Journal of Clinical Epidemiology 108 (2019) 1–9

Journal of Clinical Epidemiology

REVIEWS

Methodological review showed correct absolute effect size estimates for time-to-event outcomes in less than one-third of cancer-related systematic reviews

Nicole Skoetz[a,*,1], Marius Goldkuhle[a,1], Aaron Weigl[a], Kerry Dwan[b], Valérie Labonté[c], Philipp Dahm[d], Joerg J. Meerpohl[e], Benjamin Djulbegovic[f], Elvira C. van Dalen[g]

[a]*Department I of Internal Medicine, University Hospital of Cologne, Kerpener Str. 62, Cologne 50937, Germany*
[b]*Editorial and Methods Department, Cochrane, St Alban's House, 57-59 Haymarket, London SW1Y 4QX, UK*
[c]*Institute for Evidence in Medicine (for Cochrane Germany Foundation), Medical Center, Faculty of Medicine, University of Freiburg, Breisacher Str. 153, Freiburg 79110, Germany*
[d]*Department of Urology, Minneapolis Veterans Administration Medical Center and University of Minnesota, Minneapolis VA Health Care System, Urology Section 112D, One Veterans Drive, Minneapolis, MN 55417, USA*
[e]*Institute for Evidence in Medicine (for Cochrane Germany Foundation), Medical Center, Faculty of Medicine, University of Freiburg, Breisacher Str. 153, Freiburg 79110, Germany*
[f]*Department of Supportive Medicine and Department of Hematology, City of Hope, 1500 Duarte Rd, Duarte, CA 91010, USA*
[g]*Department of Pediatric Oncology, Emma Children's Hospital, Amsterdam UMC, University of Amsterdam, PO Box 22660, Amsterdam 1100 DD, the Netherlands*

Accepted 5 December 2018; Published online 10 December 2018

Abstract

Objectives: To evaluate in how many cancer-related Cochrane reviews hazard ratio (HR)-based absolute effects in summary of findings (SoF) tables have been correctly calculated and reported.

Study Design and Setting: We identified all Cochrane cancer intervention reviews that reported an HR for at least one outcome and provided a SoF table, published between January 2011 and December 2017 in the Cochrane Database of Systematic Reviews.

Results: In 28 reviews (29%) of 96 included Cochrane reviews, absolute effects in the SoF tables were calculated in a correct manner. In 23 reviews (24%), absolute effects had been correctly calculated, but there was no explanation given why authors calculated event-free survival (e.g., overall survival) throughout the review but reported number of events in SoF tables (e.g., death). Twelve reviews (13%) provided incorrect absolute effects. For seven reviews (7%), it was unclear if absolute effects were correctly calculated. In 26 (27%) reviews, no absolute effects based on the given HR were calculated.

Conclusions: In less than one-third of cancer-related Cochrane reviews, absolute effect size estimates were correctly calculated and reported. There is a need for guidance on how to calculate and report absolute effect estimates based on HR data.

Keywords: Time-to-event; Hazard ratio; Absolute effects; Summary of findings; Methodological; Review

Conflict of interest statement: All authors have completed the ICMJE uniform disclosure form at www.icmje.org/coi_disclosure.pdf and declare: authors received travel funds from Cochrane only to present findings at one international meeting and one international conference for the submitted work; no financial relationships with any organizations that might have an interest in the submitted work in the previous 3 years; no other relationships or activities that could appear to have influenced the submitted work." NS, AW and MG are part of Cochrane Haematological Malignancies, EvD part of Cochrane Childhood Cancer, VL and JJM part of Cochrane Germany, and KD staff member of the Cochrane Editorial Unit.

Funding: Partial support from Cochrane was received to present data during the Cochrane Midyear Meeting 2017 in Geneva, Switzerland, and during the Global Evidence Summit 2017 in Cape Town, South Africa.

Authors' Contributions: N.S. initiated the project, designed the data collection tools and analysis, selected studies, extracted and analysed data, and drafted and revised the article. M.G. was involved in the design of the data collection tools and analysis, selected studies, extracted and analysed data, prepared tables for presentation, and revised the article. EvD designed the data collection tools and analysis, selected studies, extracted and analysed data, and drafted and revised the article. A.W. selected studies, extracted and analysed data, and revised the article. P.D., B.G., K.D., J.J.M., and V.L. extracted data and revised the article. All authors approved the final version of the article. All authors had full access to all of the data (including statistical reports and tables) in the study and can take responsibility for the integrity of the data and the accuracy of the data analysis.

[1] Contributed equally.

* Corresponding author. Tel.: +49221 478 96651; fax: +49221 478 96654.

E-mail address: nicole.skoetz@uk-koeln.de (N. Skoetz).

https://doi.org/10.1016/j.jclinepi.2018.12.006
0895-4356/

What is new?

Key findings

- We identified errors in the presentation of absolute effect measures in summary of findings (SoF) tables of cancer-related Cochrane reviews and describe common pitfalls to avoid.
- The errors in calculation of hazard ratios can be minimized if the review authors first assess direction of effect measure (event or nonevent) and then accordingly calculate the respective corresponding absolute effects. For example, the event that is typically measured is mortality (death), but the outcome reported is often overall survival (1- mortality), which goes into opposite direction.

What this adds to what is known?

- The appropriate presentation of absolute effect size estimates based on hazard ratios in SoF tables has not been evaluated previously.

What is the implication, what should change now?

