

Sample size requirement in trials that use the composite endpoint major adverse cardiovascular events (MACE)

Marsal, Josep Ramon; Urreta-Barallobre, Iratxe; Ubeda-Carrillo, Marimar; Osorio, Dimelza; Lumbreras, Blanca; Lora, David; Fernández-Felix, Borja M.; Oristrell, Gerard; Ródenas-Alesina, Eduard; Herrador, Lorena; Ballesteros, Mónica; Zamora, Javier; Pijoan, Jose I.; Ribera, Aida; Ferreira-González, Ignacio

DOI:

[10.1186/s13063-022-06977-4](https://doi.org/10.1186/s13063-022-06977-4)

License:

Creative Commons: Attribution (CC BY)

Document Version

Publisher's PDF, also known as Version of record

Citation for published version (Harvard):

Marsal, JR, Urreta-Barallobre, I, Ubeda-Carrillo, M, Osorio, D, Lumbreras, B, Lora, D, Fernández-Felix, BM, Oristrell, G, Ródenas-Alesina, E, Herrador, L, Ballesteros, M, Zamora, J, Pijoan, JI, Ribera, A & Ferreira-González, I 2022, 'Sample size requirement in trials that use the composite endpoint major adverse cardiovascular events (MACE): new insights', *Trials*, vol. 23, no. 1, 1037. <https://doi.org/10.1186/s13063-022-06977-4>

[Link to publication on Research at Birmingham portal](#)

General rights

Unless a licence is specified above, all rights (including copyright and moral rights) in this document are retained by the authors and/or the copyright holders. The express permission of the copyright holder must be obtained for any use of this material other than for purposes permitted by law.

- Users may freely distribute the URL that is used to identify this publication.
- Users may download and/or print one copy of the publication from the University of Birmingham research portal for the purpose of private study or non-commercial research.
- User may use extracts from the document in line with the concept of 'fair dealing' under the Copyright, Designs and Patents Act 1988 (?)
- Users may not further distribute the material nor use it for the purposes of commercial gain.

Where a licence is displayed above, please note the terms and conditions of the licence govern your use of this document.

When citing, please reference the published version.

Take down policy

While the University of Birmingham exercises care and attention in making items available there are rare occasions when an item has been uploaded in error or has been deemed to be commercially or otherwise sensitive.

If you believe that this is the case for this document, please contact UBIRA@lists.bham.ac.uk providing details and we will remove access to the work immediately and investigate.

REVIEW

Open Access



Sample size requirement in trials that use the composite endpoint major adverse cardiovascular events (MACE): new insights

Josep Ramon Marsal^{1,2}, Iratxe Urreta-Barallobre^{2,3,4}, Marimar Ubeda-Carrillo⁵, Dimelza Osorio^{2,6}, Blanca Lumbreras^{2,7}, David Lora^{2,8,9}, Borja M. Fernández-Felix^{2,10}, Gerard Oristrell^{11,12}, Eduard Ródenas-Alesina¹¹, Lorena Herrador¹¹, Mónica Ballesteros^{2,6}, Javier Zamora^{2,10,13}, Jose I. Pijoan^{2,14,15}, Aida Ribera^{1,2*}  and Ignacio Ferreira-González^{1,2*}

Abstract

Background: The real impact of the degree of association (DoA) between endpoint components of a composite endpoint (CE) on sample size requirement (SSR) has not been explored. We estimate the impact of the DoA between death and acute myocardial infarction (AMI) on SSR of trials using the CE of major adverse cardiac events (MACE).

Methods: A systematic review and quantitative synthesis of trials that include MACE as the primary outcome through search strategies in MEDLINE and EMBASE electronic databases. We limited to articles published in journals indexed in the first quartile of the Cardiac & Cardiovascular Systems category (Journal Citation Reports, 2015–2020). The authors were contacted to estimate the DoA between death and AMI using joint probability and correlation. We analyzed the SSR variation using the DoA estimated from RCTs.

Results: Sixty-three of 134 publications that reported event rates and the therapy effect in all component endpoints were included in the quantitative synthesis. The most frequent combination was death, AMI, and revascularization ($n = 20$; 31.8%). The correlation between death and AMI, estimated from 5 trials, oscillated between -0.02 and 0.31 . SSR varied from 14,602 in the scenario with the strongest correlation to 12,259 in the scenario with the weakest correlation; the relative impact was 16%.

Conclusions: The DoA between death and AMI is highly variable and may lead to a considerable SSR variation in a trial including MACE.

Keywords: Composite endpoints, Sample size, Correlation, MACE

Background

The use of combined or composite endpoints (CEs) as primary outcomes in randomized clinical trials (RCT) is common. In the field of cardiovascular (CV) medical

literature, the CE “major adverse cardiovascular event” (MACE) is widely extended [1–9]. MACE usually includes the endpoints cardiovascular or all causes of death, non-fatal acute myocardial infarction (AMI), non-fatal stroke, and, in many instances, revascularization.

The strengths and weaknesses of the use of CEs as primary outcomes have been broadly discussed [8, 9]. Their potential to reduce sample size requirement (SSR) and their ability to capture the net effect of an intervention are the most popular attributed benefits [8, 10–13]. In a

*Correspondence: aida.ribera@vhir.org; iferreir@vhebron.net

¹ Cardiovascular Epidemiology and Research Unit, Vall d’Hebron Institut de Recerca (VHIR), Vall d’Hebron Barcelona Hospital Campus, Pg. Vall d’Hebron, 119-129, 08035 Barcelona, Spain
Full list of author information is available at the end of the article



recent study [14], we described how the expected occurrence rates of the individual component endpoints, the expected magnitude of the effect of the intervention on each endpoint, and the degree of association (DoA) or correlation between the component endpoints can modulate the ability of the CE to reduce the SSR [10, 14–16]. Although the expected event rate and the expected magnitude of effect on each endpoint can be estimated from prior trials, the expected DoA between endpoints of the CE is unknown. Even though this parameter is usually not incorporated in the calculation of the sample size, it may eventually dictate the ability of the CE to actually reduce the SSR [7, 9, 17–23]. Therefore, ignoring this parameter could lead to imprecise estimation of the required sample size which, in turn, can negatively impact the ethical soundness and logistical efficiency of the trial.

