *base rate neglect* [39]. In clinical medicine, this confusion may manifest itself in the misinterpretation of diagnostic test results, but the analogous situation arises in the interpretation of clinical trial results. The frequentist 95% confidence intervals (CI) indicate that, under repeated sampling, 95% of likewise calculated intervals would contain the true parameter value. Whether this particular interval does or does not include the parameter is an all-or-none proposition, and simply not known. The CI does not provide the probability of the parameter falling inside the interval, which is only available under a Bayesian framework. Given the authors' and readers' focus on this probabilistic interpretation, a proper Bayesian re-analysis of PROTECTED TAVR seems desirable.

For this re-analysis, we constructed a hierarchical literature-based prior by pooling evidence from all previous relevant RCTs (**Figure 2A**, **Table 5**). The combination of this prior and current evidence, results in a posterior median OR of 0.84 (95% CrI 0.58-1.23, **Figure 2B**). The posterior probability of *any* beneficial effect in stroke (OR <1.0) in favour of the CEP device is 81.3% (**Figure 2B, Table 6**). However, a stroke specialist expert-consensus group led by Cranston and colleagues considered a -1.1% ARD to be clinically relevant (the MCID) for stroke-reducing therapies [40, 41]. Building on the baseline risk of stroke of the control group

(2.9%), a 1.1% absolute risk reduction would result in an MCID of 0.54 on the OR scale. Consequently, based on the posterior probability distribution in **Figure 2**, the posterior probability of the treatment effect being at least this size (1.1% absolute risk reduction or an OR of 0.54) is only 5.0% (**Figure 2B**). We evaluated the robustness of this conclusion using weakly informative, skeptical, enthusiastic, and pessimistic priors. Under these reference priors, the probability of a clinically relevant effect ranged between 5.5-14.7% (**Table 6**). These estimations imply that, although some CEP treatment effect is likely, it is far less probable that this effect is clinically relevant. The robustness of these conclusions is confirmed by the relative insensitivity of the posterior to the various priors.

Currently, the BHF PROTECT TAVI is being performed in the United Kingdom [42], with a prospected sample size of >8000 patients undergoing transfemoral TAVR. The p-value resulting from this trial will be smaller than PROTECTED TAVR's p-value, if the same effect size is present (-0.6% absolute risk difference), purely by augmentation of the sample size. As discussed previously, an exceedingly small difference will always reach statistical significance in an infinitely large sample size. Consequently, it will be important to adequately interpret the clinical relevance of the to-be observed effect of this trial.

https://samuelheuts.github.io/Bayes_in_RCTs/docs/PROTECTED_TAVR.html presents a walkthrough for all these analyses (https://github.com/samuelheuts/Bayes_in_RCTs , page PROTECTED TAVR) and contains all separate figures of the various priors, including downweighting, facilitating an intuitive interpretation of the influence of the prior on the posterior.

Putting a trial into the context of previous evidence: The MINT trial

The MINT trial was a multicenter RCT hypothesizing that a liberal compared to a restrictive transfusion strategy would improve outcomes in myocardial infarction (MI) [37]. The

primary composite endpoint was all-cause mortality and MI at 30 days, which occurred in 14.5% versus 16.9% of patients, respectively (adjusted relative risk [RR] 1.15, 95%CI 0.99-1.34, p=0.07). According to the authors' conclusion: 'In patients with acute myocardial infarction and anemia, a liberal transfusion strategy did not significantly reduce the risk of recurrent myocardial infarction or death at 30 days. However, potential harms of a restrictive transfusion strategy cannot be excluded'. Of note, in MINT, an RR/OR >1.0, or ARD >0%, represents harm in the restrictive group (or conversely; benefit for a liberal transfusion strategy).

Many readers are likely disturbed by this dichotomization of the MINT trial results into the simple statement 'did not significantly reduce the risk'. Readers are likely more interested in the **probability** that a liberal transfusion policy is associated with *any* benefit, and particularly a *clinically meaningful* one. Still, as is often the case, a consensual MCID is lacking. However, the Bayesian posterior distribution probabilities can be reported at several varying thresholds corresponding to an individual reader's MCID viewpoint. The MINT-trial sample size calculations were powered to detect a +1.8% absolute risk difference, and therefore this may be considered a potential MCID. Furthermore, a +1.0% ARD has previously been applied as the MCID in studies reporting similar 'hard' clinical endpoints [43].

For this Bayesian re-analysis, we derived a literature-based prior from a self-conducted hierarchical meta-analysis of the previously reported trials investigating the same research question (**Figure 3A**, **Table 5**). Under this literature-based prior, the posterior probability of *any* difference in favor of the liberal transfusion strategy was 96.5%, while the probabilities of the MCIDs were 66.6% and 86.6% (for +1.8% and +1.0% ARD, **Figure 3B, Table 6**). These findings suggest that - despite statistical significance not being reached - a clinically relevant treatment effect of a liberal transfusion strategy is entirely likely. However, many would not consider these probabilities to definitively argue against a restrictive strategy,

or justify the adoption of a liberal strategy, but rather speak to the need for further research where the results of the MINT-trial could serve as the prior, demonstrating the Bayesian adage that '*today's posterior is tomorrow's prior*'.