- There is an urgent need for additional training materials and guidance for authors on how to calculate and present absolute effects based on time-to-event data.

1. Introduction

Absolute effect estimates are more understandable to patients, clinicians, and other users of systematic reviews than relative effect measures and are the recommended effect measure to communicate risks [1]. They reflect the clinical importance of outcomes and can ground exaggerated outcome perceptions of clinicians and patients, which may occur if solely relative effects are reported [2−4]. Absolute effects provide important supplementary information that considers risk-specific control event rates over a given time period. Absolute effect estimates are a routine part of the user-friendly format of 'summary of findings" (SoF) tables or evidence profiles [5]. Reviews published by Cochrane, which is widely known for establishing methodological standards for conducting and reporting high quality systematic reviews, regularly include such SoF tables. SoF tables are prepared according to the GRADE guidance papers and can be calculated using software products such as GRADEpro GDT (gradepro.org) or MAGICapp (app.magicapp.org) [5].

In many fields of health care, in particular oncology, analyses that assess the time to a given event for one or several groups of patients are commonly used. For patients with cancer, one of the most relevant outcomes is overall survival (OS). It describes the survival time of patients until death for any reason which occurs within a certain period of follow-up. In addition, another outcome like progression-free survival (PFS), that is the survival time without detectable worsening of disease (progress, relapse, death) over a considered time-period, is often assessed. This outcome measure provides complimentary information for OS. Both outcomes are so called time-to-event outcomes, as they involve the assessment of both whether a particular event occurs, and also when it occurs [6]. To compare time-to-event outcomes of two groups of patients, hazard ratios (HR) with corresponding confidence intervals that provide relative effect size estimates are used.

The calculation of absolute effects based on HR is error prone because both beneficial (event-free survival) and adverse effects (events) can easily be confused and because calculation of HR is based on difficult to interpret exponential functions. As there is currently no written guidance on how to calculate absolute effects based on HR, and how to best present these in SoF tables, it might be especially difficult for review authors to do this properly, as well as for journal editors and peer reviewers to identify mistakes. Another potential challenge arises around the consistent definition of time-to-event outcomes across all parts of the review including the abstract, results section, and the SoF table. Because of that absolute effects based on reported time-to-event outcomes are difficult to calculate, present, and interpret. Although often the event is measured (e.g., death), the event-free survival (e.g., OS) is reported throughout individual studies and the corresponding systematic review, review authors must be aware to calculate the respective absolute effect (for the event or for event-free survival). Until a recent update (September 2018), the GRADEpro GDT software allowed calculation of absolute effects based on HR only for outcomes and baseline risks corresponding to events (such as mortality) but not for event-free survival (like overall survival).

In this methodological review, we evaluated in how many current cancer-related Cochrane reviews absolute effects based on HR in SoF tables have been correctly calculated and reported.

2. Materials and methods

We report our methodological review according to the Preferred Reporting Items for Systematic Reviews and Meta-analyses (PRISMA) [7]. The project was conducted according to an a priori developed protocol. As this is a methodological review, it was not eligible for a registration in the International Prospective Register of Systematic Reviews (PROSPERO). The protocol can be accessed on request from the review authors.

2.1. Eligibility criteria

Cancer-related Cochrane intervention reviews and overviews of reviews were eligible for inclusion if they

provided an SoF table in which the effect size for at least one time-to-event outcome based on a pooled HR was included. Reviews in which an HR for an outcome was given from a single study only were also eligible. Reviews that reported HR from several studies for the same outcome but did not pool the respective HR were excluded. This was to ensure that only one single HR per outcome in the review was reported, which could be used to calculate one corresponding absolute effect estimate. The results for this time-to-event outcome must have been mentioned in at least one of the following sections: abstract, methods, or results. Therapeutic, preventive, or prophylactic intervention reviews were eligible. Reviews not meeting all these criteria were excluded. We excluded reviews in which effects were presented in risk ratios or odds ratios (ORs) only. We used the original English version of each Cochrane review for data extraction and assessment.

2.2. Study identification and selection

We systematically identified all Cochrane intervention reviews that examined questions in the context of oncology, irrespective of type of cancer, stage of disease, type of intervention, outcomes assessed, or study design of included studies. This was done by using the function: ‘‘Browse by topic’’ and by choosing the following options: ‘‘Cancer’’ and ‘‘Stage: Review’’ in the Cochrane Database of Systematic Reviews. This function is based on tags, which are manually applied by the operators of the Cochrane Library.

The current version of SoF tables was first described in 2010, and the first GRADE guidelines were published in 2011 [5,8]. We did not expect any Cochrane Reviews to include SoF tables before 2011. Therefore, we restricted the included Cochrane Reviews to a 6-year period between January 2011 and December 2017. In case a review was published more than once during this time period, for example, as primary publication and as an update, we included only the most recent publication. Three authors (N.S., A.W., and M.G.) independently screened titles, abstracts, and full texts identified in the Cochrane Database of Systematic Reviews for eligibility according to the inclusion criteria. Review screening was carried out in one step because it is necessary to view a review full text to assess the availability of SoF tables and time-to-event outcomes. If any disagreement regarding the inclusion of reviews occurred, the authors tried to resolve it by discussion or involved another author (EvD) until consensus was achieved.

