Considerations on the strengths and limitations of using disease-related mortality as an outcome in clinical research

Janus Christian Jakobsen , Jørn Wetterslev , Christian Gluud

10.1136/bmjebm-2018-111154

Copenhagen Trial Unit, Centre for Clinical Intervention Research,Rigshospitalet, Copenhagen University Hospital, Copenhagen, Denmark

Correspondence to:
Dr Janus Christian Jakobsen, Copenhagen Trial Unit, Centre for Clinical Intervention Research,Rigshospitalet, Copenhagen University Hospital, Copenhagen, Denmark; janusjakobsen@mac.com

To cite: Jakobsen JC, Wetterslev J, Gluud C. *BMJ Evidence-Based Medicine* 2021;**26**:127–130.

Abstract

Disease-related mortality (eg, cardiovascular mortality or breast-cancer mortality) is often used as an outcome in randomised clinical trials and systematic reviews. The rationale why disease-related mortality might be used in addition to, or instead of, all-cause mortality seems to be that disease-related mortality may more readily detect the experimental intervention effects. Disease-related mortality is theoretically what most interventions aim at influencing; disease-related intervention effects are not 'diluted' by events unrelated to the disease that may be occurring in both the experimental group and the control group (eg, traffic accidents). Intervention–effect estimates are indeed theoretically diluted and affected if events unrelated to the disease or the trial interventions are occurring. Although sounding attractive, we will in the present paper consider the several methodological limitations of using disease-related mortality instead of all-cause mortality as an outcome. When mortality is a relevant outcome, we recommend using all-cause mortality as a primary outcome and disease-specific mortality as a secondary or exploratory outcome depending on power.

Key messages

What is already known about this subject?

- Disease-related mortality is often used in clinical intervention research presumably because disease-related intervention effects are (theoretically) not 'diluted' by events unrelated to the disease that are occurring in both intervention groups (eg, traffic accidents).

What are the new findings?

- There are several methodological limitations of using disease-related mortality: (1) assuming no change in relative risk reduction, there is a theoretical loss of power because of a lower proportion of participants with disease-related mortality compared to all-cause mortality, (2) disease-related mortality results might be misinterpreted, (3) the decision to classify a death as disease-specific might be influenced by knowledge of treatment allocation or differences in the degree of observation between the compared groups and lead to biased results, (4) 'true' deadly adverse events caused by trial interventions might not be classified as disease-related and (5) it may not be valid to pool different definitions of disease-related outcomes in a meta-analysis.

How might these results change the focus of research or clinical practice?

- Disease-related mortality should primarily be used as a secondary or an exploratory outcome.
- All-cause mortality should always be reported in addition to disease-related mortality.

It is of utmost importance to choose patient-centred outcomes (outcomes that are clinically relevant from a patient perspective) when conducting a randomised clinical trial or a systematic review of such trials.[1–3] Mortality will, however, not always be the most important patient-centred outcome. For example, improved quality of life might be more important than a marginal survival gain in certain incurable cancer patients, anxiety symptoms might be the most important outcome for children with psychological problems, and pain might be the most important outcome for children with middle ear infection. Patient-reported outcome measures (PROMs) have been defined for some specific conditions; PROMs aim at assessing the quality of care delivered to patients from the patient perspective.[4] Nevertheless, all-cause mortality remains a reliable and patient-centred primary outcome in most circumstances.[1–3] When all-cause mortality is used as an outcome, all beneficial and harmful effects that a given intervention might have on risk of death are summed up and shown in the result. However, due to the relatively low occurrences of deaths in many trial intervention groups, a large number of randomised participants and a large number of randomised clinical trials (large required sample sizes) may be needed to confirm or reject a realistic intervention effect on all-cause mortality—and most often the information sizes are not large enough even in

Table 1 Examples of required sample sizes of trials with an experimental intervention and a control intervention (1:1) in which the intervention effect is similar on all-cause mortality and disease-related mortality

Outcome	Proportion dying in control group	Assumed relative risk reduction of experimental intervention	Required sample size (n participants)
All-cause mortality	10%	Relative risk reduction (RRR)=20%	8604
Disease-specific mortality	7%	RRR=20%	12660
All-cause mortality	20%	RRR=20%	3874
Disease-specific mortality	14%	RRR=20%	5902
All-cause mortality	30%	RRR=20%	2298
Disease-specific mortality	21%	RRR=20%	3650

In all scenarios, alpha=0.05 and beta=0.10 were used. Required samples sizes estimated in sample size and power (http://biostat.mc.vanderbilt.edu/wiki/Main/PowerSampleSize).

systematic reviews where several trials are combined.[5–7] Hence, it will not always be pragmatic to choose all-cause mortality as a primary outcome because of a presumed low statistical power.