Please see https://samuelheuts.github.io/Bayes_in_RCTs/docs/mint_trial.html (https://github.com/samuelheuts/Bayes_in_RCTs, page MINT) for the full rationale and coding of the MINT trial. This links also contains all separate figures of the various priors, analyses, and examples of the application of Bayesian analyses to inform personalized treatment decisions.

Comparing effectivenesses: The CLEAR trial

The CLEAR trial randomized statin-intolerant patients at high risk for cardiovascular disease to bempedoic acid or placebo [38]. The four-point composite endpoint included cardiovascular death, MI, stroke, and revascularization. After a median follow-up of 40 months, the composite primary endpoint occurred in 11.7% and 13.3% of patients (HR 0.87, 95%CI 0.79-0.96, p=0.004), respectively. This 1.6% absolute risk reduction (95%CI 0.5-2.7%) over 40 months was mainly driven by MI and revascularization, and not by death or stroke, and came with an increase of several non-cardiac adverse events.

In CLEAR, bempedoic acid's effect was studied as a substitute for statin in statin-intolerant patients. The trial found a statistically significant reduction in clinical events with administration of bempedoic acid. However, clinicians may be naturally be interested in the comaparative efficacy of bempedoic acid to statin therapy. The MCID was consequently based on a recent expert consensus evaluating statin therapy for primary and secondary prevention [44], stating a 5% absolute risk reduction of atherosclerotic cardiovascular disease (ASCVD) in primary prevention patients, and a 10% reduction in secondary prevention patients, over the course of five years. Of note, these differences in ARD arise from the primary and secondary

preventions groups' differing baseline risk, and actually correspond to similar risk reductions on relative scales such as OR/RR. The CLEAR-trial had a median follow-up of 40 months, and incorporated 30% primary prevention patients (expected -5% ARD in 5 years [44]), and 70% secondary prevention patients (expected -10% ARD in 5 years [44]), resulting in a weighted effect size threshold (defined as MCID in this analysis) of -2.8% ARD in 40 months.

Three other similar placebo-controlled RCTs were conducted before CLEAR and may be used to construct a literature-based hierarchical pooled prior (**Figure 4A**, **Table 5**). Under this literature-based prior, the posterior probability of any effect of bempedoic acid was 99.5%, while the probability of an effect similar to the ARD of statin therapy in these high-risk patients was only 0.3% (**Figure 4B, Table 6**). Furthermore, reference priors were used, under which the probability of *any* effect ranged between 98.7-99.0% (**Table 6**), and the probability of an effect similar to that of statins (-2.8% absolute risk difference) ranged between 0.1-0.2%. These findings imply that bempedoic is likely to reduce ASCVD, but it is highly unlikely that this reduction approaches the effect of statin therapy.

Nevertheless, we should emphasize that the evidence from prior statin studies stems from trials performed several years ago. Furthermore, the fact that the effect bempedoic acid does not seem to approach the effect of statins, does not render bempedoic acid *ineffective.* As such, the conclusion in the original publication by Nissen and colleagues is still valid (*treatment with bempedoic acid during a median follow-up of 40.6 months significantly lowered the risk of major adverse cardiovascular events*), but its *clinical* relevance remains to be determined. Interestingly, a recent re-analysis by the CLEAR-authors found bempedoic to be as effective as statins in reducing clinical events per 1mmol/L LDL-reduction [45]. Still, far greater LDL-reduction were achieved with statins, as compared to bempedoic acid.

A further in-depth assessment of the CLEAR trial, including the full process, coding, and additional conjugate analyses, can be found in

https://samuelheuts.github.io/Bayes_in_RCTs/docs/CLEAR.html through https://github.com/samuelheuts/Bayes_in_RCTs (CLEAR page). Notably, this link also contains a presentation of the performance of conjugate analyses, and a detailed walkthrough for the construction of a Bayesian hierarchical model.

OTHER CONSIDERATIONS

There are some other applications of Bayesian inference to RCTs worth mentioning regarding the application of Bayesian statistical inference to non-inferiority trials, stopping rules for RCTs, and Bayesian meta-analyses, which are summarized below.

- *Non-inferiority trials:* Non-inferiority trials are used when comparing a new treatment to an active control. The intricacies of this design have been previously considered [46]. In frequentist analyses, non-inferiority is met when the limit of the 95% confidence interval does not cross the non-inferiority margin, requiring specific analysis. Under the Bayesian framework, non-inferiority is met when the posterior probability of the non-inferiority margin – as in the MCID - exceeds a certain predefined value [11, 15]. Of note, unlike frequentist non-inferiority analyses, Bayesian estimations are the same as with efficacy studies, but only relate to a different interval on the x-axis under the PDF [46]. In these instances, sceptical priors may actually be considered enthusiastic, as they favour the presence of non-inferiority.
- *Stopping rules for RCTs:* The Bayesian approach, with its ability to incorporate prior knowledge, is well suited to determine whether to stop a trial early, because of either futility or efficacy, in a similar manner, based on a predefined margin. As Bayesian inference depends solely on observed and not unobserved data, no statistical penalty is required for multiple looks at the data.
- *Meta-analysis:* Finally, the Bayesian approach lends itself perfectly for a meta-analysis of RCTs [21]. With the inclusion of multiple trials, the first trial is analysed under a minimally informative prior. This analysis is then hierarchically updated by the succeeding trials. In other words, the posterior of every single analysis serves as a prior for the following analysis, resulting in a sequentially updated posterior of the

totality of the trial data. Also in such analyses, the probabilities of treatment effect sizes can be estimated, including any benefit or harm, or clinically relevant benefit and harm [47]. Of note, it is generally discouraged to use literature-based priors in meta-analyses of RCTs, as the likelihood is generally based on the totality of randomized evidence, and the use of a prior derived from a non-randomized study design introduces additional bias. Finally, Bayesian techniques are especially suited to network meta-analyses.

