Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Uptake of Bayesian robust information borrowing in Phase II clinical trials: a scoping review

Vivienn Weru

vivienn.weru@dkfz-heidelberg.de

German Cancer Research Center (DKFZ)

Silvia Calderazzo

German Cancer Research Center (DKFZ)

Leonhard Held

University of Zurich

Manuel Wiesenfarth

Cogitars GmbH

Annette Kopp-Schneider

German Cancer Research Center (DKFZ)

Research Article

Keywords: External information; Robust mixture prior, Power prior, Bayesian hierarchical model, Effective sample size, Basket trials

Posted Date: June 2nd, 2026

DOI: https://doi.org/10.21203/rs.3.rs-9783037/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

Uptake of Bayesian robust information borrowing in Phase II clinical trials: a scoping review

Vivienn Weru[1,2*†], Silvia Calderazzo[1†], Leonhard Held[3], Manuel Wiesenfarth[4], Annette Kopp-Schneider[1]

[1]Division of Biostatistics, German Cancer Research Center (DKFZ), Heidelberg, Germany.
[2]Medical Faculty, University of Heidelberg, Heidelberg, Germany.
[3]Epidemiology, Biostatistics & Prevention Institute (EBPI), University of Zurich, Zurich, Switzerland.
[4]Cogitars GmbH, Heidelberg, Germany.

*Corresponding author(s). E-mail(s): vivienn.weru@dkfz-heidelberg.de;
†These authors contributed equally to this work.

Abstract

Background and objective: Incorporation of external information in clinical study designs has received increasing attention in recent years. When external data can be considered consistent with the data collected in the current trial, efficiency gains and a reduction in the required sample size can be achieved. This makes it particularly attractive when, *e.g.*, recruitment of patients is difficult. However, unexpected inconsistency cannot be ruled out, and various robust borrowing approaches have been developed to limit potential losses. These include, for example, the power prior, meta-analytic predictive prior, robust mixture prior, and Bayesian hierarchical model. However, their use in actual clinical trials remains underexplored. In this scoping review, we investigated the use of external information in Phase II clinical trials.

Methods: Publications were identified by literature search in PubMed central, the Cochrane Central Register of Controlled Trials and complemented with a citation search in some of the main methodological papers on Bayesian borrowing in clinical trials.

Results: Of the 1282 articles retrieved, 37 were eventually included. We investigated, among others, whether and how the suitability of external data was assessed, the borrowing method used, and the specification and justification of tuning parameters that influence the degree of borrowing. Fixed downweighting was found to be the most commonly used method for external information borrowing, with only slightly more than a third of the studies providing details on its implementation.

Conclusion: Overall, we found poor reporting of various aspects relating to Bayesian borrowing, signaling a gap between the methodological developments and the practical uptake and reporting of such methods.

Keywords: External information; Robust mixture prior; Power prior; Bayesian hierarchical model; Effective sample size; Basket trials.

1 Introduction

The use of external data to augment clinical trial information has gained popularity in recent years. Such an approach is typically referred to as 'external information borrowing' and holds the potential to significantly increase the efficiency of a given trial. Areas in which borrowing has seen increasing application are those in which there is interest or need to reduce the sample size, *e.g.* pediatric trials, control or placebo arms, or in the analysis of subgroups, for example, multiple study arms in basket trials. Basket trials are a specific type of a clinical trial where a single therapy is evaluated across different types of tumors with a common mutation (see *e.g.* [21, 22, 37]). In such a setting, rather than conducting multiple independent trials, only one study protocol is needed.

However, external data to be incorporated into the current trial requires careful selection and guidelines have been provided to aid in this selection, for example, by Pocock [39]. In addition, when covariates are available, regression or propensity score approaches can be used to adjust for measured confounding [3, 11, 16, 18, 24, 30]. However, unobserved heterogeneity between current and external data can still be present, and for this, robust Bayesian and frequentist methods have been developed

(*e.g.* [9, 10, 20, 24, 35, 44]) that downweight external information in case of such heterogeneity. Although fixed downweighting approaches are not considered "robust" as they do not adapt to unforeseen prior-data conflict, we included them in this review since they do acknowledge potential heterogeneity and adjust beforehand for this.

Although many methods have been developed to facilitate external data borrowing, their use in actual clinical trials is understudied. Systematic reviews are available that have focused on different aspects of Bayesian methods in clinical trials. For example, the systematic review of Barret et al. [4] investigated the type of priors used in Phase II and III randomized clinical trials, as well as the use of decision thresholds and whether such thresholds were pre-specified. In their systematic review, Ferreira et al. [14] investigated reporting of replicability items in Phase III clinical trials, focusing on BayesWatch [48] and ROBUST [50] guidelines. The more exploratory nature of Phase II lends itself more naturally to the use of external information, and this has not been investigated systematically so far, to the best of our knowledge. This scoping review aims to fill this gap. We identified Phase II clinical trials that use robust Bayesian borrowing. We investigated their reporting of the used borrowing method and respective parameter settings. We focused on a scoping review to highlight gaps in current knowledge concerning all potential aspects related to the use of Bayesian robust borrowing approaches.

2 Methods

The objective of this scoping review is to investigate the use and reporting of Bayesian robust borrowing methods in clinical trials, with a focus on Phase II trials. The review followed the PRISMA guideline and the PRISMA-ScR [52] checklist is provided as supplementary material.