In the present study, we aimed to estimate the DoA between two common component endpoints of MACE (all-cause death and AMI) from a sample of clinical trials published in the CV medical literature that included MACE as their primary outcome. The ultimate purpose was to assess the impact of the DoA between these two component endpoints (all-cause death and AMI) on the SSR of a trial with a hypothetical CE combining these endpoints. A strong variability in the SSR depending on the DoA between death and AMI would justify the consideration of this parameter for a more precise calculation of sample size.

Methods

We performed a systematic review following Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) recommendations [24]. This study was sponsored by CIBER of Epidemiology and Public Health (CIBERESP).

Literature search

To identify studies for inclusion in this review, we developed comprehensive search strategies for MEDLINE and EMBASE electronic databases. We applied no restrictions to language or gender. The literature search was restricted to RCT using MACE as the primary outcome, published in core medical journals and indexed in the first quartile of the Cardiac & Cardiovascular Systems category in the Journal Citation Reports from 2015 to 2020.

The search strategies used, and additional core medical journals added to the criteria of selecting titles indexed in the first quartile of the Cardiac & Cardiovascular Systems category, are available in Additional file 1: Supplementary material.

Eligibility criteria

We included RCT with (a) two-arm parallel design (*i.e.*, excluding crossover, factorial or more than two treatments/groups); (b) primary outcome defined as the combination of at least 2 of the following events: cardiovascular (or all-causes) death, non-fatal AMI, and non-fatal stroke, allowing a combination with one or more other events; and (c) sample size of 500 patients or more. We excluded the manuscripts referring to interim analysis or not including the main analysis of the RCT (*i.e.*, analysis of subgroups).

Selection of studies and data extraction

All references' titles and abstracts were independently screened by at least two reviewers. Second, the full article of those selected in the first screening was reviewed. This second round of screening involved two reviewers independently and was based on the application of the selection criteria. Two independent reviewers extracted the following data from each study included: population, intervention groups and outcome characteristics, intervention effects on the different outcomes, composition of the MACE, and numbers of patients with the events of interest were independently extracted by two reviewers. Finally, we contacted by email the corresponding author/s of each included RCT to request for further information and the necessary data to estimate the association between the events conforming to the MACE (see Additional file 1: Supplementary material for further details).

Any disagreement between the reviewers at all stages was solved by a third reviewer. The risk of bias was not assessed because the aim of the study was not to estimate the unbiased effect of the study interventions.

Analytical plan

We first described the CEs of the clinical trials included in the quantitative synthesis. Then, we used the generic inverse variance method [25] to estimate the pooled event rates and the pooled risk ratios (RRs) for the CEs included in the selected trials and for those component endpoints more often included in these CEs: death (cardiovascular and all-causes death), AMI, stroke, and revascularization. This was done overall and by subgroups depending on certain RCTs characteristics: number of components of the CE, number of patients included in the trial, follow-up time, and the mean age of the population.

From the RCTs for which we could obtain the necessary data to estimate the DoA, we calculated the joint probability and Pearson's correlation between the pair of outcomes death (all causes or cardiovascular) and AMI.

We define the joint probability between the components i and j (π_{ij}) as the number of patients with both components at the final assessment divided by all patients at the initial time. Pearson’s correlation between the components i and j ($\rho_{i,j}$) was calculated as:

$$\rho_{i,j} = \frac{\pi_{i,j} - \pi_i\pi_j}{\sqrt{\pi_i(1 - \pi_i)\pi_j(1 - \pi_j)}}$$

where π_i and π_j are the proportion of deaths for any cause and proportion of AMI in both treatment groups (we pooled the results of both treatments since we consider unlikely different associations in each group), respectively, and $\pi_{i,j}$ is the joint probability mentioned before.

In addition, the standardized joint probability (SJP) was calculated as $\pi_{ij}/\min(\pi_i, \pi_j)$ because the SJP has better metric properties than joint probability or correlation [26]. A value of 0 indicates no overlapping, and 1 indicates that all subjects with the less prevalent event also had the most prevalent event. We then selected the RCTs that included the outcomes death (all-causes death or CV death) and AMI as component endpoints. We estimated, from these trials, the pooled event rate in both baseline ($t = 0$) and final ($t = 1$) from the proportion of each event and the proportions of subjects with both events as follows: $\pi_t = \pi_{i,t} + \pi_{j,t} - \pi_{ij,t}$, the pooled effect of the therapies and the SSR of a hypothetical clinical trial that included these outcomes as individual primary outcomes. SSR was calculated using the normal approximation to the binomial test [14, 27, 28] as follows:

$$n \geq \left(\frac{Z_{\alpha/2} \sqrt{2\pi(1 - \pi)} + Z_{\beta} \sqrt{\pi_{t=0}(1 - \pi_{t=0}) + \pi_{t=1}(1 - \pi_{t=1})}}{\pi_{t=0} - \pi_{t=1}} \right)^2, \pi = \frac{\pi_{t=0} + \pi_{t=1}}{2}$$

where $Z_{\alpha/2}$ and Z_{β} are the standardized normal quantiles for error types I and II, respectively.

