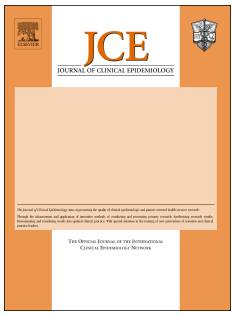


Meta-epidemiological review identified variable reporting and handling of time-to-event analyses in publications of trials included in meta-analyses of systematic reviews

Marius Goldkuhle, Caroline Hirsch, Claire Iannizzi, Ana-Mihaela Bora, Ralf Bender, Elvira C. van Dalen, Lars G. Hemkens, Marialene Trivella, Ina Monsef, Nina Kreuzberger, Nicole Skoetz

PII: S0895-4356(23)00136-1

DOI: https://doi.org/10.1016/j.jclinepi.2023.05.023

Reference: JCE 11098

To appear in: *Journal of Clinical Epidemiology*

Received Date: 7 April 2023

Revised Date: 17 May 2023

Accepted Date: 25 May 2023

Please cite this article as: Goldkuhle M, Hirsch C, Iannizzi C, Bora A-M, Bender R, van Dalen EC, Hemkens LG, Trivella M, Monsef I, Kreuzberger N, Skoetz N, Meta-epidemiological review identified variable reporting and handling of time-to-event analyses in publications of trials included in meta-analyses of systematic reviews, *Journal of Clinical Epidemiology* (2023), doi: https://doi.org/10.1016/j.jclinepi.2023.05.023.

Meta-epidemiological review identified variable reporting and handling of time-to-event analyses in publications of trials included in meta-analyses of systematic reviews

Marius Goldkuhle[a], Caroline Hirsch[b], Claire Iannizzi[c], Ana-Mihaela Bora[d], Ralf Bender[e], Elvira C van Dalen[f], Lars G Hemkens[g], Marialene Trivella[h], Ina Monsef[i], Nina Kreuzberger[j]*, Nicole Skoetz[k]*

[a] Evidence-based Medicine, Department I of Internal Medicine, Faculty of Medicine and University Hospital Cologne, University of Cologne, Kerpener Str. 62, 50937 Cologne, Germany; e-mail: marius.goldkuhle@uk-koeln.de

[b] Evidence-based Medicine, Department I of Internal Medicine, Faculty of Medicine and University Hospital Cologne, University of Cologne, Kerpener Str. 62, 50937 Cologne, Germany; e-mail: caroline.hirsch@uk-koeln.de

[c] Evidence-based Medicine, Department I of Internal Medicine, Faculty of Medicine and University Hospital Cologne, University of Cologne, Kerpener Str. 62, 50937 Cologne, Germany; e-mail: claire.iannizzi@uk-koeln.de

[d] Evidence-based Medicine, Department I of Internal Medicine, Faculty of Medicine and University Hospital Cologne, University of Cologne, Kerpener Str. 62, 50937 Cologne, Germany; e-mail: ana-mihaela.bora@uk-koeln.de

[e] Department of Medical Biometry, Institute for Quality and Efficiency in Health Care, Im Mediapark 8, D-50670 Cologne; e-mail: ralf.bender@iqwig.de

[f] Princess Máxima Center for Pediatric Oncology, Heidelberglaan 25, 3584CS Utrecht, The Netherlands; e-mail: E.C.vanDalen@prinsesmaximacentrum.nl

[g] Research Center for Clinical Neuroimmunology and Neuroscience Basel (RC2NB), University Hospital Basel and University of Basel, Basel, Switzerland. Department of Clinical Research, University Hospital Basel, University of Basel, Basel, Switzerland; Meta-Research Innovation Center at Stanford (METRICS), Stanford University, Stanford, CA, USA; Meta-Research Innovation Center Berlin (METRIC-B), Berlin Institute of Health, Berlin, Germany; e-mail: Lars.Hemkens@usb.ch

[h] Division of Cardiovascular Medicine, John Radcliffe Hospital, Oxford, OX3 9DU, UK; Department of Population Health, London School of Hygiene and Tropical Medicine, London, WC1E 7HT, UK; e-mail: m.trivella@btinternet.com

[i] Evidence-based Medicine, Department I of Internal Medicine, Faculty of Medicine and University Hospital Cologne, University of Cologne, Kerpener Str. 62, 50937 Cologne, Germany; e-mail: ina.monsef@uk-koeln.de

[j] Evidence-based Medicine, Department I of Internal Medicine, Faculty of Medicine and University Hospital Cologne, University of Cologne, Kerpener Str. 62, 50937 Cologne, Germany; e-mail: nina.kreuzberger@uk-koeln.de

[k] Evidence-based Medicine, Department I of Internal Medicine, Faculty of Medicine and University Hospital Cologne, University of Cologne, Kerpener Str. 62, 50937 Cologne, Germany; e-mail: nicole.skoetz@uk-koeln.de

**contributed equally*

Corresponding author:

Marius Goldkuhle
Evidence-based Medicine
Department I of Internal Medicine
University Hospital Cologne
Kerpener Str. 62
50637 Cologne
E-Mail: marius.goldkuhle@uk-koeln.de
Telefone: +49 221 478-62032

Abstract

Objective: Previous findings indicate limited reporting of systematic reviews with meta-analyses of time-to-event (TTE) outcomes. We assessed corresponding available information in trial publications included in such meta-analyses.