2.1 Data sources

A search of the literature was performed in PubMed Central and the Cochrane Central Register of Controlled Trials until 22 May, 2025. The search terms are listed in Supplementary Appendix A. In addition, a search for citations in relevant statistical methodological articles was done to identify clinical trials using these methods. The list of the methodological articles whose citations were searched is included in Table A.3.

2.2 Inclusion and exclusion criteria

We included Phase II clinical trials with a publication in English that involved robust Bayesian information borrowing for the analysis of the primary endpoint. Trials that used Bayesian information borrowing without discounting or consistency checks based on any of the Pocock's criterion and/or available covariates were excluded. Also excluded were studies that used only non-informative/weakly informative priors or used robust Bayesian information borrowing only for sensitivity analyses or re-analyses of a past trial.

All titles and abstracts were independently screened by VW and SC, thereafter independently reviewing the full-text of relevant articles. Conflicts were discussed.

2.3 Data extraction

Data was extracted from the articles by VW using a standardized extraction form. SC checked the data for a random 21.6% ($n = 8$) of the articles (based on alphabetical author order) and an additional 8.1% ($n = 3$) in which the information required for extraction was ambiguous. The complete list of data extracted is given in Table A.4. The extracted data was descriptively analyzed using R (version 4.3.0).

3 Overview of key aspects on borrowing external data

When borrowing external data, various key aspects must be considered. These aspects are outlined below. A detailed methodological review for each of these aspects is provided in Supplementary Appendix B.

3.1 External data

In borrowing data, various aspects of the external data need to be considered: the source of the borrowed data, suitability of the external data and target of borrowing. The sources of borrowing can be external or internal. External sources include previous trials, Real-World evidence (RWE) or information elicited from experts. Internal sources include borrowing between arms within a trial, for example, in basket trials. Distinguishing between different sources of external data is important as different assumptions, for example when evaluating operating characteristics, may be applicable depending on the source of data. With respect to the target of borrowing, external information may be borrowed for the control arm of a study, leading to so-called hybrid-control designs, or the external information available may be on the treatment effect directly. Other times information is borrowed between the treatment arms of the same trial. Another key point is that the external data need careful selection. Pocock [39] outlined six conditions for comparability, including that treatment and its evaluation, as well as patient eligibility, must be the same between external and current studies. The detailed list of Pocock's conditions is provided in Supplementary Appendix B.3.

3.2 Borrowing methods and assessment

Even if the external data has been careful selected, there might still be differences between the current and external information. To mitigate risks associated with prior-data conflict, different methods have been proposed for incorporating the external

data. These include the power prior approach [24, 25], the meta-analytic predictive (MAP) prior [35], the robust mixture prior [44], Bayesian hierarchical model [6, 51] among others. A detailed description of these borrowing methods is provided in Supplementary Appendix B.4.

All borrowing methods rely on parameter settings that regulate the amount of information borrowed. The power prior downweights the likelihood of the external data according to a power parameter, while for Bayesian hierarchical models, the between-study variance parameter determines the degree of borrowing such that if it is zero all the data is pooled while if it goes to infinity, no data is borrowed. The MAP prior is based on a hierarchical model with a between-study heterogeneity parameter influencing the borrowing. The robust mixture prior is a weighted combination of an informative prior, derived from the external data, and a robust component to robustify the prior in case of inconsistencies between the current and external data. Specification of the prior weight as well as parameters for the robust component is, therefore, also needed.

Adjusting for covariates, *e.g.* through linear regression or matching approaches, to correct for measured confounding between the current and external data can also be used. Nonetheless, the dynamic borrowing methods above help remedy problems associated with unmeasured confounding.

Effective sample size (ESS)

In this context of Bayesian information borrowing, the informativeness of a prior can be translated into a more intuitive metric called the prior effective sample size (ESS) which represents how many observations the prior is worth. In simple settings, the ESS can be easily obtained. For example, for a binomial likelihood with a conjugate $Beta(a, b)$ prior, the ESS is easily obtained as $(a + b)$. For more complex settings, different approaches have been suggested for calculating the ESS of a prior, with computations becoming non-trivial for non-conjugate priors. A description of the common

ESS methods is provided in Supplementary Appendix B.5. For some prior distributions, different methods might lead to different ESS values. It is, therefore, essential to report the method used to obtain the ESS, especially for non-conjugate settings.

3.3 Reporting guidelines in Bayesian analysis

Clear reporting of Bayesian analysis, as with any statistical analysis, is crucial to ensure understanding and transparency of the analysis to readers. When an informative prior is used, clear explanation of the choice of prior is needed. To aid in standardized reporting of Bayesian analyses, several guidelines have been developed. These include ROBUST (Reporting Of Bayes Used in clinical STudies)[50], BayesWatch [48] and BARG (Bayesian Analysis Reporting guidelines) [29]. More details on these guidelines are highlighted in Supplementary Appendix B.6. None of the above guidelines, however, is targeted to our application of information borrowing in clinical trials. When using Bayesian borrowing methods, for example those outlined in Section 3.2, reporting and justification of corresponding tuning parameters and assumptions is essential.