Limitations of the Bayesian approach

As statistical inference is the process of using data analysis to infer properties about a population parameter from noisy data samples, any inferential paradigm, including a Bayesian approach, will be accompanied by limitations. The choice of Bayesian prior is often seen as being subjective and a major limitation. However, as we have demonstrated in this Review, there are procedures to minimize this subjectivity. Bayesian analyses are associated with perceived complexity and computational intensitity, but recent computer science advances have largely overscome this limitation. In any case, this computational price is well paid for by the Bayesian ability to (i) make direct probability statements, (ii) integrate prior knowledge, (iii) have a complete picture of the uncertainty around parameter estimates, while (iv) avoiding the limitations of dichotomized reasoning that accompany null hypothesis significance testing. It should also be recalled that the frequentist paradigm is also associated with subjectivity, especially in the selection of statistical model.

CONCLUSIONS

The application of Bayesian statistical methodology to cardiovascular clinical trials facilitates an intuitive interpretation of their results, with particular emphasis on the incorporation of prior evidence and the estimation of clinically relevant treatment effects. By this approach, common cognitive biases of the frequentist approach may be mitigated, and intuitive probability estimates for treatments effects are provided. This current Review on *Methods in Cardiovascular Research and Practice* provides guidance to a clinical readership on the performance and interpretation of Bayesian analyses of cardiovascular clinical trials.

ACKNOWLEDGEMENTS

None.

FUNDING

None received.

DISCLOSURES

None.

PATIENT CONSENT STATEMENT

The authors confirm that patient consent is not applicable to this article as it comprises a review of existing literature and statistical methodology.

DATA AVAILABILITY STATEMENT

All data is openly available through a dedicated website, specifically constructed for the reproduction of our data. This site can be found through:

https://github.com/samuelheuts/Bayes_in_RCTs

EDITORIAL DISCLAIMER

Given their role as Associate Editor, James Brophy had no involvement in the peer review of this article and has no access to information regarding its peer review.

595 REFERENCES

596 **1.** Kaul S, Diamond GA. Trial and error. How to avoid commonly encountered limitations
597 of published clinical trials. *J Am Coll Cardiol.* 2010;55:415-427.
598 **2.** Brophy JM. Key Issues in the Statistical Interpretation of Randomized Clinical Trials. *Can*
599 *J Cardiol.* 2021;37:1312-1321.
600 **3.** Wasserstein RL, Lazar NA. The ASA Statement on p-Values: Context, Process, and
601 Purpose. *Am Stat.* 2016;70:129-133.
602 **4.** Wasserstein RL, Schirm AL, Lazar NA. Moving to a World Beyond "p < 0.05". *Am Stat.*
603 2019;73.
604 **5.** Goodman SN. Toward evidence-based medical statistics. 1: The P value fallacy. *Ann*
605 *Intern Med.* 1999;130:995-1004.
606 **6.** Wijeysundera DN, Austin PC, Hux JE, Beattie WS, Laupacis A. Bayesian statistical
607 inference enhances the interpretation of contemporary randomized controlled trials. *J*
608 *Clin Epidemiol.* 2009;62:13-21 e15.
609 **7.** Diamond GA, Forrester JS. Analysis of probability as an aid in the clinical diagnosis of
610 coronary-artery disease. *N Engl J Med.* 1979;300:1350-1358.
611 **8.** Brophy JM. Bayesian Interpretation of the EXCEL Trial and Other Randomized Clinical
612 Trials of Left Main Coronary Artery Revascularization. *JAMA Intern Med.* 2020;180:986-
613 992.
614 **9.** Heuts S, de Heer P, Gabrio A, et al. The impact of high versus standard enteral protein
615 provision on functional recovery following intensive care admission: Protocol for a pre-
616 planned secondary Bayesian analysis of the PRECISe trial. *Clin Nutr ESPEN.*
617 2024;59:162-170.
618 **10.** Holmes DR, Reddy VY, Turi ZG, et al. Percutaneous closure of the left atrial appendage
619 versus warfarin therapy for prevention of stroke in patients with atrial fibrillation: a
620 randomised non-inferiority trial. *Lancet.* 2009;374:534-542.
621 **11.** Reardon MJ, Van Mieghem NM, Popma JJ, et al. Surgical or Transcatheter Aortic-Valve
622 Replacement in Intermediate-Risk Patients. *N Engl J Med.* 2017;376:1321-1331.
623 **12.** Van Mieghem NM, Deeb GM, Sondergaard L, et al. Self-expanding Transcatheter vs
624 Surgical Aortic Valve Replacement in Intermediate-Risk Patients: 5-Year Outcomes of
625 the SURTAVI Randomized Clinical Trial. *JAMA Cardiol.* 2022;7:1000-1008.
626 **13.** Van Mieghem NM, Popma JJ, Deeb GM, et al. Complete 2-Year Results Confirm
627 Bayesian Analysis of the SURTAVI Trial. *JACC Cardiovasc Interv.* 2020;13:323-331.
628 **14.** Popma JJ, Deeb GM, Yakubov SJ, et al. Transcatheter Aortic-Valve Replacement with a
629 Self-Expanding Valve in Low-Risk Patients. *N Engl J Med.* 2019;380:1706-1715.
630 **15.** Fischlein T, Folliguet T, Meuris B, et al. Sutureless versus conventional bioprostheses for
631 aortic valve replacement in severe symptomatic aortic valve stenosis. *J Thorac*
632 *Cardiovasc Surg.* 2021;161:920-932.
633 **16.** Rajkumar CA, Foley MJ, Ahmed-Jushuf F, et al. A Placebo-Controlled Trial of
634 Percutaneous Coronary Intervention for Stable Angina. *N Engl J Med.* 2023;389:2319-
635 2330.
636 **17.** Simader FA, Rajkumar CA, Foley MJ, et al. Symptoms as a Predictor of the Placebo-
637 Controlled Efficacy of PCI in Stable Coronary Artery Disease. *J Am Coll Cardiol.*
638 2024;84:13-24.