Study Design and Setting: We extracted data from all randomized trials in pairwise, hazard ratio (HR)-based meta-analyses of primary outcomes and overall survival of 50 systematic reviews systematically identified from the Cochrane Database and Core Clinical Journals. Data on methods and characteristics relevant for TTE analysis of reviews, trials and outcomes were extracted.

Results: Meta-analyses included 235 trials with 315 trial analyses. Most prominently assessed was overall survival (91%). Definitions (61%), censoring reasons (41%) and follow-up specifications (56%) for trial outcomes were often missing. Available TTE data per trial were most frequently survival curves (83%), log-rank P-values (76%) und HRs (72%). When trial TTE data recalculation was reported, reviews mostly specified HRs or P-values (each 5%). Reviews primarily included intention-to-treat analyses (64%) and analyses not adjusted for covariates (25%). Except for missing outcome data, TTE-relevant trial characteristics, e.g., informative censoring, treatment switching and proportional hazards, were sporadically addressed in trial publications. Reporting limitations in trial publications translate to the review level.

Conclusion: TTE (meta-)analyses, in trial and review publications, need clear reporting standards.

Keywords

Systematic review; Meta-analysis; Randomized trials; Time-to-event outcomes; Survival analysis; Reporting quality

What is new?

Key findings

- We identified variable and often insufficient reporting of time-to-event outcomes and associated methods in publications of randomized trials included in aggregate data meta-analyses of current systematic reviews.
- Limited reporting included critical information such as outcome definitions, methods and trial characteristics relevant for assessing the certainty of time-to-event analyses, e.g. informative censoring and proportional hazards. Available time-to-event data varied substantially between trial publications.
- Limitations in trial reporting translate to review publications as well.

What this adds to what is known

- Previous methodological research suggested shortcomings in the reporting of time-to-event outcomes and analyses in study publications. Focusing on trials included in meta-analyses, we showed that these limitations have relevance for meta-analyses in current systematic reviews.

What are the implications and what should be changed?

- Trial authors should strictly adhere to available reporting guidelines for time-to-event analyses in randomized trial publications. Reporting standards for meta-analyses of time-to-event outcomes based on aggregate data are urgently needed.

1. Introduction

Researchers interested in effects of interventions on longer-term outcomes or outcomes that occur in all participants at some point often employ time-to-event outcomes (1, 2). Time-to-event analyses measure the occurrence of an event, e.g., death, disease progression or wound healing, together with the time until its occurrence and, for individuals without an observed event (censored observation), accounts for their time under observation. Survival plots and probabilities estimated by using the method by Kaplan and Meier (3), hazard ratios (HRs) estimated by using the Cox model and various statistical tests, most prominently the log-rank test, constitute the most frequently used methods for time-to-event analyses (4, 5). Meta-analyses of time-to-event outcomes from aggregate trial data are commonly performed based on the HR, which, for individual trials, can be included directly or derived from various data sources in trial publications (4, 6, 7).

Because time-to-event analyses are complex, authors of evidence syntheses depend on rigorous reporting in trial publications to determine the credibility of their meta-analyses. Trial HRs are frequently estimated by using Cox models which assume at least approximate proportionality of the hazards of compared groups (proportional hazards) over the observation time (5, 8-10). Missing outcome data, competing events and treatment switching impose challenges on interpretation of the results, especially when they lead to naive censoring of trial participants (1, 11-15). Finally, information on more general analytical trial characteristics are particularly relevant for time-to-event outcome meta-analyses and their interpretation. Unfortunately, previous studies have indicated that reporting of trials including time-to-event outcomes is often deficient (16-21).

We explored the characteristics, methodology and handling of time-to-event analyses of trials included in meta-analyses of current systematic reviews.

2. Methods

We report our assessment in accordance with an adaption of the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) checklist to meta-epidemiological research (22). The project is registered under: osf.io/5qxbd.

2.1. Eligibility criteria

We assessed publications of trials that were included in systematic reviews with a meta-analysis based on aggregate data from a minimum of two RCTs that evaluated a health-related time-to-event outcome by means of the HR. We did not impose limitations regarding intervention types, medical fields, or settings, but reviews should have been available as full-text articles published in English. Network meta-analyses, previous versions of updated reviews and co-publications of Cochrane reviews (CR) were excluded.

2.2. Identification and selection of reviews and trials

The reviews were part of a separate study on the handling of time-to-event outcomes in systematic reviews (Goldkuhle et al. 2023, *submitted*). Briefly, we randomly selected 25 CR from a sample of CR until August 2020 and 25 non-Cochrane (nCR) reviews from a corresponding sample published during the same time (28/02/2017 to 18/08/2020). Non-Cochrane reviews were identified in a systematic search performed by an experienced information specialist (IM) (appendix A1; 08/02/2021) and limited to reviews published in Core Clinical Journals, as defined by the U.S National Library of Medicine, to ensure relevance of included reviews (26).