Finally, we estimated the pooled event rate, the pooled effect of the therapies, and the SSR in a theoretical RCT that included these endpoints as a primary CE considering two scenarios according to DoA between component outcomes: the most favorable scenario for SSR (weakest correlation) and the less favorable scenario (strongest correlation). All these calculations were further carried out considering also four hypothetical scenarios depending on the event rate of the original MACE and the relative effect of the intervention on the original MACE. Thus, scenario 1 included those studies with an event rate of the CE lower than the median event rate in the whole sample (< 10.9%) and an effect of the therapy on the CE lower than the median effect in the whole sample (RR >

0.9); scenario 2 contained those studies with an event rate of the CE lower than the median event rate in the whole sample (< 10.9%) and an effect of the therapy higher than the median effect in the whole sample (RR < 0.9); scenario 3 included those studies with an event rate > 10.9% and effect on the CE > 0.9; scenario 4 included those studies with an event rate > 10.9% and effect RR < 0.9.

Results

Literature search

Our literature search identified 761 articles, and 37 additional articles were detected after a manual search. A total of 548 articles were excluded after reading the titles and abstracts, and 13 additional articles were excluded because they were duplicates or because they were not an RCT (Fig. 1). The remaining 237 articles corresponded to several publications of 134 RCTs. Sixty-three out of these 134 primary publications reported both the event rate and the effect of the therapy on each specific endpoint of the MACE. These 63 were included in the quantitative synthesis. None of them reported the DoA between the endpoints of CE. The selection process is detailed in Additional file 1: Supplementary material. We sent a request for information to the corresponding authors of the 63 RCTs, and we could obtain the necessary data to estimate the DoA for only 5 RCTs.

CE combinations

The 63 RCTs included a total of 332,199 patients (ranging from 501 to 27,564). Most trials ($n = 38$; 60.3%) included

a CE with three endpoints, being AMI the most frequent ($n = 62$; 98.4%), followed by stroke ($n = 37$; 58.7%), cardiovascular death ($n = 35$; 55.6%), and revascularization ($n = 33$; 52.4%) (Table 1). The most frequent combination was death for any reason or cardiovascular death, AMI, and revascularization ($n = 20$; 31.8%). CV death and AMI were used in 32 RCTs, more than 50% of the identified trials. Death and AMI appeared in 29 RCT. In most of the analyzed trials, predetermined follow-up was longer than 1 year ($n = 53$; 84.1%).

Table 2 shows the pooled MACE CE event rate and the pooled effect of the different interventions in the study, expressed as risk ratio, for the overall 63 RCTs and by subgroups created according to the number of endpoints included in the CE, sample size of the trial, time of follow-up, baseline average age of the

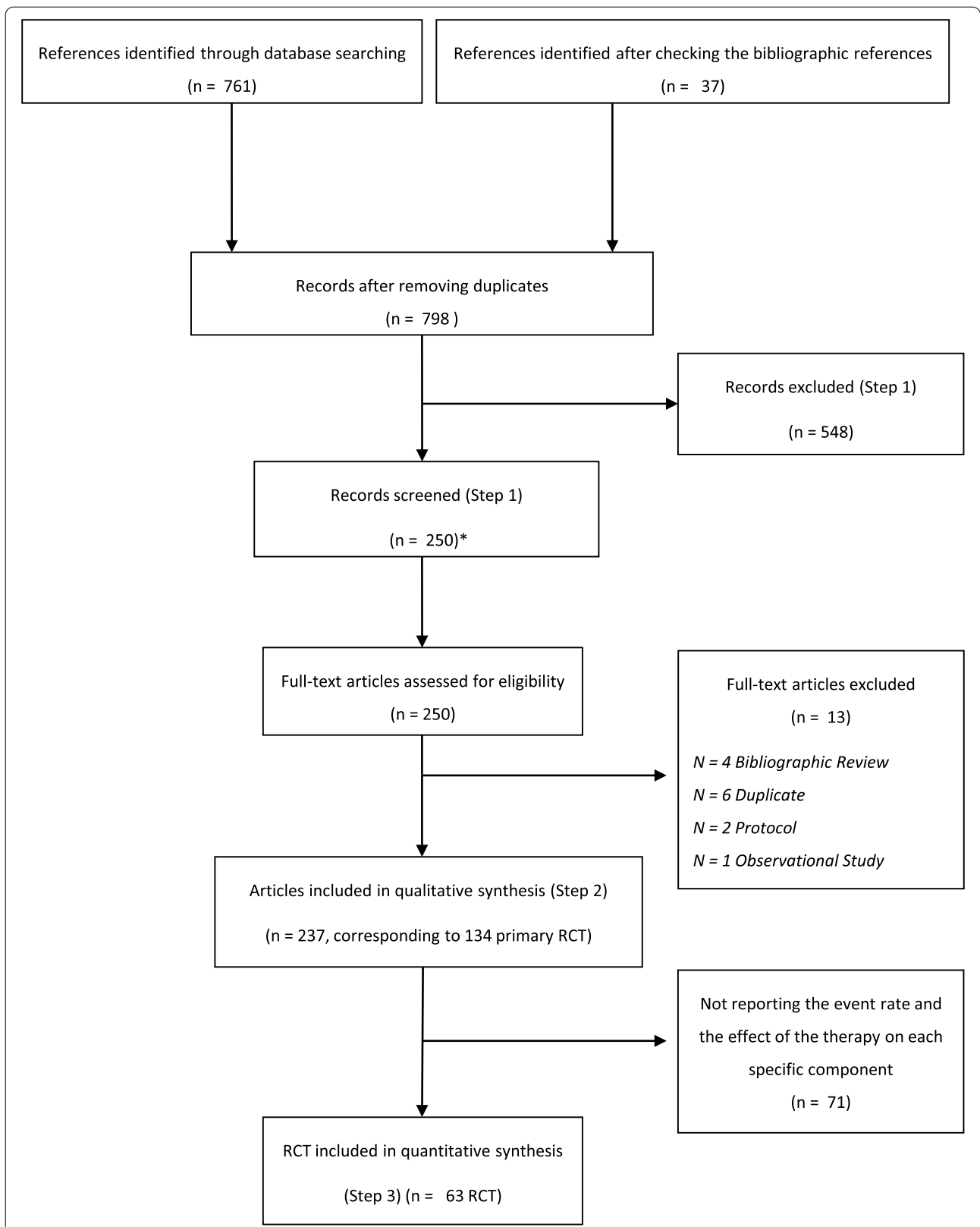


Fig. 1 A total of 548 articles were excluded after reading the titles and abstracts, and 13 additional articles were excluded because they were duplicates or because they were not an RCT