4 Results

4.1 Characteristics of the trials

Of the 1282 articles retrieved, 37 were finally included. Figure 1 shows the PRISMA flow chart of the articles selected and included. The excluded reports are in the order of exclusion. The list of included articles is provided in Supplementary Appendix C. To obtain the relevant information for this review, we also checked supplementary material of the articles, as well as `clinicaltrials.gov` for the accompanying protocols. Table 1 shows a summary of the year of publication of the articles. Of these, 86.5% ($n = 32$) were reports while 13.5% ($n = 5$) were protocols of yet unpublished studies. Most of the trials were conducted in oncology (40.5%, $n = 15$), 16.2% ($n = 6$) were in autoimmune diseases, while 10.8% ($n = 4$) were in infectious diseases with 2

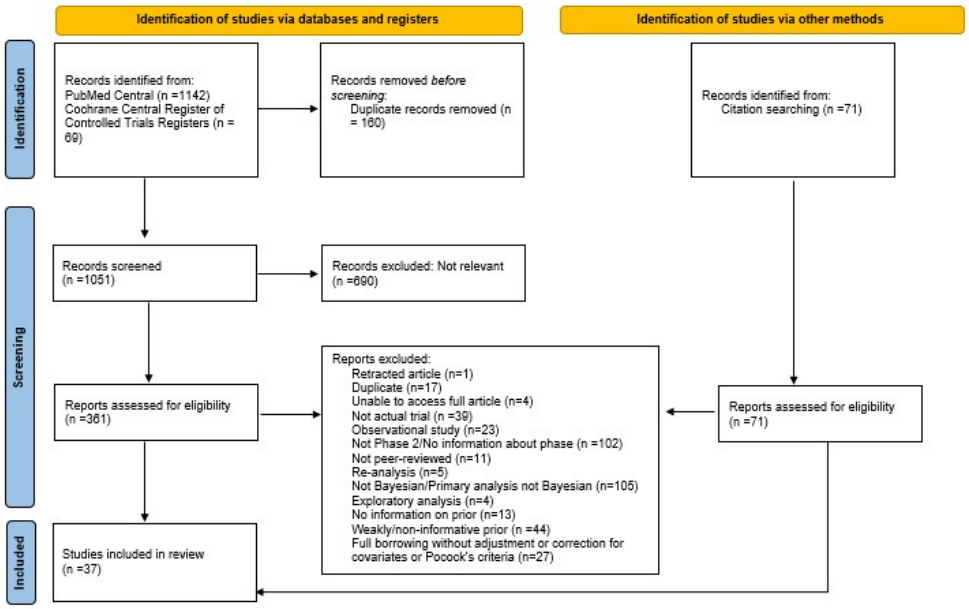


Fig. 1 PRISMA flow-diagram of the articles selected and included.

of those investigating Covid-19. Of the included trials, 37.8% ($n = 14$) had a master protocol, mainly basket trials ($n = 9$).

Table 1 Year of publication (N=37)

Publication year	n (%)
Before 2015	4 (10.8)
Between 2015 and 2020	18 (48.6)
After 2020	15 (40.5)

4.2 Key aspects on borrowing external data

As mentioned above, various key aspects must be considered when borrowing external data. In the following, we summarize their distribution in the reviewed studies.

4.2.1 Source of borrowed data

In this review, we distinguish between studies that borrow information from external and internal data sources. A majority of the trials borrowed data external to the current trial (62.2%, $n = 23$). Of these, 78.3% ($n = 18$) explicitly cited the sources of the external data while 21.7% ($n = 5$) did not. Seven (30.4%) of the articles that borrowed data external to the current trial explicitly mentioned the source in the protocol, 13.0% ($n = 3$) did not, while it was unclear for 56.5% ($n = 13$) of the articles mainly due to the lack of availability of the protocol online. External data can come from different sources, *e.g.*, historical/concurrent trials, real world evidence (RWE), or expert opinion. Of the articles citing the source of the external data, none used RWE or elicited the prior information from experts. The studies that borrowed internal information were mainly basket trials.

4.2.2 Target of borrowing

Most of the trials borrowed data to the control arm (56.8%, $n = 21$) while 32.4% ($n = 12$) borrowed between arms of the same trial, with the latter mainly being basket trials. When borrowing information on the treatment effect, one study borrowed separately information on the treatment and control arm parameters [46], one borrowed information on the control arm and the treatment effect [8], while two trials only borrowed information on the treatment effect [1, 15], one of those [15] being a single-arm trial.

4.2.3 Assessment of external data suitability

The external data need careful selection. Table 2 shows a summary of the approaches used to assess the suitability of the external data. In relation to Pocock's criteria, the majority of the trials (73.9% ($n = 17$)) assessed at least one of the conditions. One study [17] stated that prior-data conflict would be assessed using graphical displays,

one performed pair-matching [36], while one study [43] performed a prior-predictive check to assess the concordance between the data and the prior. Details were not provided for 21.7% ($n = 5$) of the studies.

Table 2 Assessment of data suitability for trials that borrowed external information (N=23). Note some categories are not mutually exclusive. The detailed list of Pocock's conditions is provided in Supplementary Appendix B.3.

Assessment of the suitability of external data	n (%)
Pocock's criteria	
Same treatment	17 (73.9)
Same patient eligibility	4 (17.4)
Same treatment evaluation	15 (65.2)
Same patient characteristics	7 (30.4)
Same organization and clinical investigators	0 (0)
No other indications	1 (4.3)
Used graphical displays	1 (4.3)
Used pair-matching	1 (4.3)
Used prior-predictive check	1 (4.3)
Not mentioned	5 (21.7)

4.2.4 Borrowing method used

As mentioned in Section 3.2, different methods can be used to incorporate external data. Below we present the borrowing methods identified in the reviewed studies. Again, we distinguish between studies that borrowed data external to the current trial and that borrowed internally. Table 3 shows a summary of the borrowing methods used. The majority of the trials (39.1%, $n = 9$) that borrowed externally used a fixed downweighting approach that included the fixed power prior [8, 46]. The robust mixture prior was used in 21.7% ($n = 5$) of the studies [8, 31, 32, 34, 40, 41], for three of which the informative component of the mixture prior was a meta-analytic predictive (MAP) prior. One study [8] used an approach that combined the mixture prior and a power prior to incorporate an aggregate treatment effect as well as individual historical control data, respectively. Three trials used the MAP prior alone [2, 13, 23] without robustification.