18. Foley MJ, Rajkumar CA, Ahmed-Jushuf F, et al. Coronary sinus reducer for the treatment of refractory angina (ORBITA-COSMIC): a randomised, placebo-controlled trial. *Lancet.* 2024;403:1543-1553.

19. Rajkumar CA, Foley MJ, Ahmed-Jushuf F, et al. N-of-1 Trial of Angina Verification Before Percutaneous Coronary Intervention. *J Am Coll Cardiol.* 2024;84:1-12.

20. Brophy JM. Bayesian Analyses of Cardiovascular Trials-Bringing Added Value to the Table. *Can J Cardiol.* 2021;37:1415-1427.

21. Diamond GA, Kaul S. Prior convictions: Bayesian approaches to the analysis and interpretation of clinical megatrials. *J Am Coll Cardiol.* 2004;43:1929-1939.

22. 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.

23. Goligher EC, Heath A, Harhay MO. Bayesian statistics for clinical research. *Lancet.* 2024;404:1067-1076.

24. Kruschke JK. Bayesian Analysis Reporting Guidelines. *Nat Hum Behav.* 2021;5:1282-1291.

25. Doi SA, Furuya-Kanamori L, Xu C, et al. The Odds Ratio is "portable" across baseline risk but not the Relative Risk: Time to do away with the log link in binomial regression. *J Clin Epidemiol.* 2022;142:288-293.

26. Granholm A, Alhazzani W, Derde LPG, et al. Randomised clinical trials in critical care: past, present and future. *Intensive Care Med.* 2022;48:164-178.

27. 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.

28. de Grooth HJ, Elbers P. Pick your prior: scepticism about sceptical prior beliefs. *Intensive Care Med.* 2022;48:374-375.

29. Spiegelhalter DJ, Freedman LS, Pamar MKB. Bayesian Approaches to Randomized Trials. *Journal of the Royal Statistical Society.* 1994;157:357-416.

30. Sterne JAC, Savovic J, Page MJ, et al. RoB 2: a revised tool for assessing risk of bias in randomised trials. *BMJ.* 2019;366:l4898.

31. Depaoli S, van de Schoot R. Improving transparencyand replication in Bayesian statistics: the WAMBS-Checklist. *Psychological Methods.* 2017;22:240-261.

32. Althouse AD, Below JE, Claggett BL, et al. Recommendations for Statistical Reporting in Cardiovascular Medicine: A Special Report From the American Heart Association. *Circulation.* 2021;144:e70-e91.

33. https://www.fda.gov/regulatory-information/search-fda-guidance-documents/guidance-use-bayesian-statistics-medical-device-clinical-trials. Vol 2024.

34. van de Schoot R, Depaoli S, King R, et al. Bayesian statistics and modelling. *Nat Rev Methods Primers.* 2021;1.

35. Brooks S, Gelman A, Jones GL, Meng X-L. *Handbook of Markov Chain Monte Carlo*: Chapman & Hall/CRC.; 2011.

36. Kapadia SR, Makkar R, Leon M, et al. Cerebral Embolic Protection during Transcatheter Aortic-Valve Replacement. *N Engl J Med.* 2022;387:1253-1263.

37. Carson JL, Brooks MM, Hebert PC, et al. Restrictive or Liberal Transfusion Strategy in Myocardial Infarction and Anemia. *N Engl J Med.* 2023.

38. Nissen SE, Lincoff AM, Brennan D, et al. Bempedoic Acid and Cardiovascular Outcomes in Statin-Intolerant Patients. *N Engl J Med.* 2023;388:1353-1364.

39. Westreich D, Iliinsky N. Epidemiology visualized: the prosecutor's fallacy. *Am J Epidemiol.* 2014;179:1125-1127.

40. Cranston JS, Kaplan BD, Saver JL. Minimal Clinically Important Difference for Safe and Simple Novel Acute Ischemic Stroke Therapies. *Stroke.* 2017;48:2946-2951.

41. Heuts S, Gabrio A, Veenstra L, et al. Stroke reduction by cerebral embolic protection devices in transcatheter aortic valve implantation: a systematic review and Bayesian meta-analysis. *Heart.* 2023.

42. Kharbanda RK, Perkins AD, Kennedy J, et al. Routine cerebral embolic protection in transcatheter aortic valve implantation: rationale and design of the randomised British Heart Foundation PROTECT-TAVI trial. *EuroIntervention.* 2023;18:1428-1435.