Table 1 MACE's design and characteristics of the 63 RCT included in the analysis

	Number	Percent
RCT included	63	
Type of intervention		
Drug treatment (vs placebo)	23	36.5
Drug treatment (vs other drug or dose)	5	7.9
Procedure (type of device or invasive procedure)	23	36.5
Diagnostic (to guide intervention)	9	14.1
Other types of intervention (process of care, rehab)	4	6.3
No. of component outcomes		
3 components	38	60.3
4 components	12	19.0
5 components	13	20.6
No. of patients included		
≤ 1500	21	33.3
1500–5000	22	34.9
> 5000	20	31.7
Individual component outcomes		
AMI	62	98.4
Stroke	37	58.7
CV death	35	55.6
Revascularization	33	52.4
Death for any reason	28	44.4
Angina	13	20.6
Others	9	14.3
Stent thrombosis	3	4.8
HF	3	4.8
Bleeding	4	6.3
Most frequent MACE		
CV death - AMI - revascularization	11	17.5
CV death - AMI - stroke	11	17.5
Death - AMI - revascularization	9	14.3
Death - AMI - stroke	6	9.5
CV death - AMI - stroke - angina	4	6.3
CV death - AMI - stroke - revascularization - angina	4	6.3
Death - AMI - revascularization - stroke	4	6.3
CV death - AMI - revascularization - angina - chest pain	1	1.6
CV death - AMI - stent thrombosis	1	1.6
CV death - AMI - stroke - angina - HF	1	1.6
CV death - AMI - stroke - bleeding	1	1.6
Death - AMI - stent thrombosis - stroke - revascularization	1	1.6
Death - AMI - angina - complicated procedure	1	1.6
Death - AMI - angina - CV admission - stroke	1	1.6
Death - AMI - HF	1	1.6
Death - AMI - revascularization - ACS	1	1.6
Death - AMI - stroke - bleeding	1	1.6
Death - AMI - stroke - bleeding - revascularization	1	1.6
Death - AMI - stroke - cardiovascular hospitalization	1	1.6
Death - low cardiac output - AMI - cardiac arrest - HF	1	1.6
Death - stroke - bleeding - vascular complication - coronary obstruction	1	1.6

Table 1 (continued)

	Number	Percent
Follow-up		
≤ 1 year	10	15.9
From 1 to 5 years	47	74.6
> 5 years	6	9.5
Mean age over 64 years	31	47.6

AMI acute myocardial infarction, CV death cardiovascular death, HF heart failure with hospitalization, ACS acute coronary syndrome

participants, and the specific endpoints included in the CE. The pooled event rate and the pooled magnitude of effect (risk ratio) for the CEs in the 63 RCT were 11.6% (95% CI 9.7–13.4) and 0.89 (95% CI 0.86–0.92), respectively. There was not any association between these parameters and the number of endpoints included in the CE. The magnitude of the effect of interventions was higher in RCTs with smaller sample sizes. However, neither time of follow-up nor the baseline mean age was associated to the effect of therapies. The largest magnitude of the effect of therapies was on the endpoint revascularization (0.75; 95% CI 0.68–0.82). Concerning the global event rate of the CEs, it was higher in those clinical trials with longer follow-up and in those with a higher baseline mean age of the population included, without differences depending on the sample size of the clinical trials. Again, the largest event rate corresponded to the endpoint revascularization (6.5%; 95% CI 7.7–8.3).

Impact of correlation between death and MI on SSR

Table 3 shows the DoA between death and AMI estimated from 5 studies and expressed as the joint probability of the distributions of both endpoints and the correlation, which ranged between –0.02 and 0.31. Figure 2 shows the impact of the correlation between death and acute myocardial infarction on the event rate and on the SSR of a theoretical trial that includes these outcomes in a composite endpoint. All calculations were performed with data from the overall 63 clinical trials (left columns in the figure) and classifying these trials in four scenarios according to the pooled magnitude of the effect of the intervention on the original CE (RR < 0.9 and RR > 0.9) and the pooled CE event rate (< 10.9% and > 10.9%).

Considering the full set of RCT, the impact of correlation on the event rate was a variation of 19% from the less favorable scenario (strongest correlation, event rate 7.17%) to the most favorable scenario (weakest correlation, event rate 8.53%). This led to a SSR variation of 16% from the less favorable scenario (strongest correlation, SSR 14,602 patients) to the most favorable scenario (weakest correlation, SSR 12,259 patients), which means a

Table 2 MACE proportion and risk ratio depending on the different characteristics

MACE	Number	Event rate			Risk ratio		
		%	95% CI	<i>p</i>	RR	95% CI	<i>p</i>
No. of components	63	11.6	9.7–13.4%	0.456	0.89	0.86–0.92	0.499
3 components	38	10.7	8.5–12.8%		0.91	0.88–0.94	
4 components	12	13.2	9.3–17.1%		0.87	0.81–0.94	
5 components	13	12.8	7.4–18.1%		0.87	0.81–0.94	
No. of patients included				0.428			0.002
≤ 1500	21	13.0	9.2–16.8%		0.77	0.71–0.85	
1500–5000	22	11.9	9.5–14.3%		0.91	0.89–0.94	
> 5000	20	9.9	6.8–13.1%	0.91	0.86–0.95		
Follow-up time				0.002			0.650
≤ 1 year	10	11.9	6.7–17.1%		0.85	0.78–0.94	
1 to 5 years	47	10.2	8.4–11.9%		0.89	0.86–0.93	
> 5 years	6	22.0	15.7–28.3%	0.89	0.83–0.98		
Age				0.039			0.907
≤ 64 years	32	9.8	7.3–12.4%		0.89	0.85–0.92	
> 64 years	31	13.6	11.1–16.2%	0.89	0.85–0.93		
Individual outcome							
CV death		3.5	2.7–4.3%		0.9	0.87–0.94	
All-causes death		3.8	2.6–5%		0.88	0.82–0.94	
Acute myocardial infarction		4.3	3.4–5.1%		0.87	0.83–0.91	
Stroke		1.7	1.4–2.1%		0.84	0.79–0.88	
Revascularization		6.5	4.7–8.3%		0.75	0.68–0.82	

difference of 2343 patients in absolute numbers. Although the relative impact of the DoA across the four scenarios was similar, in absolute values, SSR may vary up to 7192 patients, depending upon the specific combination of absolute risk of the event in the study follow-up and the magnitude of the relative effect of the intervention.