Table 3 Borrowing method used (N=37)

Methods used by trials that borrowed externally to current study	n (%)
Fixed downweighting/fixed Power prior	9 (39.1)
Robust mixture prior (w/ or w/o Meta-analytic predictive (MAP))	5 (21.7)
Meta-analytic predictive (MAP) (without robustification)	3 (13.0)
Bayesian hierarchical model	1 (4.3)
Robust MAP and fixed Power prior	1 (4.3)
Matching external control to current treatment based on covariates	1 (4.3)
Full borrowing with covariates adjustment	1 (4.3)
Dynamic borrowing/Robust assessment but unclear which method	2 (8.7)
Methods used by trials that borrowed from within the same trial	**n (%)**
Bayesian hierarchical model	10 (71.4)
Fixed downweighting/fixed Power prior	2 (14.3)
Fujikawa's design	1 (7.1)
Dynamic borrowing but unclear which method	1 (7.1)

Two trials used covariate adjustment to correct for bias [36, 42]. The majority of the trials that borrowed within the same trial used Bayesian hierarchical models, with a majority of these being basket trials. Of these, four trials [38, 45, 47, 49] clustered groups with similar response rates. One basket trial [27] reported using Fujikawa's design, although it did not provide comprehensive details.

4.2.5 Parameters influencing degree of borrowing

All borrowing methods are based on parameters that regulate the degree of borrowing as summarized in Table B.5. For trials using Bayesian hierarchical models, the majority (63.6%, $n = 7$) did not provide further information on how the heterogeneity parameter was estimated, while the rest assigned an inverse gamma prior to it. Of the trials using fixed downweighting, only 33.3% ($n = 4$) explained how the discounting weight was arrived at. Amongst these, two based the decision on simulations evaluating operating characteristics (specifically Type I error (TIE) rate) and two elicited the weight from experts [1, 43]. For example, in the ODYSSEY trial investigating dolutegravir-based antiretroviral therapy in children with HIV infection [1], the borrowing weight α was elicited from 13 experienced pediatricians upfront. Brard et al. [8] used $\alpha = 0.3$ for the power parameter of the fixed power prior, McIlroy et al. [33]

predefined a weight of 0.75 for the power prior while the study of Seccia et al. [46] investigated different power parameters for the power prior ($\alpha = 1, 0.5, 0$).

Only one of the included trials [41] that used the MAP, either alone or robustified, provided comprehensive details on the between-trial heterogeneity parameter prior and the derivation of the MAP prior, including the approximation of the MAP prior by a mixture of conjugate distributions. This study used a half-normal prior for the between-trial heterogeneity parameter. The robust mixture prior requires specification of the prior weight, location and dispersion for the robust component, and these parameters need careful thought [53]. All the studies that used the robust mixture prior did specify the prior weight assigned, with three studies using a prior weight of 0.5 while the other three studies assigned a prior weight of $0.1, 0.95$ and 0.90 to the informative component of the mixture. However, 50% ($n = 3$) did not give information on how the prior weight was chosen, two based the decision on simulations that evaluated operating characteristics, while one mentioned that they chose the 50% prior weight to avoid prior domination after considering the size of the historical studies and the placebo group. Regarding the informativeness of the robust component of the mixture prior, only three provided the exact prior specification. For trials with a normal endpoint and using the robust mixture prior, one used the same location as historical data, one had the location at the null hypothesis while one gave no information.

4.2.6 Effective sample size

As outlined in Section 3.2, the ESS provides a more accessible measure of the informativeness of the prior. For trials that borrowed externally to the current trial, 82.6% explicitly stated the ESS or, if not explicitly stated, the ESS can be easily obtained because a conjugate non-mixture model was used. Seven trials (36.8%) explicitly stated the ESS but since the model was not conjugate, it was not clear which ESS method they used. Figure 2 shows the relative ESS [19]. Of the trials where an ESS could

be obtained, eight trials had a relative ESS greater than 0.5. One trial [36] had only historical controls but matched patients to the treatment arm to correct for bias.

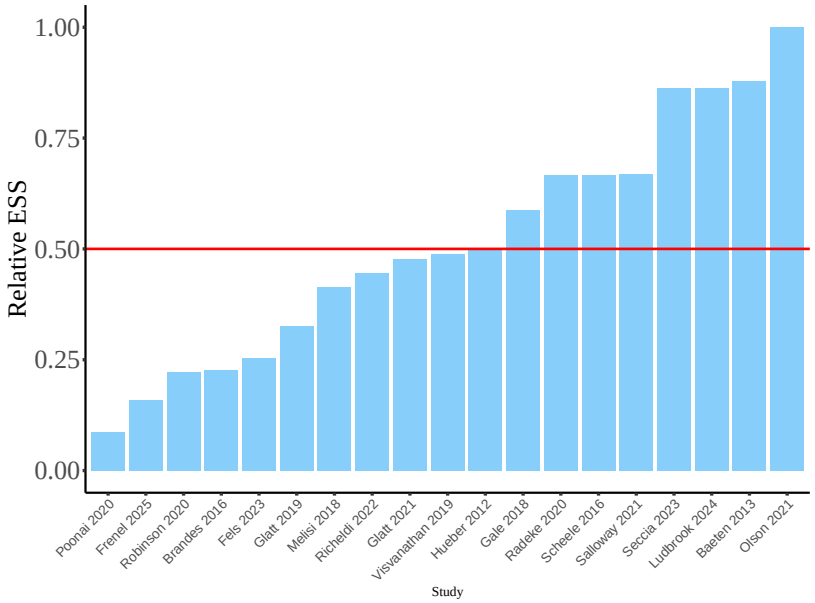


Fig. 2 Relative ESS of studies borrowing external data, obtained by dividing the prior ESS by the sum of the prior ESS and current sample size.