We assessed primary review outcomes or, if not applicable, the first time-to-event outcome reported in the abstract. If a review included overall survival/all-cause mortality in a time-to-event outcome meta-analysis, we included this analysis as well, because it is often considered

the most relevant outcome of a study. For feasibility, we excluded analyses with more than 20 trials.

For the selected review outcomes, we identified all trial publications from which time-to-event data were included in applicable meta-analyses: Systematic reviews often cite multiple publications of individual included trials. In such cases, we prioritized publications and outcome data as reported by review authors. Otherwise, we selected trial publications including a HR and confidence interval that corresponded to a review's forest plot HR, or could be inverted accordingly. If no corresponding trial HR was reported or if it differed from the review reported trial HR in any of the referenced trial publications, and other time-to-event data that were reported could not be directly matched, we selected the publication that corresponded in follow-up duration and number of participants if possible. Where a corresponding trial publication reported multiple sources of time-to-event data and it was unclear which was selected by review authors, we noted this information.

Selection took place in duplicate and independently (NK, MG). Potential discrepancies were resolved by consulting a third author (NS).

2.3. Data extraction and statistical analysis

All extractions were performed in duplicate by two authors (NK, CI, CH, AB, MG) with involvement of a third author (NS) in case of potential discrepancies. The data extraction sheet was developed and piloted a-priori (appendix A2). Data were analyzed descriptively by means of absolute and relative frequencies for categorical data and medians, means and variability measures for count data. To illustrate how review authors approached items associated with those extracted on trial level, we present data extracted for reviews in the tables along applicable trial level results.

3. Results

3.1. Search results for reviews and their included trials

A flow diagram (appendix A3) illustrates our search.

The identified reviews included 235 trials in their primary and overall survival time-to-event outcome analyses, resulting in 315 individual trial analyses of time-to-event outcomes included in review meta-analyses. For outcomes of 18 trials, we did not extract data because either time-to-event data were not available in cited publications, or it was unclear which data were included in the review, or publications were not accessible, or data were received from a secondary source (appendix A4).

3.2. Characteristics of included reviews

Appendix-table A5 presents the characteristics of included reviews in detail. Most reviews were published in 2019, addressed questions on neoplasms, and compared biologics/drugs to biologics/drugs. Reviews included a median of four studies (interquartile range (IQR) 2.25–5) and 1521 participants (571–4580.5) in time-to-event outcome meta-analyses. They compared a median of five outcomes (IQR 4–8), among them a median of two (IQR 2–2) of time-to-event outcomes.

3.3. Characteristics of trials included in review time-to-event outcome meta-analyses

Domain		Trial			Review		
		Overall (N = 235)	**Cochrane (n = 102)**	**Non-Cochrane (n = 133)**	**Overall (N = 50)**	**Cochrane (n = 25)**	**Non-Cochrane (n = 25)**
Publication							
Publication year	≤2000	9% (19)	18% (18)	1% (1)	24% (12)	44% (11)	0% (1)
	2001-2005	14% (32)	15% (15)	13% (17)	34% (17)	36% (9)	32% (8)
	2006-2010	20% (46)	25% (26)	15% (20)	48% (24)	60% (15)	36% (9)
	2011-2015	31% (74)	25% (25)	37% (49)	64% (32)	52% (13)	76% (19)
	2016-2020	27% (64)	18% (18)	35% (46)	54% (27)	36% (9)	72% (18)
Publication format	First full publication/NOS	84% (197)	76% (78)	89% (119)	100% (50)	100% (25)	100% (25)
	Updated analysis	9% (20)	11% (11)	7% (9)	22% (11)	20% (5)	24% (6)
	Abstract	4% (10)	8% (8)	2% (2)	14% (7)	20% (5)	8% (2)
	Other (e.g. final analysis, letter)	3% (8)	5% (5)	3% (3)	16% (8)	20% (5)	12% (3)
Trial design	Superiority/NOS	87% (204)	83% (85)	89% (119)	96% (48)	92% (23)	100% (25)
	Non-inferiority	11% (27)	13% (13)	11% (14)	26% (13)	28% (7)	24% (6)
	Equivalency	1% (3)	3% (3)	0% (0)	4% (2)	8% (2)	0% (0)
	Other (i.e. equivalency, combined analysis)	2% (4)	4% (4)	0% (0)	6% (3)	12% (3)	0% (0)
Data availability							
Multiple references		31% (73)	61% (62)	8% (11)	60% (30)	76% (19)	44% (11)
*Data in primary publication**	Yes	29% (67)	55% (56)	8% (11)	58% (29)	88% (22)	28% (7)
	No	3% (8)	8% (8)	0% (0)	14% (7)	28% (7)	0% (0)
	No primary publication defined by review authors	7% (17)	4% (4)	10% (13)	22% (11)	8% (2)	36% (9)
	Single publication referenced	63% (147)	33% (34)	85% (113)	76% (38)	56% (14)	96% (24)
Origin of TTE data clear[#]*	Review HR is trial HR	61% (143)	43% (44)	74% (99)	80% (40)	64% (16)	96% (24)
	Reported by review authors	16% (38)	23% (23)	11% (15)	20% (10)	28% (7)	12% (3)
	Single data source in cited publication(s)	15% (35)	21% (21)	11% (14)	28% (14)	28% (7)	28% (7)
	HR recalculated but source not reported	12% (28)	18% (18)	8% (10)	40% (20)	56% (14)	24% (6)
Trial population							
Sample size of randomized population	Median (IQR)	266 (120-620)	219 (108 – 605)	310 (149 – 627)	1531 (499 – 3318)	593 (358 – 1692)	1935 (1473 – 3766)
	Mean (range)	663 (20 – 17160)	602 (20 – 8113)	707 (40 – 17160)	2946 (83 – 31703)	2216 (83 – 10988)	3676 (349 – 31703)
	Not reported	6% (13)	10% (10)	2% (3)	18% (9)	24% (6)	12% (3)
Proportion of randomized participants not in analysis (%)[§]	Median (IQR)	2.3 (0.8 – 7.5)	3.7 (1 – 10.9)	1.7 (0.5 – 4.8)			
	Mean (range)	9.1 (0 – 63.3)	7.4 (0 – 60.9)	10.7 (0 – 63.3)			
	Unclear/ Not reported	10% (31)	14% (18)	7% (13)	36% (18)	32% (8)	40% (10)
	All randomized analyzed	55% (174)	44% (58)	63% (116)	94% (47)	92% (23)	96% (24)
Outcomes in trial publication							
Number of TTE event outcomes	Median (IQR)	2 (2 – 3)	2 (2 – 3)	2 (2 – 3)			
	Mean (range)	3.27 (1 – 49)	2.22 (1 – 6)	4.07 (1 – 49)			
	Not reported/ Unclear	1% (3)	2% (2)	1% (1)			
Assessed TTE outcomes	ACM/ OS	91% (214)	87% (89)	94% (125)			
	Progression-free survival	37% (88)	16% (16)	54% (72)			
	Disease-free survival	19% (44)	31% (32)	9% (12)			
	Duration of response	8% (18)	0% (0)	14% (18)			
	Time to progression	7% (16)	7% (7)	7% (9)			
	Other[$]	349	82	267			
Safety data as TTE data		3% (8)	0% (0)	6% (8)	6% (3)	0% (0)	12% (3)