Discussion

In this systematic review of 63 randomized clinical trials in the CV field published in first-quartile journals, the CE most often used (*n* = 20 trials) included the

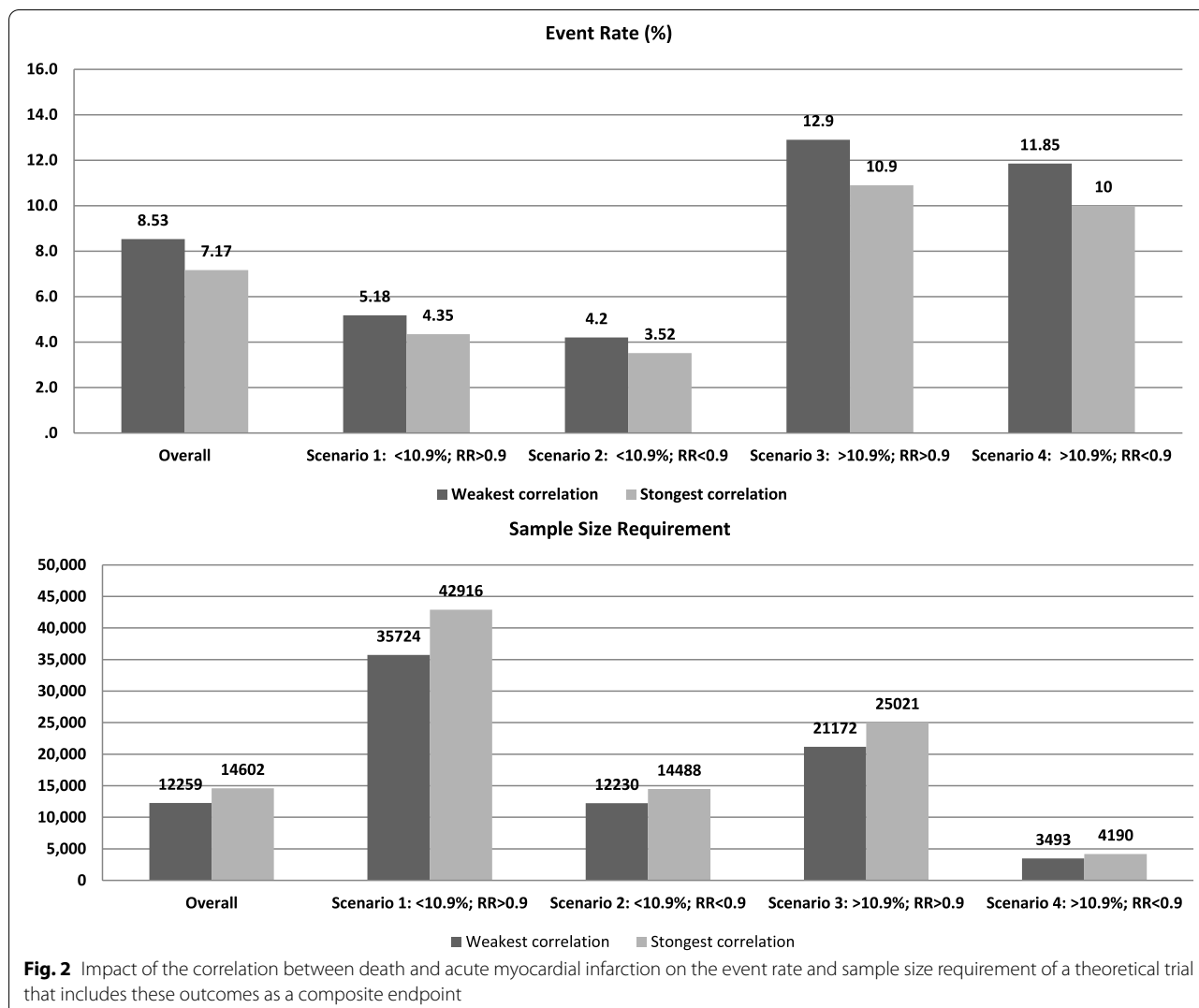
endpoints death (all-cause death or CV death), AMI, and revascularization. In the pooled analysis of these RCTs, the pooled effect of therapies and the pooled event rate were higher in the component endpoint revascularization, which is clinically the less important outcome. The DoA between the endpoints death and AMI, which could be estimated from 5 RCTs, was highly variable. This variability led to a SSR variation of a hypothetical CE including these two outcomes from 2343 patients (16%). The relative impact of the DoA on the variation of SSR was similar in different settings regardless of the event rate

Table 3 Degree of association between death and acute myocardial infarction, expressed by joint probability and correlation, in the 5 RCT with complete information

Study	Number	Death (<i>n</i> , (%))	AMI (<i>n</i> , (%))	Death and AMI (<i>n</i>)	Joint probability (%)	Correlation (95% CI)
ACCELERATE [29] ^a	12,092	309 (2.6)	517 (4.3)	74	0.61	0.157 (0.14–0.17)
GIK [30]	930	20 (2.2)	2 (0.2)	2	0.22	0.31 (0.25–0.37)
PROMISE [31]	10,003	149 (1.5)	70 (0.7)	3	0.03	0.02 (0–0.04)
Yoga-Ca [32]	3,959	144 (3.6)	28 (0.7)	0	0.00	–0.02 (–0.05–0.01)
MR.INFORM [33]	918	6 (0.7)	19 (2.1)	1	0.11	0.08 (0.02–0.14)

AMI Acute myocardial infarction

^a Cardiovascular death



and magnitude of the effect of the intervention, but, in absolute terms, SSR variation may reach a maximum of 7192 patients.