4.2.7 Success criteria used

To conduct hypothesis testing in Bayesian designs, one common success criterion is the posterior probability that the true treatment benefit exceeds some threshold. For example, if the alternative hypothesis is $\theta > 0$, where θ is the treatment effect parameter, a trial would be declared successful if $\Pr(\theta > 0 \,|\, \text{data}) > \kappa$, where κ is the minimum probability threshold (*e. g.* 95%). When using the posterior probability for success criterion, pre-specification of κ is needed. Another approach includes checking whether the posterior credible interval at some pre-specified κ includes θ_0, where θ_0 is the null hypothesis value. Most of the studies (86.5%, $n = 32$) used the posterior probability as a success criterion with one of those relying on a normal approximation

for the posterior. Two trials used credible intervals, while the success criterion was unclear for three trials.

4.2.8 Essential items to report in Bayesian analyses

Although various guidelines have been proposed for reporting a Bayesian analysis of a clinical study, in this review we focus on a list of seven items known as ROBUST, as outlined in Section 3.3, pertaining to the prior distribution, analysis, and results. The results are shown in Table 4. For studies that used the Bayesian hierarchical model, details of the prior pertain to the heterogeneity parameter. Specification of the analytical technique and sensitivity analysis on priors were the least reported items (21.6% and 29.7%, respectively). A rationale for the prior specification was provided in 72.9% of the articles, but an explicit specification of the parameter values of the prior was reported in 56.7% of the articles.

Table 4 Reporting of ROBUST items (N=37)

Items	N	%
Prior distribution		
Specified	21	56.7
Justified	27	72.9
Sensitivity analysis	11	29.7
Analysis		
Statistical model	35	94.6
Analytical technique	8	21.6
Results		
Central tendency	29	78.4
Standard deviation or credible interval	27	72.9

5 Discussion

We conducted a scoping review to investigate the uptake of Bayesian robust borrowing in Phase II clinical trials. We evaluated various aspects of borrowing, for example, the borrowing method used and the reporting of associated parameters. A potential

limitation of this scoping review is that the search was performed only in PubMed central and the Cochrane Central Register of Controlled Trials. Nevertheless, this search was complemented by a citation search in the most well-known statistical methodological papers on Bayesian borrowing.

When external data are incorporated into a clinical trial, they need to be carefully selected. Guidelines have been proposed to aid in such selection, for example Pocock's criteria [39]. Although some studies indicated that the historical data was obtained from studies similar to the current study, many did not provide additional details.

Different methods can be used to formulate the external data as a prior for the current data. Although fixed borrowing methods are simpler, they preclude adaptation to unforeseen prior-data conflict. However, even though dynamic borrowing methods adapt to heterogeneity conflict, they still incur losses in terms of estimation and testing performance if moderate conflict is present. Furthermore, most of the dynamic borrowing methods rely on tuning parameters that impact how much information is discarded based on a given observed drift.

We found poor reporting of the parameters that regulate the amount of information borrowed in terms of both providing explicit values and rationales for such values. For example, when using the Bayesian hierarchical model, many studies did not give details about the heterogeneity parameter that drives the borrowing. As detailed in Supplementary Appendix B.4, this parameter is crucial. For studies using a fixed downweighting approach, crucial details on the downweighting factor were mainly unavailable. For the robust mixture prior, the prior weight and the robust component drive the borrowing, and these choices need careful thought as detailed in Weru et al. [53]. Again, full details on such parameters were not reported in the majority of the studies.

In addition to the values of the tuning parameters, justification for such choices needs to be provided. When justification was given, the majority of the studies evaluated operating characteristics, followed by an elicitation from clinical experts. In this context, TIE rate and power are the operating characteristics most frequently evaluated. In basket trials, in addition to the basket-specific TIE rate, the family-wise TIE rate is also usually investigated. If a simulation study is used to select a borrowing method or parameter value, appropriate pipelines and reporting should be followed to ensure transparency and fairness for all evaluations [28]. Although operating characteristics are usually evaluated in a frequentist framework, Bayesian metrics have also been proposed [5, 7]. Although none of the studies looked at the latter, it is expected that these evaluations will become more common following the recent FDA draft guidance "Use of Bayesian Methodology in Clinical Trials of Drug and Biological Products Guidance for Industry" [12], which paves the way towards acceptance of the Bayesian paradigm in drug development. Such evaluations will require specification and reporting of additional quantities, most notably the design prior used to generate the simulated data.