43. Sabatine MS, Bergmark BA, Murphy SA, et al. Percutaneous coronary intervention with drug-eluting stents versus coronary artery bypass grafting in left main coronary artery disease: an individual patient data meta-analysis. *Lancet.* 2021;398:2247-2257.

44. Collins R, Reith C, Emberson J, et al. Interpretation of the evidence for the efficacy and safety of statin therapy. *Lancet.* 2016;388:2532-2561.

45. Lincoff AM, Ray KK, Sasiela WJ, et al. Comparative Cardiovascular Benefits of Bempedoic Acid and Statin Drugs. *J Am Coll Cardiol.* 2024;84:152-162.

46. Kaul S. Understanding the Merits and Drawbacks of Noninferiority Trials in Cardiovascular Medicine. *Can J Cardiol.* 2021;37:1378-1393.

47. Heuts S, Ubben JFH, Kawczynski MJ, et al. Extracorporeal cardiopulmonary resuscitation versus standard treatment for refractory out-of-hospital cardiac arrest: a Bayesian meta-analysis. *Crit Care.* 2024;28:217.

48. Kapadia SR, Kodali S, Makkar R, et al. Protection Against Cerebral Embolism During Transcatheter Aortic Valve Replacement. *J Am Coll Cardiol.* 2017;69:367-377.

49. Lansky AJ, Makkar R, Nazif T, et al. A randomized evaluation of the TriGuard HDH cerebral embolic protection device to Reduce the Impact of Cerebral Embolic LEsions after TransCatheter Aortic Valve ImplanTation: the REFLECT I trial. *Eur Heart J.* 2021;42:2670-2679.

50. Nazif TM, Moses J, Sharma R, et al. Randomized Evaluation of TriGuard 3 Cerebral Embolic Protection After Transcatheter Aortic Valve Replacement: REFLECT II. *JACC Cardiovasc Interv.* 2021;14:515-527.

51. Van Mieghem NM, van Gils L, Ahmad H, et al. Filter-based cerebral embolic protection with transcatheter aortic valve implantation: the randomised MISTRAL-C trial. *EuroIntervention.* 2016;12:499-507.

52. Lansky AJ, Schofer J, Tchetche D, et al. A prospective randomized evaluation of the TriGuard HDH embolic DEFLECTion device during transcatheter aortic valve implantation: results from the DEFLECT III trial. *Eur Heart J.* 2015;36:2070-2078.

53. Haussig S, Mangner N, Dwyer MG, et al. Effect of a Cerebral Protection Device on Brain Lesions Following Transcatheter Aortic Valve Implantation in Patients With Severe Aortic Stenosis: The CLEAN-TAVI Randomized Clinical Trial. *JAMA.* 2016;316:592-601.

54. Cooper HA, Rao SV, Greenberg MD, et al. Conservative versus liberal red cell transfusion in acute myocardial infarction (the CRIT Randomized Pilot Study). *Am J Cardiol.* 2011;108:1108-1111.

55. Carson JL, Brooks MM, Abbott JD, et al. Liberal versus restrictive transfusion thresholds for patients with symptomatic coronary artery disease. *Am Heart J.* 2013;165:964-971 e961.

56. Ducrocq G, Simon T, Steg PG, Investigators R. Effect of Restrictive or Liberal Blood Transfusion on Major Cardiovascular Events in Patients With Acute Myocardial Infarction and Anemia-Reply. *JAMA.* 2021;325:2506-2507.

57. Ray KK, Bays HE, Catapano AL, et al. Safety and Efficacy of Bempedoic Acid to Reduce LDL Cholesterol. *N Engl J Med.* 2019;380:1022-1032.

58. Laufs U, Banach M, Mancini GBJ, et al. Efficacy and Safety of Bempedoic Acid in Patients With Hypercholesterolemia and Statin Intolerance. *J Am Heart Assoc.* 2019;8:e011662.

59. Goldberg AC, Leiter LA, Stroes ESG, et al. Effect of Bempedoic Acid vs Placebo Added to Maximally Tolerated Statins on Low-Density Lipoprotein Cholesterol in Patients at High Risk for Cardiovascular Disease: The CLEAR Wisdom Randomized Clinical Trial. *JAMA.* 2019;322:1780-1788.

60. Carson JL, Brooks MM, Chaitman BR, et al. Rationale and design for the myocardial ischemia and transfusion (MINT) randomized clinical trial. *Am Heart J.* 2023;257:120-129.

750 TABLES

751 **Table 1.** Overview of frequentist and Bayesian terms.