Disease-related mortality (eg, cardiovascular mortality or breast-cancer mortality) is often used as a primary outcome in randomised clinical trials and systematic reviews. The rationale why disease-related mortality is used instead of all-cause mortality seems to be that disease-related mortality may more readily detect the experimental intervention effects. Disease-related mortality is theoretically what most interventions aim at influencing; disease-related intervention effects are not 'diluted' as all-cause mortality by events unrelated to the disease that may be occurring in both the experimental group and the control group (eg, traffic accidents). Intervention–effect estimates on all-cause mortality are indeed theoretically diluted and affected if events unrelated to the disease or the trial interventions are occurring. Although using disease-related mortality sounds attractive, we will in the following consider some methodological limitations of using this outcome instead of all-cause mortality.

Disease-related mortality may introduce loss of power

The required sample size to confirm or reject a minimal important difference when assessing risk of death, depends on the proportion of control participants dying. The proportion of control participants with a disease-related death will be less than (or equal to) the proportion of control participants who dies from all causes. If the same relative risk reduction (eg, a relative risk reduction of 20%) is used in sample size estimations, the required sample size will be larger (or equal to if the results are identical) when assessing disease-related mortality compared with assessing all-cause mortality (table 1).

Consequently, assessing disease-related mortality with a fixed sample size and a fixed relative risk reduction will result in a loss of power compared to using all-cause mortality.

On the other hand, if one assumes that a given absolute risk reduction is caused predominantly by an intervention effect on risk of disease-related deaths, it might be deduced that a smaller sample size will be sufficient when assessing disease-related mortality compared with assessing all-cause mortality (table 2). Furthermore, if there seems to be an absolute risk increase of the intervention on presumed non-disease-related deaths there will obviously be a higher power to detect an apparent beneficial effect on 'disease-related deaths'. However, this apparent advantage, of the disease-related mortality as an outcome, comes with the price that deaths caused by the intervention goes on undetected or even unreported unless the effect on all-cause mortality is reported.

It will always be unclear if the absolute risk reduction, when assessing disease-related mortality, is larger or smaller than (or similar to) the absolute risk reduction for all-cause mortality. If the absolute risk reduction is increased (indicating that some deaths related to the intervention are not registered as disease-specific) when using disease-related mortality instead of all-cause mortality, the loss of power due to fewer events may be partly counteracted. However, there might be a higher risk that it will result in a loss of power if disease-related mortality is used as an outcome instead of all-cause mortality, and the theoretical potential advantage of using disease-related mortality instead of all-cause mortality will often be questionable.

Risk of misinterpretation

Clinical events that are apparently unrelated to a disease might in fact be caused by the disease or the experimental trial

Table 2 Examples of required sample sizes of trials with an experimental intervention and a control intervention (1:1) in which the intervention effect is larger on disease-related mortality than on all-cause mortality.

Outcome	Proportion dying in control group	Absolute risk reduction	Assumed proportion dying in experimental group	Assumed relative risk reduction of experimental intervention	Required sample size (n participants)
All-cause mortality	10%	2%	8%	RRR=20%	8604
Disease-related mortality	7%	2%	5%	RRR=28.6%	5912
All-cause mortality	20%	4%	16%	RRR=20%	3874
Disease-related mortality	14%	4%	10%	RRR=28.6%	2764
All-cause mortality	30%	6%	24%	RRR=20%	2298
Disease-related mortality	21%	6%	15%	RRR=28.6%	1716

In all scenarios, alpha=0.05 and beta=0.10 were used. Required samples sizes estimated in sample size and power (http://biostat.mc.vanderbilt.edu/wiki/Main/PowerSampleSize).