Although some of the aspects we investigated are specific to clinical trial designs that involve robust borrowing, general guidelines for the reporting of Bayesian analyses also exist. We focused on ROBUST [50], although other guidelines are available. Despite some of the guidelines being available for many years, studies are yet to fully embrace them. Various important considerations and aspects in the context of external data borrowing have also been discussed in ICH and FDA guidelines. For example, in pediatric extrapolation, the ICH guidance E11A [26] outlines important aspects to consider, including evaluating the definition and course of disease between the reference and the target populations, as well as considerations for extrapolation

of safety. Further aspects include emphasis on deciding on the sources of information in consultation with regulatory authorities when possible, pre-specification of the borrowing method, and incorporation of sensitivity analyses to evaluate impact of different prior assumptions. The recent FDA draft guidelines [12] detail the necessary information that should be submitted to the FDA at the design stage of a trial that aims to use Bayesian methods. These include explicitly specifying the prior parametrization. When using informative priors, the FDA draft guidance asks for justification for borrowing and why the available information is considered relevant for the current study. The sources of this information should be clearly reported as well as the techniques used to synthesize the data if the data come from multiple sources. The borrowing method should be described and justified, as well as any parameters that regulate the amount of information borrowed. The full list can be found in the FDA guidance. The sensitivity of the results to different prior assumptions and degrees of borrowing should also be reported [12, 26]. Since Bayesian methods can be expected to be increasingly used in the future in clinical trials, a reporting guideline will be needed. Its development can start from items listed in the FDA guidance.

Supplementary information. An online Supplementary Appendix accompanies this manuscript.

Declarations

Funding

No funding is associated with the work featured in this article.

Competing interests

The authors have declared no conflict of interest.

Ethics approval and consent to participate

Not applicable

Availability of data and materials

Available on request

Authors' contributions

VW: Conceptualization, Formal analysis, Writing-original draft, Writing- review & editing. **SC:** Conceptualization, Formal analysis, Writing-original draft, Writing-review & editing. **LH:** Conceptualization, Writing- review & editing. **MW:** Conceptualization, Writing- review & editing. **AK:** Conceptualization, Writing- review & editing.

Consent for publication

Not applicable

Acknowledgements

Not applicable

References

[1] Amuge P, Lugemwa A, Wynne B, et al (2022) Once-daily dolutegravir-based antiretroviral therapy in infants and children living with HIV from age 4 weeks: results from the below 14 kg cohort in the randomised ODYSSEY trial. The Lancet HIV 9(9):e638–e648. https://doi.org/10.1016/S2352-3018(22)00163-1

[2] Baeten D, Baraliakos X, Braun J, et al (2013) Anti-interleukin-17A monoclonal antibody secukinumab in treatment of ankylosing spondylitis: a randomised,

double-blind, placebo-controlled trial. The Lancet 382(9906):1705–1713. https://doi.org/10.1016/s0140-6736(13)61134-4

[3] Banbeta A, Lesaffre E, Van Rosmalen J (2022) The power prior with multiple historical controls for the linear regression model. Pharmaceutical Statistics 21(2):418–438. https://doi.org/10.1002/pst.2178

[4] Barret L, Liaigre L, Hlavaty A, et al (2025) Priors and decision thresholds in phase 2 and phase 3 randomized controlled trials evaluating drug efficacy using Bayesian methods: a systematic review. Journal of Clinical Epidemiology 179:111651. https://doi.org/10.1016/j.jclinepi.2024.111651

[5] Berry SM, Carlin BP, Lee JJ, et al (2010) Bayesian Adaptive Methods for Clinical Trials. CRC press, https://doi.org/10.1201/ebk1439825488

[6] Berry SM, Broglio KR, Groshen S, et al (2013) Bayesian hierarchical modeling of patient subpopulations: efficient designs of phase II oncology clinical trials. Clinical Trials 10(5):720–734. https://doi.org/10.1177/1740774513497539

[7] Best N, Ajimi M, Neuenschwander B, et al (2025) Beyond the classical type I error: Bayesian metrics for Bayesian designs using informative priors. Statistics in Biopharmaceutical Research 17(2):183–196. https://doi.org/10.1080/19466315.2024.2342817

[8] Brard C, Piperno-Neumann S, Delaye J, et al (2019) Sarcome-13/OS2016 trial protocol: a multicentre, randomised, open-label, phase II trial of mifamurtide combined with postoperative chemotherapy for patients with newly diagnosed high-risk osteosarcoma. BMJ Open 9(5):e025877. https://doi.org/10.1136/bmjopen-2018-025877

[9] Calderazzo S, Wiesenfarth M, Kopp-Schneider A (2024) Robust incorporation of historical information with known type I error rate inflation. Biometrical Journal 66(1):2200322. https://doi.org/10.1002/bimj.202200322

[10] Calderazzo S, Wiesenfarth M, Weru V, et al (2025) Principled type I error rate inflation in two-arm clinical trial designs with external control information borrowing. arXiv preprint arXiv:250816348 2508.16348

[11] Dang LE, Tarp JM, Abrahamsen TJ, et al (2022) A cross-validated targeted maximum likelihood estimator for data-adaptive experiment selection applied to the augmentation of RCT-control arms with external data. arXiv preprint arXiv:221005802 https://doi.org/10.1515/jci-2024-0041

[12] FDA (2026) Use of Bayesian methodology in clinical trials of drug and biological products guidance for industry. Draft guidance URL https://www.fda.gov/regulatory-information/search-fda-guidance-documents/use-{Bayesian}-methodology-clinical-trials-drug-and-biological-products

[13] Fels LM, Costescu D, Vieira CS, et al (2023) The effect of a combined indomethacin and levonorgestrel-releasing intrauterine system on short-term postplacement bleeding profile: a randomized proof-of-concept trial. American Journal of Obstetrics and Gynecology 228(3):322–e1. https://doi.org/10.1016/j.ajog.2022.10.025