Frequentist terms	*Explanation*
Frequentist inference	A method of statistical inference that views probability as one of an infinite sequence of possible data repetitions of the same experiment. In this paradigm the probability of the data is calculated under the assumption that the null hypothesis is true.
Confidence interval (95%)	The interval in frequentist inference in which, under repeated sampling, 95% of such intervals would contain the true parameter value. The 95% confidence interval is **not** a probability distribution. While providing long-term assurance that the true effect will be in the interval, this provides little assistance to any particular study at the present time, as the true effect is either in the interval, or not.
p-value	The probability of observing similar, or more extreme data, than the observed data, under the assumption that the null hypothesis is true.
Bayesian terms	*Explanation*
Bayesian inference	A method of statistical inference that views probability as a subjective state with a measure, ranging from zero to 1, of the degree of belief in a hypothesis. As such, Bayesian inference estimates the probability of an hypothesis, given the observed data.
Bayes Factor	Quantifies the support for a model over another (i.e. the alternative hypothesis over the null), regardless of whether these models are

	correct, based uniquely on the observed data and independently from subjective opinions. This is also known as the likelihood ratio.
Bayes' Theorem	The direct relation between the prior probability, the likelihood of the observed data, and the posterior probability of the hypothesis given the data, denoted as: $$P(A\|B) = \frac{P\ (B\|A)\ P(A)}{P(B)}$$
Conjugate analysis	When the likelihood function follows the same parametric distribution as the prior, direct analytical solutions without numeric (simulation) techniques are possible.
Credible interval (CrI)	In Bayesian analysis, refers to an interval within which an unobserved parameter falls with a particular probability, commonly denoted as the 95% credible interval. This **is** a probability distribution. The Bayesian 95% credible is therefore reflective of the interval for which there is a 95% probability that it contains the true treatment effect, for the current data.
Enthusiastic prior	A reference prior which assumes a beneficial effect of the intervention under investigation, with a relatively small probability of harm.
Highest posterior density (HPD) interval	The shortest interval, containing the highest density under the probability density function, commonly denoted as the 95% HDI. For the current article, we will use the HPD interval in normally distributed data, which corresponds to the 95% credible interval.

Likelihood function	The probability of the observed data for various values of the unknown model parameters. A likelihood function can be produced in both Bayesian and frequentist inference.
Markov Chain Monte Carlo	A class of algorithms for sampling from a probability distribution.
Minimal clinically important difference (MCID)	An effect size threshold considered the smallest treatment effect relevant to patients, caregivers, and society. This threshold is patient-, procedure-, and outcome-specific. The MCID should preferably be derived from consensus statements, questionnaires, or have another scientific basis. The MCID can be used in both frequentist and Bayesian inference.
Diffuse/flat/improper/ weakly-/non-/minimally-informative prior	A type of reference prior containing negligible prior information. It will therefore assume no difference between group, with a very wide (almost flat) distribution of probability. Using this prior, the posterior will be dominated by the likelihood (i.e, the trial results).
Pessimistic prior	A reference prior which assumes a harmful effect of the intervention under investigation, with a relatively small probability of benefit.
Prior (distribution)	Probability calculated from past data, theory, or judgment before the current study is analyzed.
Posterior (distribution)	Probability of a belief conditioned on both prior beliefs and current data (likelihood), quantified by Bayes' Theorem.
Probability density function (PDF)	The probability that a continuous random variable is in any range of values can be calculated as the area under a specific curve, known as the probability density function of the random variable.

Skeptical prior	A reference prior with a relatively firm believe in no effect of the intervention under investigation, with a relatively small probability of benefit.

752

753

754 **Table 2.** A summary of the features, advantages, and downsides of the frequentist an Bayesian
755 statistical methodologies.

	Frequentism	**Bayesianism**
Features		
Type of reasoning	'Deductive' *Testing whether observed data are consistent with a given and assumed hypothesis*	'Inductive' *Estimating to which degree (various) hypothesis correspond with prior beliefs and observed data*
Mathematical notation	P (data | H)	P (H | data)
Incorporation of prior data	Absent, as inferences should only be drawn based on the current data	A prerequisite to analyze the current data (likelihood) in the light of prior belief
Probability	The expected *future* frequency of events	The degree of a certain 'belief'
Quantification of probability	Based on the p-value, representing the probability of similar – or more extreme – data in infinitely repeated future trials, under the assumption that H_0 is true	The posterior probability, derived from the posterior distribution, representing the probability of the hypothesis based on the currently obtained data
Thresholds for inference	Based on the α-level (generally set on <0.05) irrespective of the type trial, patients, or intervention	No general thresholds as this can depend on the type of trial, patients, or intervention
Advantages		
	Well known to the general public and historically widely applied	Resembles clinical reasoning

	Computationally less intensive	Provides interpretable answers
	Perceived easy interpretation of the data, based on the p-value and α, leading to simple conclusions	Opportunity to incorporate prior belief/data and continuously updating information
	Perceived objectivity	Opportunity to estimate the probability of various treatment effect sizes, including clinically relevant ones
	Reproducibility	Handling of multiple testing
Disadvantages		
	Inability to incorporate prior information	Less known to the general public
	Heavily relying on H_0 (nullism)	Perceived complexity
	Inadequate interpretation of the p-value	Computationally intensive
	Leads to unnuanced conclusions	Subjectivity in the choice of the prior
	Adjustments required for multiple testing	Potential heavy influence of the prior on the posterior
	Often foregoes clinical relevance	

H_0: null hypothesis.

758 **Table 3.** Most frequently applied software and packages with their characteristics.