intervention. For example, if a trial participant suddenly dies in a traffic accident the underlying cause might be a stroke, transitory cerebral ischaemia, hypoglycaemia, nausea or dizziness caused by the disease, or it might be adverse reactions caused by the experimental intervention. Performing blinded postmortem examinations of all dead trial participants is often impossible for pragmatic reasons, and even if blinded postmortem examinations of all dead trial participants are performed, the risk of erroneous classifications of causes of death will only be decreased, but not eliminated (eg, the participant's cause of death could be injuries caused by a traffic incident, but the traffic accident might be caused by an adverse reaction). When assessing disease-related mortality, the 'true' causality of a disease-related death will often be unclear and there is, therefore, a risk of misleading results. When results on disease-related mortality are interpreted in a clinical context, patients and uninformed clinicians might have the impression that all deaths caused by the disease are included in the disease-related mortality outcome. However, there may be a risk of erroneous classifications of certain causes of deaths. In addition, when assessing disease-related mortality outcomes, trial results might be biased if a given intervention has an effect on risk of non-disease-related deaths. For example, if an intervention decreases the risk of non-disease-related deaths the same group might experience more disease-related deaths because this 'competing risk' (risk of non-disease-related deaths) is reduced, that is, the participants live to experience the disease-related deaths. Caution is advised if the intervention effect estimates between disease-related mortality and all-cause mortality differ.[8] All-cause mortality should most often be considered the most valid mortality outcome,[2 3] and whenever disease-related mortality is reported one should always be informed about results of all-cause mortality as well.

Knowledge of treatment allocation might influence the decision to classify a death as disease-related and then lead to biased results

Classification of cause of death, disease-related or not, is often based on subjective assessments. This is for obvious reasons particularly a problem if the outcome assessors in the trial are not adequately blinded.[2] Independent 'blinded' outcome assessors (part of a blinded adjudication committee) might, for example, become unblinded when reading clinical records, or by obvious serious adverse reactions caused by the experimental intervention (eg, chemotherapy or surgery) might also compromise the validity of the blinded outcome assessment. If outcome assessors are not adequately blinded, results of disease-related mortality should always be interpreted with great caution.[9–11]

If the degree of observation or timing differs between the experimental group and control group

In a randomised clinical trial, there might be a difference between intervention groups in how and when the trial participants are observed, even if it is planned to assess outcomes similarly in the compared groups. For example, because of adverse reactions caused by an intervention, experimental participants might be followed more closely than the control participants due to more hospital admissions. If the participants in one of the compared groups are observed more closely than in the other group, it will be more plausible that the true reason for a death is recorded in that group, and this might result in a difference between the groups when assessing disease-related deaths when there might not be an actual difference. In other words, adverse reactions caused by the experimental intervention might lead to differences in the amount of data recorded between the two groups. Hence, even a perfectly blinded outcome assessor might judge similar deaths in the compared groups differently (as disease-related deaths or not) because there is more information about the trial participants in one of the groups leading to potentially biased results.

Deadly adverse events caused by trial interventions might not be classified as disease-related

Deadly adverse effects of a trial intervention for a given disease should be included in a disease-related mortality outcome. Otherwise, the disease-related mortality results will not reflect a balance between the beneficial and harmful effects of an experimental intervention on the disease in question. There is a risk that such adverse effects of trial interventions might not be classified as disease-related. For example, chemotherapy might cause cardiac arrhythmia resulting in death and such events might not be classified as 'cancer-related deaths'.

Furthermore, if a randomised clinical trial or a systematic review is warranted, it is often because the intervention effects of the intervention are unknown or uncertain. Hence, deadly unexpected true disease-specific deaths might not be adequately classified as disease-related. Results on disease-related mortality might, therefore, be biased because the impact of possible unknown adverse events and reactions might be overlooked in the results. It has, for example, been claimed that when breast cancer-related mortality is used instead of all-cause mortality, this results in biased and unreliable outcome results mainly because of differential misclassification of causes of death.[8 12]

It may not be valid to pool different definitions of disease-related outcomes in a meta-analysis

Different trials might use different definitions of how a disease-related death is defined, and this might compromise the interpretation of results across trials. If results of trials using different definitions of disease-related mortality are pooled in a meta-analysis, the validity of such a meta-analysis result might be questionable—such a meta-analysis may mix 'apples and pears'. Similar problems arise if different definitions of composite outcomes are pooled in a meta-analysis.[1] Even if the definitions of the disease-related mortality are described similarly in each trial, the subjective assessment of whether a death is disease-related or not might still differ substantially between trials. Visual inspection of forest plots and statistical tests might not have the sufficient power to show true statistical heterogeneity, and post hoc data driven choices of whether or not to pool trial results should be limited or avoided.[2 13]