Abbreviations: ACM = all-cause mortality; HR = Hazard ratio; IQR = interquartile range; NOS = Not otherwise specified; OS = Overall survival; TTE = time-to-event

Table 1: Characteristics of trials included in the reviews time-to-event outcome meta-analyses. (These data must be interpreted as "trials including at least one outcome fulfilling the respective item (e.g., 29% of trials included at least one trial outcome for which data was available in the primary trial publication, as indicated by the review authors; [#] refers to whether the origin of time-to-event data for an extracted trial was completely clear and, if so, how. The origin was clear if the forest plot HR and confidence interval in a review publication for an individual trial outcome corresponded to a HR and confidence interval reported for that outcome in a respective trial publication, if the review authors explicitly reported the source of time-to-event data for that trial outcome (e.g., in case of data recalculation) or if only a single source of time-to-event data for a trial outcome was available in any trial publication cited in a review; [§] These data are presented per trial outcome: N=315 (CR: n=131; nCR: n=184); [$] Other included, e.g., cardiovascular death, event-free survival, relapse/recurrence-free survival and myocardial infarction)*

Most trials were published between 2011 and 2015 (table 1). Time-to-event data in reviews were predominately available in first full-text publications of trials and from trials addressing superiority. When multiple publications were cited and primary publications defined by the review authors (e.g., by an asterisk in the list of references for individual trials in Cochrane reviews), we most often located applicable time-to-event data in these. Overall, the original publications of time-to-event data were completely clear for 89% (279/315 trial outcomes) of trial outcomes in that either the trial HRs in reviews corresponded to those in trial publications, or the source was explicitly reported by review authors, or only single publications were referenced.

The median population randomized per trial was 266 (IQR 120-620). For 44% (141/315 trial outcomes) of all trial outcomes, the analyzed population differed from the randomized population. If reported, the analyzed population differed by a median of 2.3% (IQR 0.8%-7.5%) and up to 63.3% of the randomized population. Trials analyzed a median of 2 (IQR 2-3) time-to-event outcomes. Most prominent time-to-event outcome per trial was overall survival or all-cause mortality (91%; 214/235 trials). Few trials assessed safety data with time-to-event methods.