General software with flexible application of Bayesian inference		
Software	*Description*	*Popular packages*
R	A programming language and software environment for statistical computing and graphics. Bayesian inference is available via the *Stan* programming language and HMC sampling algorithms (cmdstanr, rstan, rstanarm, brms), or alternatively with the BUGS programming language and MCMC (rjags), or Gibbs (r2winbugs) sampling algorithms.	cmdstanr, rstan, rstanarm, brms, bayesplot, rjags, r2winbugs
Python	General-purpose programming language used for data science and scientific computing. Bayesian inference may be implemented using *Stan* or *JAGS* as extension software.	pystan, pymc, pymc3, pyjags
Julia	Relatively new programming language tailored for numerical and scientific computing. Recognized for its high performance and capability to interface seamlessly with existing C and Fortran libraries.	Turing.jl, stan.jl, mamba.jl, dynamicHMC.jl
Software with a predefined Bayesian framework		
Software	*Description*	
JASP	An open-source statistical program with modules for both frequentist and Bayesian analyses. These modules are in turn based on *R* packages. Despite the predefined nature, the program provides a great deal of flexibility, such as various settings for MCMC sampling. The module "Bayesian A/B test" is particularly applicable to the analysis of clinical trial data.	
SPSS	Widely used statistical software used for data management and statistical analysis. It has several basic models for application of Bayesian inference. There is limited flexibility in customizing the model parameters.	
STATA	A versatile statistical software package designed for statistical analysis and data visualization. STATA offers an extensive array of functions to conduct Bayesian analysis within its environment.	
MATLAB	Proprietary programming language and software environment designed for numerical computing. MATLAB is renowned for its robust array-based syntax and its capability to seamlessly interface with established C and Fortran libraries. There are multiple toolboxes within the MATLAB environment for performing Bayesian inference.	

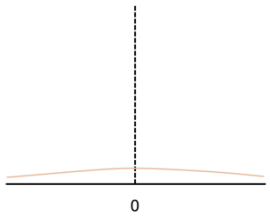
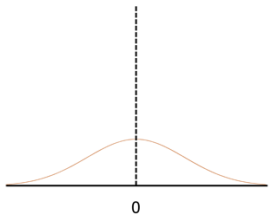
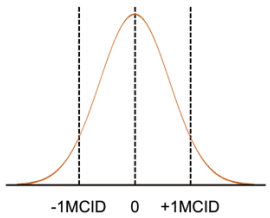
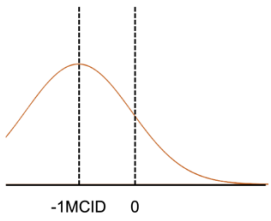
759

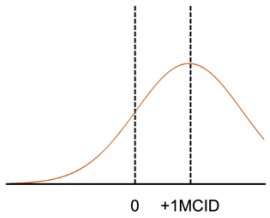
760 *HMC: Hamilton Monte Carlo, MCMC: Markov Chain Monte Carlo.*

761
762

763 **Table 4.** Proposal for standardized prior elicitation and formulation.

764

Prior type	Mean (μ) log OR	SD (σ) log OR	Rationale	Visual presentation of the prior
Priors assuming no effect, with large uncertainty				
Non-informative*	0	20	This prior assumes no difference between groups, with a probability distribution in which any effect - even extreme - is equally likely	
Weakly informative	0	2	This prior assumes no difference between groups (log OR μ = 0), with a wide distribution that captures plausible effect sizes (log OR σ = 2)	
Reference priors				
Sceptical	0	Based on 10% probability of MCID	The sceptical prior assumes no difference between groups with a relatively high degree of certainty, reflected by the mere 10% probability of a clinically relevant effect	
Enthusiastic	-1MCID	Based on a 30% probability of 'harm'	The enthusiastic prior centres around the MCID for benefit, and only assumes a 30% probability of 'harm'.	

Pessimistic	+1MCID	Based on a 30% probability of 'benefit'	If there is reasons for pessimism, the pessimistic prior may centre around the MCID for harm, and only assumes a 30% probability of 'benefit'.	
Literature-based priors				
Derived from one RCT	Data	Data	This prior is based on the data of one specific RCT to inform the current analysis[+]	Depending on the trials
Derived from multiple RCTs	Pooled data	Pooled data	This prior is constructed by pooling of multiple available RCTs	Depending on the trials
Derived from observational data	Data	Data	This prior is informed by potentially important observational data, when the is a lack of previous RCTs. It should be noted that the use of such a prior is debatable in the analysis of an RCT. To adjust for the observational character of such studies, penalties (downweighting) can be applied to the prior	Depending on the trials

765

766 We assume a normal distribution of the priors in this table, for basic interpretation. Settings are
767 presented as log odds ratios. Also, the gradation in 'belief' of the various priors can be adapted
768 by variations in the SD. A smaller SD constitutes a firmer belief, while a larger SD reflects a
769 higher degree of uncertainty.

770 *Non-informative priors can also interchangeably be termed as 'flat' priors, or 'diffuse' priors.

771 [+]Multiple priors informed by various singular RCT can be used, or a pooled analysis can be
772 applied, as in the lower row.

773 *MCID: minimal clinically important difference, OR: odds ratio, RCT: randomized controlled*
774 *trial, SD: standard deviation.*

775

776 **Table 5.** Considerations for the current Bayesian re-analysis including references for the
777 literature-based priors.