Conclusions

Disease-related mortality outcomes may be used to assess the effects of the intervention within a 'mechanistic paradigm', but we have summarised several methodological limitations that must be considered when results on disease-related mortality outcomes are interpreted. Results of disease-related mortality outcomes should especially be interpreted with caution when the intervention effect estimates differ between disease-related mortality and all-cause mortality. For example, if results of disease-related mortality indicate a beneficial effect of an intervention but the results of all-cause mortality either indicate no difference between the groups or even a harmful effect, then there is either a risk that the results of disease-related mortality are biased because of the methodological limitations we have described, or the intervention

in question might reduce the risk of some causes of deaths (the presumed 'disease-related' causes) but might increase the risk of other 'non-disease-related' risks of death—otherwise the results of disease-related mortality and all-cause mortality would not differ. Due to the many methodological limitations, we believe that disease-related mortality should not be used as a primary outcome in many circumstances. Disease-related mortality should primarily be used as a secondary or an exploratory outcome, depending on the envisaged power of the outcome, and results of disease-related mortality should always be related to results of all-cause mortality.[13]

Contributors JCJ had the idea for the article and wrote multiple drafts including the final version and is the guarantor. All authors edited, advised and made amendments.

Funding The authors have not declared a specific grant for this research from any funding agency in the public, commercial or not-for-profit sectors.

Competing interests None declared.

Patient consent for publication Not required.

Provenance and peer review Not commissioned; externally peer reviewed.

ORCID iDs
Janus Christian Jakobsen http://orcid.org/0000-0002-3642-2120
Jørn Wetterslev http://orcid.org/0000-0001-7778-1771

References

1 **Garattini S**, Jakobsen JC, Wetterslev J, *et al*. Evidence-Based clinical practice: overview of threats to the validity of evidence and how to minimise them. *Eur J Intern Med* 2016;**32**:13–21.
2 **Higgins JPT**, Green S. The Cochrane Handbook for Systematic Reviews of Interventions, Version 5.1.0. *The Cochrane Collaboration* 2011. Available: www.cochrane-handbook.org [Accessed latest access December 2016].
3 **Jakobsen J**, Gluud C. The necessity of randomized clinical trials. *Br J Med Med Res* 2013;**3**:1453–68.
4 **Wilson HA**, Middleton R, Abram SGF, *et al*. Patient relevant outcomes of unicompartmental versus total knee replacement: systematic review and meta-analysis. *BMJ* 2019;**28**.
5 **Brok J**, Thorlund K, Gluud C, *et al*. Trial sequential analysis reveals insufficient information size and potentially false positive results in many meta-analyses. *J Clin Epidemiol* 2008;**61**:763–9.
6 **Turner RM**, Bird SM, Higgins JPT. The impact of study size on meta-analyses: examination of underpowered studies in Cochrane reviews. *PLoS One* 2013;**8**:e59202.
7 **Jakobsen JC**, Gluud C, Winkel P, *et al*. The thresholds for statistical and clinical significance – a five-step procedure for evaluation of intervention effects in randomised clinical trials. *BMC Med Res Methodol* 2014;**14**.
8 **Gøtzsche PC**, Jørgensen KJ, **Cochrane Breast Cancer Group**. Screening for breast cancer with mammography. *Cochrane Database Syst Rev* 2013;**156**.
9 **Hróbjartsson A**, Emanuelsson F, Skou Thomsen AS, *et al*. Bias due to lack of patient blinding in clinical trials. A systematic review of trials randomizing patients to blind and nonblind sub-studies. *Int J Epidemiol* 2014;**43**:1272–83.
10 **Hróbjartsson A**, Thomsen ASS, Emanuelsson F, *et al*. Observer bias in randomised clinical trials with binary outcomes: systematic review of trials with both blinded and non-blinded outcome assessors. *BMJ* 2012;**344**:e1119.
11 **Hróbjartsson A**, Thomsen ASS, Emanuelsson F, *et al*. Observer bias in randomized clinical trials with measurement scale outcomes: a systematic review of trials with both blinded and nonblinded assessors. *Can Med Assoc J* 2013;**185**:E201–11.
12 **Autier P**, Boniol M. Mammography screening: a major issue in medicine. *Eur J Cancer* 2018;**90**:34–62.
13 **Jakobsen JC**, Wetterslev J, Winkel P, *et al*. Thresholds for statistical and clinical significance in systematic reviews with meta-analytic methods. *BMC Med Res Methodol* 2014;**14**:120.