2.3. Data extraction

All included reviews were randomly allocated to eight members of the research team (N.S., M.G., P.D., A.W., V.L., J.J.M., K.D., EvD) to be extracted independently in duplicate. In case one of these individuals was involved in the publication of a specific Cochrane review (e.g., as an author or member of the editorial group), this particular review was reassigned to other, nonconflicted members of the research team. We used a dedicated pilot-tested extraction form. Any discrepancies during data extraction were resolved through discussion or if necessary with involvement of a third author.

To classify the baseline characteristics of the included SRs, we extracted information on the cancer type (e.g., breast, lung, colorectal), but also ‘‘cancer in general’’ and ‘‘mixed’’ (multiple diseases, but not cancer in general) and year of publication. To examine how absolute effects were calculated, we extracted data for the first two time-to-event outcomes, which were reported in a SoF table and the description for these outcomes with corresponding HR and their 95% confidence intervals as reported in abstract, methods section, and/or results section. For Cochrane reviews in which a (pooled) HR was given, we assumed that this effect measure and its associated confidence intervals had been correctly calculated. We extracted the first two HR outcomes because in cancer reviews these are commonly OS and PFS. Overall survival as an outcome measures includes only the single event "death", whereas PFS includes the events "death," "progression," and "relapse". Comparing the reported baseline risks for these two outcomes allowed to determine whether the baseline risks applied to events or event-free survival (as described in the next paragraphs). In addition, we extracted the description of the same outcomes as reported in the SoF table and the absolute effects as well as information regarding assumption of the underlying baseline risk. If several SoF tables were included, we used data from the first SoF table that listed an eligible time-to-event outcome.

For each outcome, we interpreted the meaning of an HR $<$ 1, that is, whether this favored the control or intervention arm, based on the choice of the event as documented in the methods section of the review. If absolute effects were reported in the SoF table and review authors provided information on how the control group risk had been determined, we extracted this information. In case review authors did not provide information on how they determined the control group risk, we assessed whether they used the number of people with the event (e.g., people being dead at a specific time point) or the number of people event-free (e.g., people being alive at a specific time point) to calculate absolute effects for the intervention group. If at least two HR outcomes, like OS and PFS, were reported in the review, we compared the absolute numbers in the estimated control group risk for both outcomes, as shown in Figs. 1−3. If the absolute number for the outcome OS (based on the event people being dead) was lower than for the outcome PFS (based on the event people with progressive disease), we assumed authors had used number of people being event-free to calculate absolute effects for the intervention arm (see Fig. 1).

If the number was lower for OS than for PFS (see Fig. 2), we assumed authors had used number of people with the event to estimate the control group risk and calculate numbers for the intervention arm.

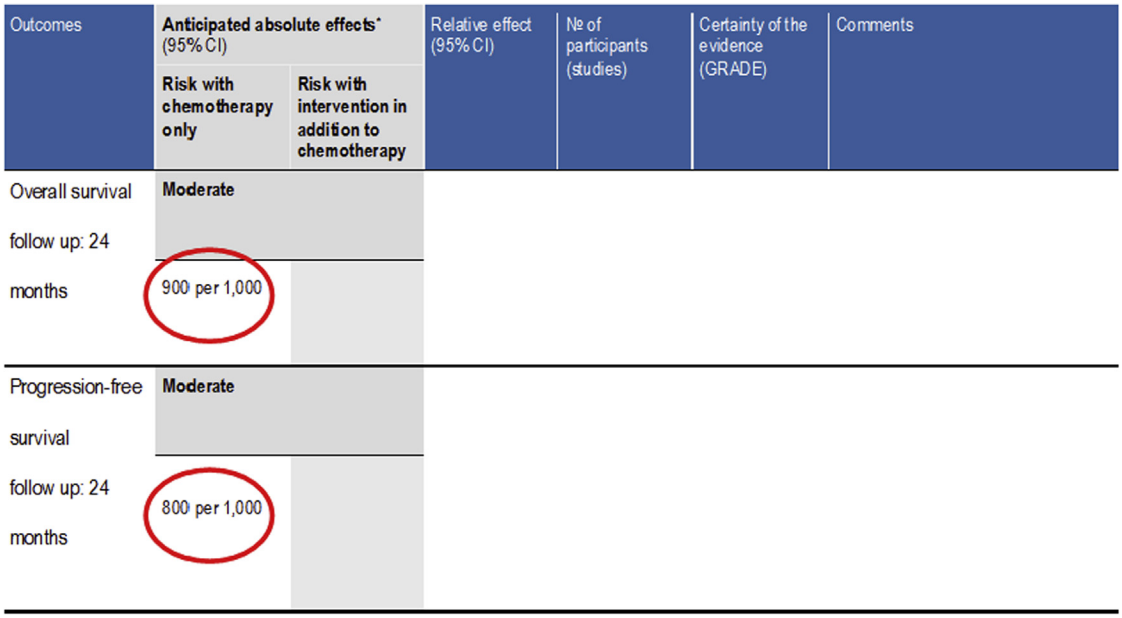


Fig. 1. Extract from an exemplar SoF table. Numbers for the estimated control group risks are marked in red (for the outcomes overall survival and the combined outcome progression-free survival). (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

If the absolute numbers in both estimated control groups were identical (see Fig. 3), it was impossible to judge whether authors used absolute numbers of people with event or people being event-free to determine control group risk for their calculations. One would expect a higher overall number of people who survived compared to the number of people who survived without progression (PFS), as OS is based on deaths only, but PFS is the sum of people being dead, with relapse, or progression.