3.4. Characteristics of trial outcomes included in this assessment

Domain		Trial outcome			Review			Handling in review
		Overall (N = 315)	Cochrane (n = 131)	Non-Cochrane (n = 184)	Overall (N = 50)	Cochrane (n = 25)	Non-Cochrane (n = 25)	
Trial outcomes included in this assessment (Primary and overall survival/ all-cause mortality review outcomes*)*	ACM/ OS	64% (201)	67% (88)	61% (113)	88% (44)	84% (21)	92% (23)	
	Progression-free survival	17% (52)	8% (11)	22% (41)	20% (10)	12% (3)	28% (7)	
	Disease-free survival	6% (20)	13% (17)	2% (3)	8% (4)	12% (3)	4% (1)	
	Local control	3% (10)	4% (5)	3% (5)	4% (2)	4% (1)	4% (1)	
	Stent failure	3% (8)	0% (0)	4% (8)	2% (1)	0% (0)	4% (1)	
	Other*	6% (20)	7% (10)	5% (10)	24% (12)	32% (8)	16% (4)	
Primary trial outcome		42% (132)	37% (48)	46% (84)	76% (38)	76% (19)	76% (19)	
Outcome definition provided		61% (192)	59% (77)	63% (115)				Heterogenous definitions mentioned - 6% (3/50) in discussion - 2% (1/50) in results
Composite outcome	Yes	26% (83)	19% (25)	32% (58)				
	No	70% (221)	76% (99)	66% (122)				
	Unclear	3% (11)	5% (7)	2% (4)				
Outcome composites defined		92% (76)	92% (23)	91% (53)				
Outcome composites consistent	Yes	42% (132)	52% (68)	35% (64)	62% (31)	64% (16)	60% (15)	
	No	5% (15)	5% (6)	5% (9)	16% (8)	16% (4)	16% (4)	
	Unclear	1% (3)	2% (3)	0% (0)	6% (3)	12% (3)	0% (0)	
	Not applicable	52% (165)	41% (54)	60% (111)	90% (45)	80% (20)	100% (25)	
Death as competing event possible	Yes	11% (35)	11% (14)	11% (21)	22% (11)	20% (5)	24% (6)	
	No	84% (264)	82% (108)	85% (156)	92% (46)	88% (22)	96% (24)	
	Unclear	5% (16)	7% (9)	4% (7)	26% (13)	28% (7)	24% (6)	
Reasons for censoring provided	Yes	41% (130)	33% (43)	47% (87)	74% (37)	68% (17)	80% (20)	
	Unclear/ Not reported	59% (185)	67% (88)	53% (97)	96% (48)	96% (24)	96% (24)	
Reasons for censoring	Participant last known event-free	23% (72)	16% (21)	28% (51)	44% (22)	36% (9)	52% (13)	
	End of follow-up	15% (47)	14% (18)	16% (29)	50% (25)	40% (10)	60% (15)	
	Loss-to-follow up	9% (28)	8% (11)	9% (17)	26% (13)	24% (6)	28% (7)	
	Other#	6% (18)	2% (3)	8% (15)	28% (1)	12% (3)	40% (10)	
	Unclear	1% (2)	0% (0)	1% (2)	2% (1)	0% (0)	4% (1)	
Follow-up start reported	Yes	56% (175)	53% (70)	57% (105)	80% (40)	68% (17)	92% (23)	
	No	34% (108)	35% (46)	34% (62)	82% (41)	72% (18)	92% (23)	
	Unclear	0% (1)	1% (1)	0% (0)	2% (1)	4% (1)	0% (0)	
	Not applicable	10% (31)	11% (14)	9% (17)	40% (20)	40% (10)	40% (10)	
Follow-up start	Randomization	43% (135)	42% (55)	43% (80)	72% (36)	60% (15)	84% (21)	Follow-up start included in any outcome definition - 38% (19/50) Randomization - 4% (2/50) Allocated treatment - 2% (1/50) Enrollment
	Allocated treatment	7% (22)	5% (6)	9% (16)	14% (7)	4% (1)	24% (6)	
	Other (e.g. enrollment, previous treatment)	6% (18)	8% (10)	5% (9)	24% (12)	24% (6)	24% (6)	
	Not applicable	44% (139)	46% (60)	43% (79)	88% (44)	80% (20)	96% (24)	

Abbreviations: ACM = all-cause mortality; IQR = Interquartile range; MACE = Major adverse cardiac events; MI = Myocardial infarction; OS = overall survival; TIMI = Thrombolysis in myocardial infarction

Table 2: Characteristics of time-to-event outcomes defined and analyzed in the included trials. (Other included event-free survival, time to wound healing, event-free survival, major adverse cardiac events (MACE), thrombolysis in myocardial infarction (TIMI) major bleeding, "composite of all-cause death, myocardial infarction or stroke", time to recurrence, biochemical relapse-free survival and time to death from prostate cancer; # Other included alternative treatment, competing events, absence of post-baseline information, participant withdrawal or withdrawal of consent, and inadequate outcome assessment)*

Outcome definitions were provided for 61% (192/315 trial outcomes) assessed trial outcomes (table 2). Death as a competing event was possible in 11% (35/315 trial outcomes) of trial outcomes. Planned reasons for censoring of study participants were reported for less than half of trial outcomes. Reasons were most frequently last known time-points of individuals being event-free and end of follow-up. Loss-to-follow-up, alternative treatments and competing events, were less often reported. Finally, a follow-up starting point was given for 56% (175/315 trial outcomes) of trial outcomes, which was most frequently randomization.