In previous publications, we and others already showed that both the CE event rate and the effect of the therapy on the CE are usually driven by the less important endpoint [6, 10, 14, 15, 34]. In this systematic review of a sample of CV trials that used the CE MACE as the primary endpoint, we analyzed the pooled event rate and the pooled effect of the therapies on the most common endpoints, such as death, AMI, stroke, and revascularization. Again, the event rate was higher, and the magnitude of the effect of therapy was larger on the endpoint revascularization, which is the less important and can be considered a clinician-driven outcome. The gradients in both importance to patients and treatment effects across component endpoints may complicate the interpretation of trial results, as was previously pointed out [7]. Regardless

of the potential problems of interpretation due to a gradient between endpoints, since one of the main uses of the CE is to reduce SSR, in this work, we explore how the interrelationship between the most popular components included in the CE MACE (i.e., death and AMI) may impact SSR.

As we illustrated in a simulation study, the interrelationship between endpoints included in a CE can help to define the SSR in some situations, at least in the cardiovascular field. Specifically, we showed that in 25 of 66 simulated scenarios the DoA between two CV endpoints influenced the decision of using a CE instead of a single endpoint [14, 16]. We also showed that the influence of the DoA in SSR decreases when the effect of the therapy on the relevant endpoint increases.

Although it is true that the smaller the DoA between endpoints the lower the SSR of the corresponding CE, the impact of the DoA on SSR is also importantly influenced

by the CE event rate and by the magnitude of the effect of the therapy on endpoints, as we previously have shown [14]. Thus, from a statistical perspective, a small DoA between two specific endpoints should not lead to the selection of these endpoints for a CE to decrease the SSR. Rather, the expected effect of the therapy on the endpoints, the expected event rate, and, importantly, the specific type of endpoints should carefully be considered during the sample size calculation. In this sense, our findings could be useful for sample size considerations in future cardiovascular trials using the CE MACE. The present work shows that the variability of the DoA between death and AMI endpoints, estimated from real CV trials, is generally high. Specifically, the required sample size for a trial with a CE combining this couple of endpoints might vary as much as 17% depending on the DoA between them, which supposes a considerable uncertainty.

We also analyzed the impact of the DoA on the SSR across four scenarios defined by the CE event rate and by the effect of the therapy on the CE (Fig. 2). This impact, in relative terms, was similar across the four scenarios, probably because the ranges of the event rates and the magnitudes of the effect of the experimental interventions among the trials included were narrow, and these parameters are closely related with the impact of the DoA on the SSR, as we previously showed [14].

Nowadays, it is impossible to know in advance the DoA between the component endpoints of a CE. For this reason, our initial approach has been to estimate the range of variation of the DoA between the most popular endpoints included in the CE MACE, as a first step to understand the potential of this parameter to drive the SSR. In any case, the DoA between CV endpoints may likely be influenced by other intrinsic characteristics of the population, such as the baseline cardiovascular risk. For instance, the probability of death in patients who had an AMI is likely to be higher in patients with diabetes, hypertension, and renal impairment than in patients without these conditions, which can lead to a different DoA estimation between these endpoints. Similarly, clinician-driven outcomes, such as revascularization, could be influenced by regional practices, which can also lead to different DoA between revascularization and other CV outcomes depending on the region. Therefore, the next step in future research should be to explore in depth, in large databases from RCT, or in large observational studies or registries, not only the DoA between other CV endpoints but which factors, if any, determine a stronger or weaker DoA between different CV endpoints. This will permit us to anticipate the expected DoA more precisely, and thus, it will lead to a more informed and precise sample size estimation in future studies. In the meantime, given

the important variation in correlation estimates between death and AMI, ranging from -0.02 to 0.31 , there are several possibilities for sample size estimation when using these outcomes. One could conservatively use the worst case, one could use some average correlation, or one could estimate the correlation in a blinded interim analysis and recalculate the sample size. Although our intention is not to make a formal recommendation regarding how to calculate SSR based on DoA, we think that the latter strategy would likely allow a more precise correction in the interim analysis regarding SSR.

Limitations

Unfortunately, we could obtain the data necessary to estimate the DoA between death and AMI only from 5 RCTs. Therefore, we cannot conclude that our DoA estimations may be representative of the full spectrum of CV trials. However, these five trials may represent, to a certain extent, a disparate trial sample (i.e., different sample sizes, event rates, magnitude of effect). We also restricted ourselves to the study of CV CE, specifically MACE, a commonly used endpoint. Relationships between component endpoints making up CE in other fields (e.g., infectious and neurological diseases) may be different, as well as their potential impact on the required sample size for trials in those clinical areas.

Another limitation of our study is that we restricted our approximation to the differences in proportions, while most clinical trials in the cardiovascular area use survival analysis to estimate hazard ratios. In addition, the fact that we do not consider censoring could yield biased estimates. Although we attempt to study how the association between the most usual outcomes in the cardiovascular field varies and how this variation may impact the SSR estimations, future studies should focus on survival data to quantify DoAs between two components of a MACE, which is probably a more realistic approach.