In case authors reported event-free survival like OS and PFS throughout the review but used number of events (i.e., people being dead or with progress) to calculate absolute effects, we extracted information on how authors commented on this in the SoF table like ''Instead of OS, mortality is reported in this SoF table, for technical reasons''.

2.4. Recalculation of absolute effect size estimates

To check whether review authors calculated the absolute effects from the HR outcome correctly, we recalculated absolute effects based on methods described by Tierney et al. [9]. This article recommends using

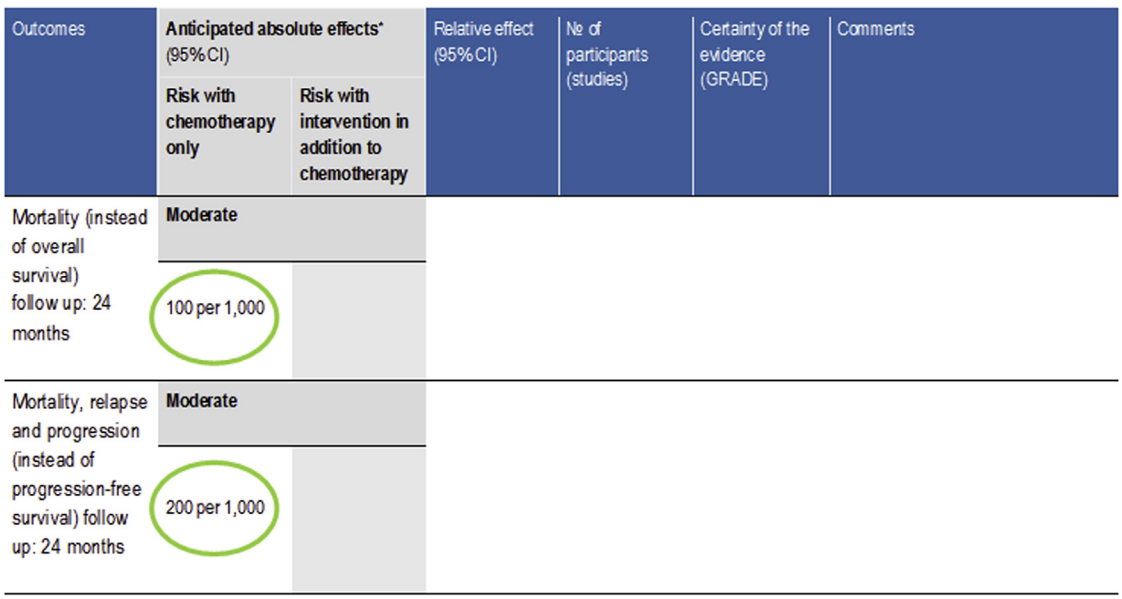


Fig. 2. Extract from an exemplar SoF table. Numbers for the estimated control group risks are marked in green (for the outcomes mortality and the outcome sum of mortality, relapse, and progress). (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

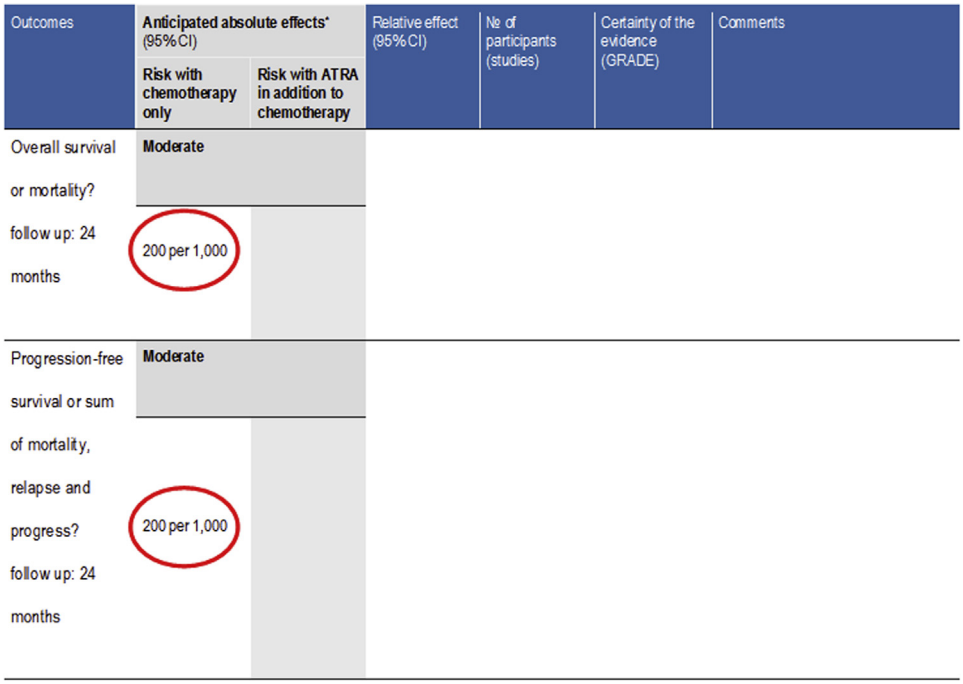


Fig. 3. Extract from an exemplar SoF table. Numbers for the estimated control group risks are marked in red (for the outcomes overall survival and progression-free survival). (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

event-free survival data to interpret HR using the following formula:

$$\text{Corresponding intervention risk, per 1000} = (\exp[\ln(\text{proportion of patients event} - \text{free}) \times HR]) \times 1000, \text{ per 1000}$$

These calculations were made based on the HR and baseline risk estimates, which were reported in the Cochrane review.