[14] Ferreira D, Vivot A, Diemunsch P, et al (2020) Bayesian analysis from phase III trials was underused and poorly reported: a systematic review. Journal of Clinical Epidemiology 123:107–113. https://doi.org/10.1016/j.jclinepi.2020.03.021

[15] Frenel JS, Mathiot L, Cropet C, et al (2025) Durvalumab and tremelimumab in combination with metronomic oral vinorelbine for recurrent advanced cervical

cancer: an open-label phase I/II study. Journal for Immunotherapy of Cancer 13(1):e010708. https://doi.org/10.1136/jitc-2024-010708

[16] Gao C, Yang S, Shan M, et al (2025) Improving randomized controlled trial analysis via data-adaptive borrowing. Biometrika 112(2):asae069. https://doi.org/10.1093/biomet/asae069

[17] Glatt S, Jemec GB, Forman S, et al (2021) Efficacy and safety of bimekizumab in moderate to severe hidradenitis suppurativa: a phase 2, double-blind, placebo-controlled randomized clinical trial. JAMA Dermatology 157(11):1279–1288. https://doi.org/10.1001/jamadermatol.2021.2905

[18] Gordon A, Schuler A (2025) A non-parametric sensitivity analysis for bounding bias in hybrid control trials. arXiv preprint arXiv:250718876 2507.18876

[19] Held L, Sabanés Bové D (2020) Likelihood and Bayesian Inference - With Applications in Biology and Medicine, 2nd edn. Springer Berlin Heidelberg, https://doi.org/10.1007/978-3-662-60792-3

[20] Hobbs BP, Carlin BP, Mandrekar SJ, et al (2011) Hierarchical commensurate and power prior models for adaptive incorporation of historical information in clinical trials. Biometrics 67(3):1047–1056. https://doi.org/10.1111/j.1541-0420.2011.01564.x

[21] Hobbs BP, Pestana RC, Zabor EC, et al (2022) Basket trials: review of current practice and innovations for future trials. Journal of Clinical Oncology 40(30):3520–3528. https://doi.org/10.1200/jco.21.02285

[22] Honap S, Sands BE, Jairath V, et al (2024) Basket, umbrella, and platform trials: the potential for master protocol–based trials in inflammatory bowel disease. Gastroenterology 167(4):636–642. https://doi.org/10.1053/j.gastro.2024.04.020

[23] Hueber W, Sands BE, Lewitzky S, et al (2012) Secukinumab, a human anti-IL-17A monoclonal antibody, for moderate to severe Crohn's disease: unexpected results of a randomised, double-blind placebo-controlled trial. Gut 61(12):1693–1700. https://doi.org/10.1136/gutjnl-2011-301668

[24] Ibrahim JG, Chen MH (2000) Power prior distributions for regression models. Statistical Science pp 46–60. https://doi.org/10.1214/ss/1009212673

[25] Ibrahim JG, Chen MH, Gwon Y, et al (2015) The power prior: theory and applications. Statistics in Medicine 34(28):3724–3749. https://doi.org/10.1002/sim.6728

[26] ICH (2024) E11A: Pediatric extrapolation guidance for industry. URL https://www.ema.europa.eu/en/ich-guideline-e11a-pediatric-extrapolation-scientific-guideline

[27] Kawamoto Y, Morizane C, Komatsu Y, et al (2024) Phase II trial of niraparib for BRCA-mutated biliary tract, pancreatic and other gastrointestinal cancers: NIR-B. Future Oncology 20(26):1901–1907. https://doi.org/10.2217/fon-2023-0348

[28] Kopp-Schneider A, Wiesenfarth M, Held L, et al (2024) Simulating and reporting frequentist operating characteristics of clinical trials that borrow external information: towards a fair comparison in case of one-arm and hybrid control two-arm trials. Pharmaceutical Statistics 23(1):4–19. https://doi.org/10.1002/pst.2334

[29] Kruschke JK (2021) Bayesian analysis reporting guidelines. Nature Human Behaviour 5(10):1282–1291. https://doi.org/10.1038/s41562-021-01177-7

[30] Li X, Miao W, Lu F, et al (2023) Improving efficiency of inference in clinical trials with external control data. Biometrics 79(1):394–403. https://doi.org/10.1111/biom.13583

[31] Ludbrook VJ, Budd DC, Thorn K, et al (2024) Inhibition of receptor-interacting protein kinase 1 in chronic plaque psoriasis: a multicenter, randomized, double-blind, placebo-controlled study. Dermatology and Therapy 14(2):489–504. https://doi.org/10.1007/s13555-024-01097-0

[32] Maertens J, Logan AC, Jang J, et al (2020) Phase 2 study of anti-human cytomegalovirus monoclonal antibodies for prophylaxis in hematopoietic cell transplantation. Antimicrobial Agents and Chemotherapy 64(4):10–1128. https://doi.org/10.1128/aac.02467-19

[33] McIlroy G, Lax S, Gaskell C, et al (2024) Investigator choice of standard therapy versus sequential novel therapy arms in the treatment of relapsed follicular lymphoma (REFRACT): study protocol for a multi-centre, open-label, randomised, phase II platform trial. BMC Cancer 24(1):370. https://doi.org/10.1186/s12885-024-12112-0

[34] Muehlemann N, Mukherjee R, Taher AT, et al (2021) Innovative adaptive study design in transfusion-dependent beta-thalassemia: Bayesian design with concurrent randomization and borrowing from historical data. Blood 138:4160. https://doi.org/10.1182/blood-2021-146512