3.5. Time-to-event methodological characteristics of the trials included in review time-to-event outcome meta-analyses

Domain		Trial outcome: Overall (N = 315)	Trial outcome: ACM/ OS (n = 198)	Trial outcome: Combined, including ACM (n = 77)	Trial outcome: Not including ACM (n = 40)	Review: Overall (N = 50)	Review: Cochrane (n = 25)	Review: Non-Cochrane (n = 25)	Handling in review
Time-to-event data available for trial outcomes in trial publications									
Time-to-event data	Survival curves	83% (263)	85% (168)	90% (69)	65% (26)	92% (46)	84% (21)	100% (25)	Five most frequent methods for TTE data - 62% (31/50) HR and CI - 42% (21/50) Set of methods (e.g., Tierney 2008 (4)) - 22% (11/50) Survival curves - 20% (10/50) log(HR) and standard error - 8% (4/50) HR and other information
	P-value (log-rank)	76% (240)	75% (148)	87% (67)	63% (25)	94% (47)	92% (23)	96% (24)	
	HR or log(HR)	72% (226)	68% (135)	95% (73)	45% (18)	90% (45)	80% (20)	100% (25)	
	Time-point specific survival (per arm)	46% (145)	48% (95)	49% (38)	30% (12)	82% (41)	76% (19)	88% (22)	
	Median survival (per arm)	40% (125)	39% (78)	51% (39)	20% (8)	58% (29)	56% (14)	60% (15)	
	Type of test unclear or not reported	6% (20)	6% (12)	5% (4)	10% (4)	26% (13)	20% (5)	32% (8)	
	Other*	10% (33)	11% (22)	4% (3)	18% (8)	46% (23)	60% (15)	32% (8)	
HR calculation	Cox model	60% (188)	57% (113)	75% (58)	43% (17)	86% (43)	72% (18)	100% (25)	HR included in meta-analyses - 88% (44/50) HR/ log(HR) NOS - 6% (3/50) Other (HR/ log HR from Cox model and HR/ log(HR) from Cox model, log-rank test and Kaplan Meier curve)
	Other[#]	3% (9)	3% (6)	1% (1)	5% (2)	14% (7)	16% (4)	12% (3)	
	Unclear/ Not reported	11% (36)	10% (20)	19% (15)	2% (1)	42% (21)	40% (10)	44% (11)	
	No HR calculated	26% (82)	30% (59)	4% (3)	50% (20)	54% (27)	64% (16)	44% (11)	
Survival plots for trial outcomes in trial publications									
Survival plots	Kaplan-Meier	79% (249)	81% (161)	88% (68)	50% (20)	92% (46)	84% (21)	100% (25)	
	Other[§]	4% (14)	3% (6)	1% (1)	16% (7)	14% (7)	16% (4)	12% (3)	
	No, no graphs were presented	17% (52)	16% (31)	10% (8)	33% (13)	60% (30)	64% (16)	56% (14)	
Number at risk reported	Yes	58% (184)	55% (108)	78% (60)	40% (16)	88% (44)	76% (19)	100% (25)	
	No	27% (86)	33% (65)	13% (10)	25% (11)	58% (29)	64% (16)	52% (13)	
	Not applicable	14% (45)	13% (25)	9% (7)	33% (13)	56% (28)	56% (14)	56% (14)	
Censoring reported	Marked on plot	38% (119)	37% (74)	49% (38)	18% (7)	68% (34)	64% (16)	72% (18)	Handling of non-administrative censoring - 2% (1/50) Mentioned as bias criterion
	On plot and with individuals at risk	3% (11)	3% (5)	8% (6)	0% (0)	14% (7)	12% (3)	16% (4)	
	No	43% (136)	45% (90)	34% (26)	50% (20)	80% (40)	80% (20)	80% (20)	
	Not applicable	16% (49)	15% (29)	9% (7)	33% (13)	62% (31)	68% (17)	56% (14)	
Censoring balanced	Yes	30% (96)	31% (61)	40% (31)	10% (4)	66% (33)	64% (16)	68% (17)	
	No	8% (24)	6% (12)	14% (11)	3% (1)	28% (14)	20% (5)	36% (9)	
	Unclear	3% (9)	3% (5)	3% (2)	5% (2)	14% (7)	4% (1)	24% (6)	
	Not applicable	59% (186)	61% (121)	43% (33)	80% (32)	88% (44)	92% (23)	84% (21)	
Data recalculation from trials reported in reviews for an individual trial outcome									
Data recalculation	HR and other information (e.g., events)	5% (15)	4% (8)	1% (1)	15% (6)	4% (2)	4% (1)	4% (1)	
	P-value and other information (e.g., events)	5% (15)	6% (12)	3% (2)	3% (1)	8% (4)	12% (3)	4% (1)	
	Other[$]	8% (25)	8% (16)	8% (6)	7% (3)	20% (10)	28% (7)	12% (3)	
	Not reported	83% (260)	82% (162)	88% (68)	75% (30)	86% (43)	76% (19)	96% (24)	

Abbreviations: AAR = Absolute risk reduction; ACM = All-cause mortality; CI = Confidence interval; HR = Hazard ratio; O-E = Observed – expected; OS = overall survival; NOS = Not otherwise specified; RMST = Restricted mean survival time; RPSFT = Rank Preserving Structural Failure Time

Table 3: Time-to-event specific methodological characteristics of trials included in the reviews time-to-event outcome meta-analyses. (Other includes median cumulative incidence (per arm), mean and standard deviation per arm, O-E events (log-rank) or hazard rates, or Wilcoxon-Gehan test; [#] Other includes HR calculated from log rank tests, HR from Cox and RPSFT models, HR from Cox and time-dependent Cox models, Cox Markov model, and Cox and Fine and Gray models; [§] Other includes cumulative incidence curves, adjusted Kaplan-Meier curves and unclear type of curves); [$] Other includes HR and confidence intervals, individual participant data (recalculated or from publication), survival curves, and time-point specific survival times)*