3. Results

3.1. Search results

As shown in the study flowchart (see Fig. 4), our search led to 483 Cochrane reviews, which were determined to be cancer-related. Study selection showed that 210 of these included an SoF table and of these, 96 reported for at least one time-to-event outcome an HR in the SoF table. The references of the included Cochrane reviews are listed in the online Appendix.

3.2. General characteristics of included reviews

The vast majority of studies, 52 (54%), was published in 2017 or 2016, see also Appendix Table 1. Only one review, without an update publication, has been published in 2011. The largest number of reviews evaluated interventions for hematological malignancies (21 reviews, 22%), 12 reviews (13%) assessed interventions for patients with breast cancer, and four reviews (4%) evaluated ‘‘cancer in general’’. No overviews of reviews were identified.

Five reviews (5%) reported only one outcome with an HR. Ten reviews (10%) mentioned in a comment or footnote, how they determined the baseline risk in the control group, but none of this information was useful to evaluate whether review authors used the number of people with an event or the number of people being event-free to calculate absolute effect size estimates. This was because authors did not report whether the patients were alive or dead at the respective time point. Also, transparent reporting of information where baseline risk data is derived from was often missing.

3.3. Presentation of absolute effect estimates in included reviews

Table 1 summarizes the presentation of absolute effect estimates in the included reviews.

3.3.1. Absolute effects correctly calculated with consistent labeling of outcomes throughout the review

Twenty-eight reviews (29%) correctly calculated absolute effects and labeled the time-to-event outcomes in a

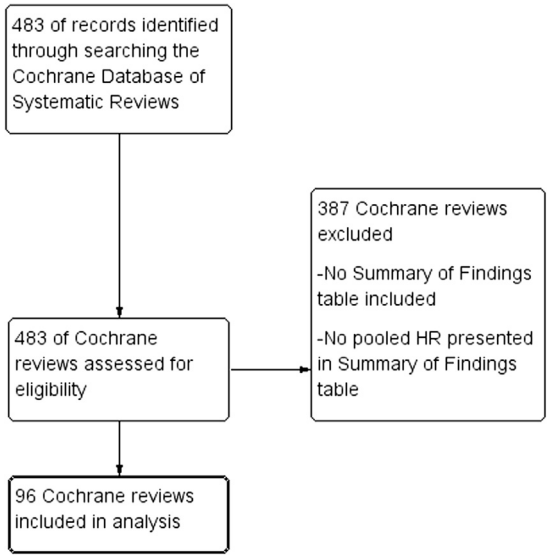


Fig. 4. PRISMA flow diagram of Cochrane reviews.

consistent manner throughout the review, that is, making a clear distinction between people being event-free (e.g., people alive at a specific time point) and people with an event (e.g., people dead at specific time point). Accordingly, time-to-event outcomes were labeled consistently throughout abstract, methods, results section, and SoF table of the respective reviews (see an example in Fig. 5).

3.3.2. Absolute effects correctly calculated but inconsistent presentation of outcomes in SoF table and other parts of the review

Twenty-three reviews (24%) correctly calculated the absolute effects. However, there was inconsistency in how labels for the time-to-event outcomes were used throughout the review. In the SoF table, events (e.g., number of deaths) were used to calculate the absolute effect, whereas in other parts of the review, event-free survival (e.g., OS) was reported, without any explanation in the comment section as shown in Fig. 5, why the name of the outcome changed within the review.

3.3.3. Incorrect calculation of absolute effects

Twelve reviews (13%) provided incorrect absolute effect estimates. The underlying reason was that instead of correctly entering the number of people with the event, the review authors entered the number of people without an event into calculation software, then applying the HR. This led to incorrect results with less people instead of more being alive in the favored arm (see Fig. 6). The review authors reported these incorrect results in the SoF table only; none of the review authors reported these incorrect numbers in the abstract, plain language summary, results, or discussion section.

3.3.4. Unclear results

In seven reviews (7%), it was unclear how review authors determined the control group risk and whether direction of results was correct. This was the case when the control group risk for both outcomes of interest (e.g., OS and PFS) was identical, as shown in Fig. 3.

3.3.5. No absolute effects calculated

Twenty-six reviews (27%) did not calculate an absolute effect. However, five of these reviews reported mean survival ranges for both the control and intervention arm or weighted mean survival with 95% confidence intervals in the SoF table but without any explanation how these had been calculated. Thus, their provenance remained unclear, and it could not be judged whether they were correct or incorrect.