[35] Neuenschwander B, Capkun-Niggli G, Branson M, et al (2010) Summarizing historical information on controls in clinical trials. Clinical Trials 7(1):5–18. https://doi.org/10.1177/1740774509356002

[36] Olson A, Lin R, Marin D, et al (2021) Third-party BK virus-specific cytotoxic T lymphocyte therapy for hemorrhagic cystitis following allotransplantation. Journal of Clinical Oncology 39(24):2710–2719. https://doi.org/10.1200/jco.20.02608

[37] Park JJH, Siden E, Zoratti MJ, et al (2019) Systematic review of basket trials, umbrella trials, and platform trials: a landscape analysis of master protocols. Trials 20(1):572. https://doi.org/10.1186/s13063-019-3664-1

[38] Piha-Paul SA, Taylor MH, Spitz D, et al (2019) Efficacy and safety of buparlisib, a PI3K inhibitor, in patients with malignancies harboring a PI3K pathway activation: a phase 2, open-label, single-arm study. Oncotarget 10(60):6526. https://doi.org/10.18632/oncotarget.27251

[39] Pocock SJ (1976) The combination of randomized and historical controls in clinical trials. Journal of Chronic Diseases 29(3):175–188. https://doi.org/10.1016/0021-9681(76)90044-8

[40] Pozzilli P, Bosi E, Cirkel D, et al (2020) Randomized 52-week phase 2 trial of albiglutide versus placebo in adult patients with newly diagnosed type 1 diabetes. The Journal of Clinical Endocrinology & Metabolism 105(6):e2192–e2206. https://doi.org/10.1210/clinem/dgaa149

[41] Richeldi L, Azuma A, Cottin V, et al (2022) Trial of a preferential phosphodiesterase 4B inhibitor for idiopathic pulmonary fibrosis. New England Journal of Medicine 386(23):2178–2187. https://doi.org/10.1056/nejmoa2201737

[42] Robinson MF, Damjanov N, Stamenkovic B, et al (2020) Efficacy and safety of PF-06651600 (Ritlecitinib), a novel Jak3/TEC inhibitor, in patients with Moderate-To-Severe rheumatoid arthritis and an inadequate response to methotrexate. Arthritis & Rheumatology 72(10):1621–1631. https://doi.org/10.1002/art.41316

[43] Scheele W, Diamond S, Gale J, et al (2016) Phosphodiesterase type 5 inhibition reduces albuminuria in subjects with overt diabetic nephropathy. Journal of the

American Society of Nephrology 27(11):3459–3468. https://doi.org/10.1681/asn.2015050473

[44] Schmidli H, Gsteiger S, Roychoudhury S, et al (2014) Robust meta-analytic-predictive priors in clinical trials with historical control information. Biometrics 70(4):1023–1032. https://doi.org/10.1111/biom.12242

[45] Schuler M, Berry DA, Burris HA, et al (2024) Nivolumab in patients with advanced or metastatic malignancies, including rare cancers: results of checkmate 627, an adaptive basket design clinical trial. JCO Oncology Advances 1:e2400044. https://doi.org/10.1200/oa-24-00044

[46] Seccia TM, Shagjaa T, Morpurgo M, et al (2023) Randomized clinical trial of nafamostat mesylate, a potent transmembrane protease serine 2 (TMPRSS2) inhibitor, in patients with COVID-19 pneumonia. Journal of Clinical Medicine 12(20):6618. https://doi.org/10.3390/jcm12206618

[47] Slosberg ED, Kang BP, Peguero J, et al (2018) Signature program: a platform of basket trials. Oncotarget 9(30):21383. https://doi.org/10.18632/oncotarget.25109

[48] Spiegelhalter DJ, Myles JP, Jones DR, et al (2000) Bayesian methods in health technology assessment: a review. Health Technology Assessment 4. https://doi.org/10.3310/hta4380

[49] Subbiah V, Kreitman R, Wainberg Z, et al (2022) Dabrafenib plus trametinib in patients with BRAF V600E-mutant anaplastic thyroid cancer: updated analysis from the phase II ROAR basket study. Annals of Oncology 33(4):406–415. https://doi.org/10.1016/j.annonc.2021.12.014

[50] Sung L, Hayden J, Greenberg ML, et al (2005) Seven items were identified for inclusion when reporting a Bayesian analysis of a clinical study. Journal of Clinical

Epidemiology 58(3):261–268. https://doi.org/10.1016/j.jclinepi.2004.08.010

[51] Thall PF, Wathen JK, Bekele BN, et al (2003) Hierarchical Bayesian approaches to phase II trials in diseases with multiple subtypes. Statistics in Medicine 22(5):763–780. https://doi.org/10.1002/sim.1399

[52] Tricco AC, Lillie E, Zarin W, et al (2018) Prisma extension for scoping reviews (prisma-scr): Checklist and explanation. Annals of Internal Medicine 169(7):467–473. https://doi.org/10.7326/m18-0850, URL http://dx.doi.org/10.7326/m18-0850

[53] Weru V, Kopp-Schneider A, Wiesenfarth M, et al (2026) Information borrowing in bayesian clinical trials: choice of tuning parameters for the robust mixture prior. Statistics in Biopharmaceutical Research p 1–21. https://doi.org/10.1080/19466315.2026.2646537, URL http://dx.doi.org/10.1080/19466315.2026.2646537

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- SupplementaryAppendix.pdf
- PRISMAScRChecklist.docx