The most frequently available time-to-event results (appendix-figure A6), for individual trial outcomes were HRs or log(HR)s, log-rank P-values and survival curves, in combination with either time-point specific survival probabilities, median survival times or both. Differences in available time-to-event data types existed between outcomes of overall survival/ all-cause mortality, composite outcomes including death from any cause and outcomes not including death from any cause (appendix A7). Other data such as cumulative incidence rates and O-E events were given rarely. When reported, HRs were primarily calculated with Cox models and sporadically from log-rank results or, for example, Rank Preserving Structural Failure Time (RPSFT) or Fine and Gray models (table 3; appendix A7).

The included reviews only scarcely reported the utilized sources of time-to-event data for an individual trial outcome, if done, most often it was recalculation from HRs or P-values, with information such as events per trial arm.

For 79% (249/315 trial outcomes) of outcomes, trials provided Kaplan-Meier curves, occasionally with the censored individuals throughout follow-up and the individuals at risk over time. Sporadically reported were cumulative incidence curves and adjusted Kaplan-Meier curves. If assessable, we perceived censoring as balanced, regarding distribution over time and proportions, in 80% (96/120 trial outcomes) of applicable trial outcomes.

3.6. General methodological characteristics of the trials included in review time-to-event outcome meta-analyses

Domain		Trial outcome: Overall (N = 315)	Trial outcome: ACM/ OS (n = 198)	Trial outcome: Combined, including ACM (n = 77)	Trial outcome: Not including ACM (n = 40)	Review: Overall (N = 50)	Review: Cochrane (n = 25)	Review: Non-Cochrane (n = 25)	Handling in review
Trial outcome analyses available in trial publications									
Available analyses	ITT	70% (220)	68% (135)	79% (61)	60% (24)	96% (48)	96% (24)	96% (24)	
	Per protocol	8% (25)	8% (16)	8% (6)	8% (3)	40% (20)	48% (12)	32% (8)	
	mITT	5% (15)	5% (9)	8% (6)	3% (1)	20% (10)	8% (2)	32% (8)	
	As treated	2% (7)	2% (3)	3% (2)	5% (2)	8% (4)	0% (0)	16% (4)	
	Unclear/ Not reported	23% (73)	25% (50)	13% (10)	33% (13)	60% (30)	72% (18)	48% (12)	
Trial outcome analyses included in review meta-analyses									Eligible analyses reported - 42% (21/50) ITT Included analyses reported - 20% (10/50) ITT - 6% (3/50) "Analysis not reported in trials" - 18% (9/50) "Not reported for all trials"
Included analysis	ITT	69% (216)	67% (133)	79% (61)	55% (22)	96% (48)	96% (24)	96% (24)	
	mITT	5% (16)	5% (9)	6% (5)	5% (2)	18% (9)	8% (2)	28% (7)	
	Other (e.g. per protocol, as treated)	3% (8)	2% (5)	1% (1)	5% (2)	30% (15)	28% (7)	32% (8)	
	Unclear/ Not reported	25% (78)	27% (53)	13% (10)	34% (15)	62% (31)	72% (18)	52% (13)	
Analysis in complete population	Yes	55% (174)	56% (110)	55% (42)	55% (22)	96% (48)	96% (24)	96% (24)	
	No	32% (100)	31% (62)	31% (24)	35% (14)	70% (35)	80% (20)	60% (15)	
	Unclear/ Not reported	9% (27)	9% (17)	9% (7)	7% (3)	14% (7)	28% (7)	40% (10)	
	Not applicable (e.g. subgroups only)	4% (14)	5% (9)	5% (4)	3% (1)	14% (7)	8% (2)	20% (5)	
Analysis in allocated arm	Yes	88% (276)	88% (175)	91% (70)	78% (31)	98% (49)	96% (24)	100% (25)	
	No	0% (1)	0% (0)	0% (0)	3% (1)	2% (1)	0% (0)	4% (1)	
	Unclear/ Not reported	10% (38)	11% (23)	8% (7)	15% (8)	44% (22)	60% (15)	28% (7)	
Adjusted, unadjusted, stratified analyses available in trial publications									
	Adjusted	27% (86)	25% (50)	36% (28)	20% (8)	62% (31)	56% (14)	68% (17)	Eligible covariate adjustments - 4% (2/50) Adjusted only - 4% (2/50) Hierarchical (adjusted before unadjusted) - 4% (2/50) Unadjusted only - 2% (1/50) Both - 2% (1/50) Hierarchical (unadjusted before adjusted) - 6% (3/50) Unclear Stratified HRs eligible: 2% (1/50) Handling differently adjusted HRs - 2% (1/50) Only adjusted included, others likely excluded - 2% (1/50) Unadjusted recalculated - 8% (4/50) Unclear Included adjustment mentioned - 2% (1/50) In results
	Stratified	20% (62)	18% (36)	29% (22)	10% (4)	36% (18)	24% (6)	48% (12)	
	Unadjusted	17% (54)	17% (34)	17% (13)	18% (7)	54% (27)	48% (12)	60% (15)	
	Unclear/ Not reported	22% (70)	21% (42)	34% (26)	5% (2)	54% (27)	40% (10)	68% (17)	
	Not applicable (No HR reported)	28% (87)	31% (62)	5% (4)	53% (21)	54% (27)	64% (16)	44% (11)	
Adjusted, unadjusted, stratified analyses included in review meta-analyses									
Covariate adjustment of included analysis	Unadjusted	25% (80)	28% (56)	16% (12)	30% (12)	60% (30)	68% (17)	52% (13)	
	Stratified	18% (56)	16% (31)	27% (21)	10% (4)	34% (17)	20% (5)	48% (12)	
	Adjusted	13% (41)	12% (24)	16% (12)	13% (5)	44% (22)	28% (7)	60% (15)	
	Unclear/ Not reported	44% (138)	44% (87)	42% (32)	48% (19)	76% (38)	72% (18)	80% (20)	