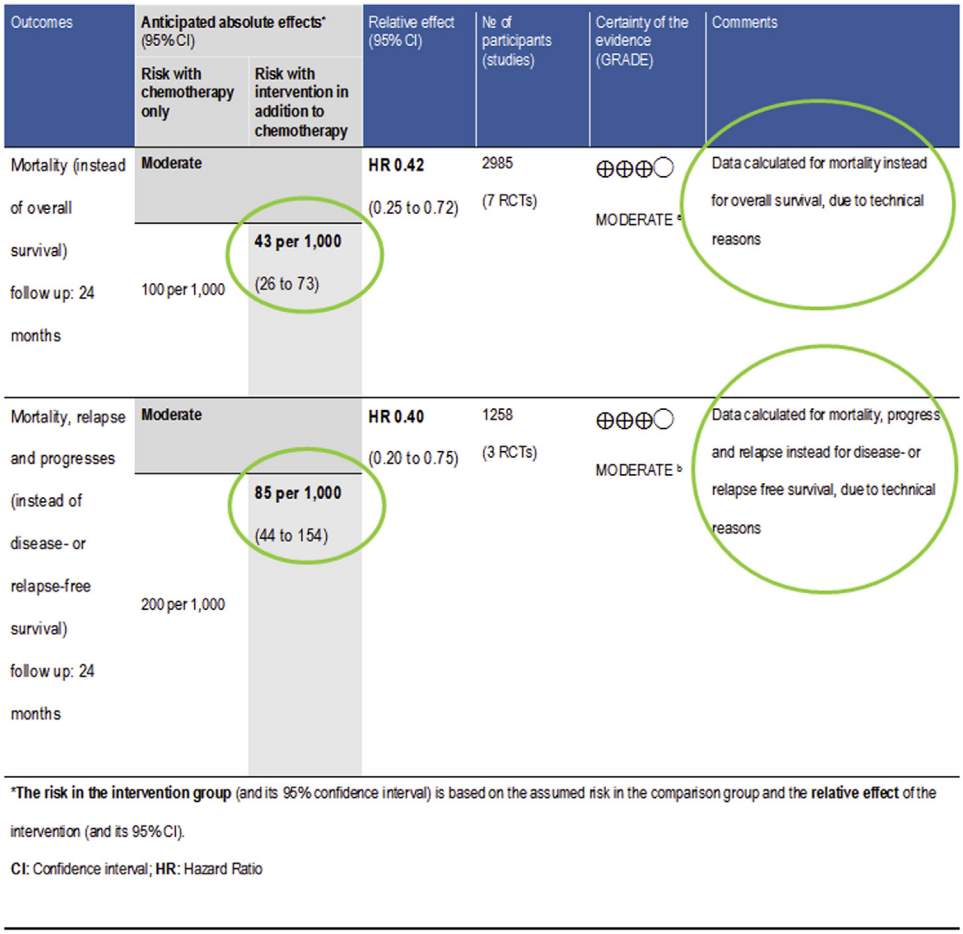
4. Discussion

4.1. Principal findings

This methodological review shows that absolute effects based on time-to-event outcomes are calculated correctly and presented in a readily interpretable way in less than one in three Cochrane reviews related to cancer (out of 96 reviews). In about every fourth review, the absolute effects were correctly calculated but the respective outcomes were labeled inconsistently and potentially misleading without any comment why authors calculated event-free survival (e.g., OS) throughout the review but reported number of events (e.g., death) in SoF tables. Twelve percent provided incorrect absolute effects in the SoF tables, because inappropriate data were entered into the calculation software. As the review authors did not report the results of the incorrect calculations in the abstract, results section, or

Table 1. Presentation of absolute effect estimates in the included Cochrane reviews

Calculation of absolute effects and labeling of outcomes	Cochrane reviews (N = 96)	Figure
Absolute effects correctly calculated	28 + 23 (53%)	5
Consistent labeling of outcomes throughout the review	28 (29%)	
Inconsistent labeling of outcomes throughout the review	23 (24%)	
Absolute effects incorrectly calculated	12 (13%)	6
Unclear results	7 (7%)	3
No absolute effects calculated	26 (27%)	



Outcomes	Anticipated absolute effects* (95% CI)		Relative effect (95% CI)	№ of participants (studies)	Certainty of the evidence (GRADE)	Comments
	Risk with chemotherapy only	Risk with intervention in addition to chemotherapy				
Mortality (instead of overall survival) follow up: 24 months	Moderate 100 per 1,000	**43 per 1,000** (26 to 73)	**HR 0.42** (0.25 to 0.72)	2985 (7 RCTs)	⊕⊕⊕◯ MODERATE [a]	Data calculated for mortality instead for overall survival, due to technical reasons
Mortality, relapse and progresses (instead of disease- or relapse-free survival) follow up: 24 months	Moderate 200 per 1,000	**85 per 1,000** (44 to 154)	**HR 0.40** (0.20 to 0.75)	1258 (3 RCTs)	⊕⊕⊕◯ MODERATE [b]	Data calculated for mortality, progress and relapse instead for disease- or relapse free survival, due to technical reasons

*The risk in the intervention group (and its 95% confidence interval) is based on the assumed risk in the comparison group and the **relative effect** of the intervention (and its 95% CI).

CI: Confidence interval; **HR**: Hazard Ratio

Fig. 5. Extract from an exemplar SoF table. Numbers for the correctly calculated absolute effects and comments are marked in green, assuming the HR < 1 favors the intervention group. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