Conclusions

The use of CE in clinical trials is a challenge for the interpretation of their results and for the estimation of the anticipated SSR. The most common CE used in CV trials combines the outcomes death, AMI, stroke, and revascularization. In this CE, the higher event rate and the higher effect of therapies usually falls on the outcome revascularization, which is the less important. The DoA between components of the CE MACE may influence the SSR. Specifically, the DoA between the endpoints death and MI is highly variable and can lead to a sizeable SSR variation of a trial with a theoretical CE including these outcomes. Therefore, the DoA among the components of the most frequent CEs used in clinical trials should be explored in depth in future studies. We recommend that

future trialists incorporate some measures of the degree of association between components in their publications. Also, initiatives for enhancing transparency in publications of health research, such as EQUATOR, could also help tackle the lack of recommendations on what information to display and how to display it when CEs are used as the main results of a clinical study. This would permit the incorporation of this parameter for a refined calculation of the sample size in future clinical trials.

Abbreviations

AMI: Acute myocardial infarction; CEs: Composite endpoints; CV: Cardiovascular; DoA: Degree of association; MACE: Major adverse cardiac events; PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses; RCT: Randomized clinical trials; RR: Risk ratio; SSR: Sample size requirement.

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s13063-022-06977-4>.

Additional file 1: Supplementary material.

Acknowledgements

We are very grateful for the comments and support from Guadalupe Gomez and Jordi Cortés. We also want to thank Carmen Bernal Soriano, Marta Puig, Diego San José, Ana Pastor Moreno, Celia Fátima Gómez, Shandra Fuentes, Iñigo Sanz, and Eñaut Aguirre for their support during the systematic reviewing process.

Authors' contributions

JR, AR, JZ, and IFG designed the study. IUB and MUB searched the manuscripts, and all authors were involved in the data extraction. JR and AR analyzed the data. IFG, JIP, AR, and JR worked on the first draft, although all authors were involved in the design, drafting, and acceptance of the final version of this article. The authors read and approved the final manuscript.

Funding

Intramural CIBER-ESP PR22 from the Center for Biomedical Research in Epidemiology and Public Health Network (CIBERESP).

Availability of data and materials

All data and materials are available from the corresponding author upon request.

Declarations

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Author details

¹Cardiovascular Epidemiology and Research Unit, Vall d'Hebron Institut de Recerca (VHIR), Vall d'Hebron Barcelona Hospital Campus, Pg. Vall d'Hebron, 119-129, 08035 Barcelona, Spain. ²CIBER Epidemiology and Public Health, Madrid, Spain. ³Biodonostia Health Research Institute, Clinical Epidemiology, San Sebastián, Spain. ⁴Osakidetza Basque Health Service, Donostialdea Integrated Health Organisation, Donostia University Hospital, Clinical Epidemiology Unit, San Sebastián, Spain. ⁵Osakidetza Basque Health Service, Donostialdea Integrated Health Organisation, Donostia University Hospital, Library

Service, San Sebastián, Spain. ⁶Health Services Research Group, Vall d'Hebron Institut de Recerca (VHIR), Vall d'Hebron Hospital Universitari, Vall d'Hebron Barcelona Hospital Campus, Passeig Vall d'Hebron 119-129, 08035 Barcelona, Spain. ⁷Public Health Department, Miguel Hernandez University, Alicante, Spain. ⁸Health Research Institute Hospital 12 de Octubre (imas12), Madrid, Spain. ⁹Statistical Studies Department, Universidad Complutense de Madrid (UCM), Madrid, Spain. ¹⁰Clinical Biostatistics Unit, Hospital Ramón y Cajal (IRYCIS), Madrid, Spain. ¹¹Cardiology Department, Vall d'Hebron University Hospital, Barcelona, Spain. ¹²CIBER Cardiovascular Diseases, Madrid, Spain. ¹³Institute of Metabolism and Systems Research, University of Birmingham, Birmingham, UK. ¹⁴Clinical Epidemiology Unit, Cruces University Hospital, Barakaldo, Spain. ¹⁵Biocruces-Bizkaia Health Research Institute, Barakaldo, Spain.

Received: 1 July 2022 Accepted: 2 December 2022

Published online: 21 December 2022

References

- Lauer MS, Topol EJ. Clinical trials - multiple treatments, multiple endpoints, and multiple lessons. *JAMA*. 2003;289(19):2575–7.
- ICH C. ICH Topic E 9: statistical principles for clinical trials. London: EMEA (European Medicines Agency); 2006.
- Moher D, Hopewell S, Schulz KF, Montori V, Gøtzsche PC, Devereaux PJ, et al. CONSORT 2010 explanation and elaboration: updated guidelines for reporting parallel group randomised trials. *BMJ*. 2010;340:c869.
- Prieto-Merino D, Smeeth L, Van STP, Roberts I. Dangers of non-specific composite outcome measures in clinical trials. *BMJ*. 2013;347(November):f6782.
- Kip KE, Hollabaugh K, Marroquin OC, Williams DO. The problem with composite end points in cardiovascular studies: the story of major adverse cardiac events and percutaneous coronary intervention. *J Am Coll Cardiol*. 2008;51(7):701–7.
- Ferreira-González I, Permanyer-Miralda G, Busse JW, Bryant DM, Montori VM, Alonso-Coello P, et al. Composite endpoints in clinical trials: the trees and the forest. *J Clin Epidemiol*. 2007;60(7):660–1.
- Montori VM, Permanyer-Miralda G, Ferreira-González I, Busse JW, Pacheco-Huergo V, Bryant D, et al. Validity of composite end points in clinical trials. *BMJ*. 2005;330(7491):594–6 Available from: <https://www.bmj.com/content/330/7491/594>. Cited 2022 Jun 30.
- Freemantle N, Calvert M, Wood J, Eastaugh J, Griffin C. Composite outcomes in randomized trials: greater precision but with greater uncertainty? *JAMA*. 2003;289(19):2554–9.
- Freemantle N, Calvert MJ. Interpreting composite outcomes in trials. *BMJ*. 2010;341(aug18 3):c3529.
- Sozu T, Sugimoto T, Hamasaki T. Sample size determination in clinical trials with multiple co-primary binary endpoints. *Stat Med*. 2010;29(21):2169–79.
- Ross S. Composite outcomes in randomized clinical trials: arguments for and against. *Am J Obstet Gynecol*. 2007;196(2):119.e1–6.
- Xiong C, Yu K, Gao F, Yan Y, Zhang Z. Power and sample size for clinical trials when efficacy is required in multiple endpoints: application to an Alzheimer's treatment trial. *Clin Trials*. 2005;2(5):387–93.
- Cannon CP. Clinical perspectives on the use of composite endpoints. *Control Clin Trials*. 1997;18(6):517–9.
- Marsal JR, Ferreira-González I, Bertran S, Ribera A, Permanyer-Miralda G, García-Dorado D, et al. The use of a binary composite endpoint and sample size requirement: influence of endpoints overlap. *Am J Epidemiol*. 2017;185(9):832–41.
- Bofill Roig M, Gómez MG. A new approach for sizing trials with composite binary endpoints using anticipated marginal values and accounting for the correlation between components. *Stat Med*. 2019;38(1):1–22.
- Marsal JR, Ferreira-González I, Ribera A, Oristrell G, Pijoan JI, García-Dorado D. Bin-CE: A comprehensive web application to decide upon the best set of outcomes to be combined in a binary composite endpoint. *PLoS ONE*. 2018;13(12):e0209000.
- Cordoba G, Schwartz L, Woloshin S, Bae H, Gøtzsche PC. Definition, reporting, and interpretation of composite outcomes in clinical trials: systematic review. *BMJ*. 2010;341:c3920.
- Ferreira-González I, Busse JW, Heels-Ansdell D, Montori VM, Akl EA, Bryant DM, et al. Problems with use of composite end points in