Considerations	PROTECTED-TAVR [36]	MINT [37]	CLEAR [38]
Relevant prior information	Six smaller RCTs were previously performed [48-53]	Three similar RCTs in patients with MI were performed [54-56]	Three similar RCTs were previously performed [57-59]
Literature-based prior	Log OR [-0.02, 0.35] OR []	Log OR [-0.01, 0.51]	Log OR [-0.16, 0.27]
MCID (ARD, NNT)	-1.1% (91)	+1.8% ARD (56)* +1.0% (100)	-2.8% (36)
MCID rationale	Based on an expert-consensus by stroke specialists [40]	The first MCID was based on the expected treatment effect*, and a second MCID was arbitrarily set at 1.0% ARD	Based on the weighted mean ARD of statin therapy for primary and secondary prevention [44]
Weakly/Non-informative prior	Log OR [0, 2] OR [1, 7.4]	Log OR [0, 2] OR [1, 7.4]	Log OR [0, 2] OR [1, 7.4]
Skeptical prior	Log OR [0, 0.39]	Log OR [0, 0.25]	Log OR [0, 0.21]
Enthusiastic prior	Log OR [-0.49, 0.94]	Log OR [0.33, 0.62]	Log OR [-0.27, 0.51]
Pessimistic prior	Log OR [0.49, 0.94]	Log OR [-0.33, 0.62]	Log OR [0.27, 0.51]

778

779 *Based on the expected treatment effect of the MINT-trial [60].

780 *ARD: absolute risk difference, OR: odds ratio, NA: not applicable, NNT: number needed to*
781 *treat, RCT: randomized controlled trial.*

782 **Table 6.** Outcomes of the included re-analyzed trials under the Bayesian framework.

Outcomes	PROTECTED TAVR [36]	MINT [37]	CLEAR [38]*
Corresponding Figure	2	3	4
Studied MCID for the trial in ARD	-1.1% ARD	+1.8% ARD	-2.8% ARD
MCID in OR	0.49	1.13	0.76
Literature-based prior			
Posterior probability any difference	81.3%	96.5%	99.5%
Posterior probability MCID	5.0%	71.1%	0.2%
Weakly/Non-informative prior			
Posterior probability any difference	86.1%	97.3%	98.9%
Posterior probability MCID	14.1%	71.0%	0.2%
Skeptical prior			
Posterior probability any difference	82.2%	92.2%	98.7%
Posterior probability MCID	5.5%	31.9%	0.1%
Enthusiastic prior			
Posterior probability any difference	88.0%	97.4%	99.0%
Posterior probability MCID	14.7%	70.1%	0.2%
Pessimistic prior			
Posterior probability any difference	82.4%	94.9%	98.7%
Posterior probability MCID	10.3%	56.6%	0.1%

783

The posterior probabilities presented in this Table are based on separate analyses under various priors of the re-analyzed trials. First, the trial is analyzed under the predefined literature-based prior, producing the probability density function of the posterior (as in Figure 1B). Then, the AUC of the PDF to adjacent to log OR 0 is calculated (representing the probability of any difference). In addition, the posterior probability of the predefined MCID is calculated as well, represented by the AUC under the PDF adjacent to the MCID. Finally, both these posterior probabilities are calculated under the non-informative, skeptical, enthusiastic, and pessimistic prior (where applicable).

* MCID presented as the probability of an effect similar to the absolute risk reduction of statins to prevent atherosclerotic cardiovascular disease in high risk patients [44].

ARD: absolute risk difference, AUC: area under the curve, MCID: minimal clinically important difference, NA: not applicable, OR: odds ratio, PDF: probability density function.

FIGURE LEGENDS

Figure 1. Graphical presentation of the interplay between the prior, likelihood, and posterior (A), with the calculation of the posterior probability of any effect, and of the MCID (B), in a hypothetical trial.

For this hypothetical trial, the fictive MCID was set at -1% ARD (benefit) and +1% ARD (harm). The AUC to the left of '0% ARD' is the posterior probability of any beneficial effect, while the AUC to the left of the -1%ARD, is the probability of the MCID for benefit. The AUC to the right of 0 is the probability of any harm while the probability of a the MCID in harm is the AUC to the right of +1% ARD. Finally, the AUC between -1% ARD and +1% ARD is known as ROPE (region of practical equivalence) in this figure.

ARD: absolute risk difference, AUC: area under the curve, MCID: minimal clinically important difference, ROPE: region of practical equivalence.

Figure 2. Components of the re-analysis of the PROTECTED TAVR trial.

A: Construction of the literature-based prior, from a hierarchical meta-analysis of RCTs published before PROTECTED TAVR. B: The combination of the prior data and the likelihood results in an updated posterior (OR 0.84, 95%CrI 0.58-1.23). The posterior probability of any difference is 81.3%, while the posterior probability of the MCID (OR 0.54) is 5.0%.

CEP: cerebral embolic protection, CrI: credible interval, MCID: minimal clinically important difference, OR: odds ratio.

Figure 3. Analyses to derive the literature-based prior (A) with estimation of the posterior under this literature-based prior (B) for the re-analysis of the MINT-trial.

A: Construction of the literature-based prior, derived from the totality of randomized evidence before the MINT-trial. B: analysis of MINT under the literature-based prior. The posterior probability of any difference is 96.5%, whule posterior probability of the MCID is 71.1%.
MCID: minimal clinically important difference, OR: odds ratio.

Figure 4. Construction of the literature-based prior (A) with estimation of the posterior (B) for the re-analysis of the CLEAR trial.
A: Construction of the literature-based prior, derived from previous randomized trials before CLEAR. B: analysis of CLEAR under the literature-based prior. The posterior probability of any difference is 99.5%, while the posterior probability of the MCID is 0.2%.
MCID: minimal clinically important difference, OR: odds ratio.