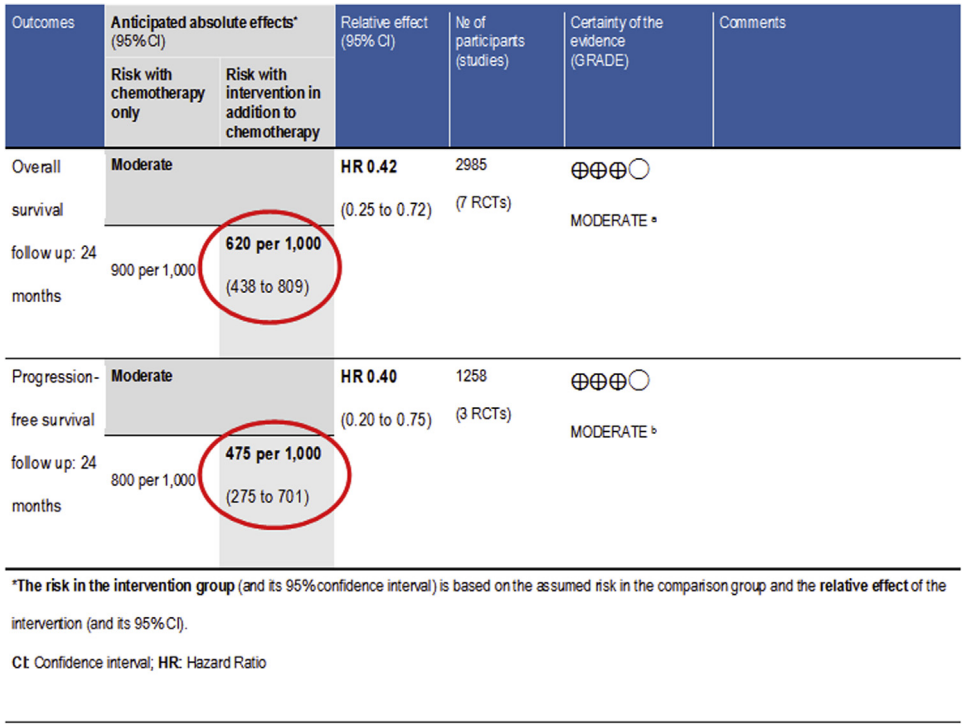
conclusion, this error is not transferred to other sections of the respective reviews. For 7%, it was completely unclear if absolute effects were correctly calculated, because it was unclear whether control group risk numbers described people with event or people without event. In the remaining 27%, no absolute effect had been calculated at all.

4.2. Strengths and limitations of this methodological review

We performed this methodological review based on an a priori developed protocol. The strength to the validity of our findings is based on a comprehensive search, a screening process done in duplicate, a piloted data extraction form, and data abstraction in duplicate. The reliability of this work is ensured through adherence to the review methods proposed by Cochrane and reporting in accordance with the PRISMA standards. The small sample size might have influenced our results; however, at the date of the search, no more cancer-related Cochrane reviews including time-to-event outcomes described as HR in a SoF table were available. Another limitation is that we evaluated only cancer-related reviews; therefore, our findings are directly applicable to cancer reviews only. Cochrane reviews evaluating patients with other types of disease and time-to-event outcomes should also be assessed, as these reviews might report other outcomes like time to hospitalization, time to discharge, or time to recovery, which could result in different findings. Although the focus of this review was on Cochrane reviews only, which has recently mandated the inclusion of absolute effects and SoFs, we expect similar issues to affect non-Cochrane reviews. We evaluated 10 high-impact cancer journals publishing systematic reviews within the same time period as mentioned above but could not identify any review out of 177 reporting absolute effects for HR outcomes or presenting SoF tables (unpublished data).

4.3. Strengths and limitations in relation to other methodological reviews

To date, there has not been other research directed to the reporting of absolute effect size estimates specifically



Outcomes	Anticipated absolute effects* (95% CI)		Relative effect (95% CI)	№ of participants (studies)	Certainty of the evidence (GRADE)	Comments
	Risk with chemotherapy only	Risk with intervention in addition to chemotherapy				
Overall survival follow up: 24 months	Moderate 900 per 1,000	**620 per 1,000** (438 to 809)	**HR 0.42** (0.25 to 0.72)	2985 (7 RCTs)	⨁⨁⨁◯ MODERATE [a]	
Progression-free survival follow up: 24 months	Moderate 800 per 1,000	**475 per 1,000** (275 to 701)	**HR 0.40** (0.20 to 0.75)	1258 (3 RCTs)	⨁⨁⨁◯ MODERATE [b]	

*The risk in the intervention group (and its 95% confidence interval) is based on the assumed risk in the comparison group and the **relative effect** of the intervention (and its 95% CI).

CI: Confidence interval; **HR**: Hazard Ratio

Fig. 6. Extract from an exemplar SoF table. Numbers for the incorrectly calculated absolute effect, assuming that HR < 1 favors the intervention group. The incorrect numbers are marked in red. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

for time-to-event outcomes in SoF tables. With regard to its specific focus and relevance to all investigators using time-to-event outcomes in SoF tables, our review is unique.

Prior methodological work assessing the frequency of reporting of absolute effects any outcome type systematic review revealed that they are rarely reported. Alonso-Coello et al. assessed 98 Cochrane and 104 non-Cochrane systematic reviews, which were published in 2010 and not limited to a certain field of disease. Overall only 36.1% of these reviews presented absolute effect estimates for the most patient-important outcome. In 32 systematic reviews that included SoF tables, all authors calculated absolute effects. Twenty-nine of these reviews were Cochrane reviews and three reviews were non-Cochrane reviews. The relative effect measures on which the absolute effects were based were primarily risk ratios or ORs; a smaller amount of only 6.4% were calculated from HRs. If absolute effects were reported, the source of the baseline risk applied for calculation was often not given [10]. This is in accordance with our findings showing sparse reporting of where baseline risk data are derived from. Agarwal et al. performed a methodological review to assess the frequency of reporting of absolute effects in the abstracts of systematic reviews. They included 96 Cochrane and 94 non-Cochrane reviews published in 2010 and revealed that absolute effects were reported in the abstracts of 22.5% of the respective systematic reviews. Again, the relative effect estimates corresponding to the calculated absolute effects where predominantly relative risks and ORs, only a small number (5.8%) were HR [2].