- cardiovascular trials: systematic review of randomised controlled trials. *BMJ*. 2007;334(7597):786.
19. Bhardwaj SS, Camacho F, Derrow A, Fleischer AB, Feldman SR. Statistical significance and clinical relevance: the importance of power in clinical trials in dermatology. *Arch Dermatol*. 2004;140(12):1520–3.
 20. Ferreira-González I, Permyer-Miralda G, Busse JW, Bryant DM, Montori VM, Alonso-Coello P, et al. Methodologic discussions for using and interpreting composite endpoints are limited, but still identify major concerns. *J Clin Epidemiol*. 2007;60(7):651–62.
 21. Gómez G, Lagakos SW. Statistical considerations when using a composite endpoint for comparing treatment groups. *Stat Med*. 2012;32(5):719–38.
 22. Gómez G, Gómez-Mateu M, Dafni U. Informed choice of composite end points in cardiovascular trials. *Circ Cardiovasc Qual Outcomes*. 2014;7(1):170–8.
 23. Chi GYH. Some issues with composite endpoints in clinical trials. *Fundam Clin Pharmacol*. 2005;19(6):609–19.
 24. Moher D, Liberati A, Tetzlaff J, Altman DG, Antes G, et al. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *PLoS Med*. 2009;6(7):e1000097.
 25. Borenstein M, Hedges LV, Higgins JPT, Rothstein HR. A basic introduction to fixed-effect and random-effects models for meta-analysis. *Res Synth Methods*. 2010;1(2):97–111.
 26. Warrens MJ. On association coefficients for 2 × 2 tables and properties that do not depend on the marginal distributions. *Psychometrika*. 2008;73(4):777 Available from: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2798022/>. Cited 2022 Sep 30.
 27. Fleiss JL. *Statistical Methods for Rates and Proportions*. 2nd ed. New York: Wiley; 1981.
 28. Casagrande JT, Pike MC. An improved approximate formula for calculating sample sizes for comparing two binomial distributions. *Biometrics*. 1978;34(3):483–6.
 29. Lincoff AM, Nicholls SJ, Riesmeyer JS, Barter PJ, Brewer HB, Fox KAA, et al. Evacetrapib and cardiovascular outcomes in high-risk vascular disease. *N Engl J Med*. 2017;376(20):1933–42 Available from: <https://www.nejm.org/doi/full/10.1056/nejmoa1609581>. Cited 2022 Sep 19.
 30. Zhao K, Zhang Y, Li J, Cui Q, Zhao R, Chen W, et al. Modified glucose-insulin-potassium regimen provides cardioprotection with improved tissue perfusion in patients undergoing cardiopulmonary bypass surgery. *J Am Heart Assoc*. 2020;9(6) Available from: <https://pubmed.ncbi.nlm.nih.gov/32151220/>. Cited 2022 Sep 19.
 31. Douglas PS, Hoffmann U, Patel MR, Mark DB, Al-Khalidi HR, Cavanaugh B, et al. Outcomes of anatomical versus functional testing for coronary artery disease. *N Engl J Med*. 2015;372(14):1291–300 Available from: <https://www.nejm.org/doi/full/10.1056/nejmoa1415516>. Cited 2022 Sep 19.
 32. Prabhakaran D, Chandrasekaran AM, Singh K, Mohan B, Chattopadhyay K, Chadha DS, et al. Yoga-based cardiac rehabilitation after acute myocardial infarction: a randomized trial. *J Am Coll Cardiol*. 2020;75(13):1551–61 Available from: <https://pubmed.ncbi.nlm.nih.gov/32241371/>. Cited 2022 Sep 19.
 33. Nagel E, Greenwood JP, McCann GP, Bettencourt N, Shah AM, Hussain ST, et al. Magnetic resonance perfusion or fractional flow reserve in coronary disease. *N Engl J Med*. 2019;380(25):2418–28 Available from: <https://www.nejm.org/doi/full/10.1056/NEJMoa1716734>. Cited 2022 Sep 19.
 34. Bofill Roig M, Gómez MG. Selection of composite binary endpoints in clinical trials. *Biometrical J*. 2018;60(2):246–61.

Publisher's Note

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

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more biomedcentral.com/submissions