Our review demonstrated incorrect calculation and reporting of absolute effects based on HR in the SoF tables of cancer-related Cochrane reviews. Prior work suggested flaws in the calculation of absolute effects of clinical studies. A review evaluating 734 randomized controlled trials, published in high impact general medical journals of which 373 investigated time-to-event outcomes, found that only half of randomized controlled trials reporting number-needed-to-treat or number-needed-to harm from such outcomes used appropriate calculation methods [11]. Prior studies have also addressed the challenge of calculating numbers-needed-to-treat for time-to-event data in the setting of competing risks and reviewed the potential issue of varying follow-up times [12,13].

4.4. Meaning of this methodological review: explanations, implications, and further research

Our methodological assessment shed light on the problems review authors face when they try to calculate absolute effects based on time-to-event data in SoF tables. There is currently no written guidance on calculating HR-based absolute effects and how best to present them in SoF tables. Therefore, it may be particularly difficult for reviewers to do this properly and also for journal editors and peer reviewers to identify mistakes. This work demonstrates the need for additional training and guidance of

review authors working with time-to-event data. A workshop addressing this issue was held at an international conference already and further workshops are planned. Also, additional training materials in the form of written materials, online modules, and webinars for review authors and other GRADE users on how to calculate absolute effects are currently being developed and will soon be disseminated. Moreover, the GRADEpro GDT software has been adapted in September 2018 to provide review authors the choice as to whether to enter the number of people with events or number of people without events for the control group risk. This will allow better consistency of use of outcomes throughout the review. In addition, there appears need for additional oversight of review authors to identify incorrect and misleading information before publication.

A further key aspect to consider is how review authors should estimate the absolute risk in the control group for time-to-event outcomes: should it be based on one study or on all included studies or on data from representative observational studies and which time point should be used? Confidence intervals of calculated absolute effects do not incorporate uncertainty in the assumed control risks and are not considered by the calculation according to Tierney et al. [6]. This is of special concern if we look at long-term survival with a low or moderate mortality and a corresponding high number of censored patients (i.e., a low number of patients under risk and a high censoring rate). These aspects will be considered in another article.

5. Conclusion

Based on our systematic review, in less than one in three cancer-related Cochrane reviews that included at least one time-to-event outcome, absolute effect size estimates were correctly calculated and appropriately reported. This was due to missing comments and/or entering incorrect numbers into the GRADEpro GDT software. There is an urgent need for additional training materials and guidance for review authors, editors, and peer-reviewers on how to calculate and present absolute effects based on HR data.

Supplementary data

Supplementary data related to this article can be found at https://doi.org/10.1016/j.jclinepi.2018.12.006.

References

[1] Gigerenzer G, Gaissmaier W, Kurz-Milcke E, Schwartz LM, Woloshin S. Helping doctors and patients make sense of health statistics. Psychol Sci Public Interest 2007;8:53−96.
[2] Agarwal A, Johnston BC, Vernooij RWM, Carrasco-Labra A, Brignardello-Petersen R, Neumann I, et al. Authors seldom report the most patient-important outcomes and absolute effect measures in systematic review abstracts. J Clin Epidemiol 2017;81:3−12.
[3] Malenka DJ, Baron JA, Johansen S, Wahrenberger JW, Ross JM. The framing effect of relative and absolute risk. J Gen Intern Med 1993;8:543−8.
[4] Naylor C, Chen E, Strauss B. Measured enthusiasm: does the method of reporting trial results alter perceptions of therapeutic effectiveness? Ann Intern Med 1992;117:916−21.
[5] Guyatt G, Oxman AD, Akl EA, Kunz R, Vist G, Brozek J, et al. GRADE guidelines: 1. Introduction-GRADE evidence profiles and summary of findings tables. J Clin Epidemiol 2011;64:383−94.
[6] van Dalen EC, Tierney JF, Kremer LCM. Tips and tricks for understanding and using SR results. No. 7: time-to-event data. Evidence Based Child Health A Cochrane Rev J 2007;2:1089−90.
[7] Moher D, Liberati A, Tetzlaff J, Altman DG. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. BMJ 2009;339:b2535.
[8] Rosenbaum SE, Glenton C, Oxman AD. Summary-of-findings tables in Cochrane reviews improved understanding and rapid retrieval of key information. J Clin Epidemiol 2010;63:620−6.
[9] Tierney JF, Stewart LA, Ghersi D, Burdett S, Sydes MR. Practical methods for incorporating summary time-to-event data into meta-analysis. Trials 2007;8:16.
[10] Alonso-Coello P, Carrasco-Labra A, Brignardello-Petersen R, Neumann I, Akl EA, Vernooij RWM, et al. Systematic reviews experience major limitations in reporting absolute effects. J Clin Epidemiol 2016;72:16−26.
[11] Hildebrandt M, Vervölgyi E, Bender R. Calculation of NNTs in RCTs with time-to-event outcomes: a literature review. BMC Med Res Methodol 2009;9:21.
[12] Bender R, Kromp M, Kiefer C, Sturtz S. Absolute risks rather than incidence rates should be used to estimate the number needed to treat from time-to-event data. J Clin Epidemiol 2013;66:1038−44.
[13] Gouskova NA, Kundu S, Imrey PB, Fine JP. Number needed to treat for time-to-event data with competing risks. Stat Med 2014;33:181−92.