Abbreviations: ACM = all-cause mortality; HR = hazard ratio; ITT = intention-to-treat; mITT = modified intention-to-treat; OS = overall survival

Table 4: General methodological characteristics of trials included in the reviews time-to-event outcome meta-analyses.

Analysis types available in trial publications for individual trial outcomes (table 4; appendix A8), were most often intention-to-treat (ITT) analyses alone and most analyses that were reported as ITT analyses were performed in the complete allocated population.

Trial outcome analyses that were included in meta-analyses of the reviews were mostly ITT analyses as referred to by the trial conductors (69%; 216/315 trial outcomes). Overall, more than half of trial outcome analyses that were included in meta-analyses were clearly performed in the complete allocated trial population and in 88% (276/315) of analyses participants were analyzed in their allocated arm.

If adjustment or stratification of trial outcome HRs was reported, the most frequently available combinations in trial publications for individual outcomes were a single HR that was adjusted for baseline characteristics (12%; 39/315 trial outcomes). An available HR was reported as unadjusted in 24% (54/228 trial outcomes) and as adjusted for 38% (86/228 trial outcomes) of trial outcomes. Yet, frequently the adjustment status of available HRs was not reported.

Trial outcome analyses that were included in meta-analyses were mostly unadjusted (45%; 80/177 trial outcomes), 23% were adjusted (41/177 trial outcomes). For 44% (138/315 trial outcomes) the adjustment status could not be determined.

3.7. Trial results and results included in review time-to-event outcome meta-analyses

The relative effect of HRs from trials included in time-to-event outcome meta-analyses, as reported for example in forest plots and including recalculated HRs, was predominately favoring the intervention that was indicated by the review authors (appendix A9). As judged by 95% confidence intervals, 26% (81/315 trial outcomes) of trial analyses were statistically significantly favoring the review authors' defined intervention. Results differed between CR and nCR and between outcomes of overall survival/ all-cause mortality, composite outcomes including death from any cause and outcomes not including death from any cause (appendix

A9). Hazard ratios which were directly reported in trial publications by trial authors showed a similar distribution.

Where a HR was directly reported in a trial publication, a HR <1 most often indicated a decreased risk of the event in the intervention group (86%; 179/208 trial outcomes) and it was predominantly calculated based on the rate of events in each group (91%; 190/208 trial outcomes) in difference to the rate of participants not experiencing the event (absence of event).

Hazard ratios reported in the trial publications were directly applicable to trial HRs in meta-analyses for 51% (120/315 trial outcomes) of trial outcomes or had to be inverted in 7% (23/315 trial outcomes). In several cases an available HR or its confidence interval differed from the HR in the meta-analysis, e.g. reviews explicitly reported not to use a trial HR, or recalculated the confidence interval.

3.8. Specific trial characteristics with relevance for time-to-event analyses and interpretation

Domain		Trial (N = 235)	Review (N = 50)	Handling in review
Follow-up in trials				
Follow-up measure available	Follow-up reported across outcomes	79% (185)	96% (48)	Foreseen follow-up time reported - 8% (4/50) Longest follow-up, - 6% (3/50) Minimum duration of follow-up required - 4% (2/50) Maximum duration of follow-up specified Handling varying follow-up reported - 10% (5/50) Sensitivity analyses (e.g. shorter/longer follow-up - 12% (6/50) Other (e.g. meta-regression, study exclusion, risk of bias) - 2% (1/50) Unclear Varying follow-up mentioned - 24% (12/50) In discussion - 8% (4/50) In results, - 6% (3/50) In results and in discussion
	Follow-up reported for outcomes	3% (6)	10% (5)	
	No follow-up measure reported	19% (44)	44% (22)	
Available follow-up measures	Median	66% (154)	92% (46)	
	Minimum	25% (59)	56% (28)	
	Maximum	23% (53)	54% (27)	
	IQR/ lower and upper range of IQR	18% (43)	56% (28)	
	Other, e.g. mean, fixed time-point, standard deviation	12% (29)	28% (14)	
Follow-up calculation	Median, surviving patients only	8% (19)	26% (13)	
	Median, all patients	5% (11)	16% (8)	
	Other*	5% (12)	18% (9)	
	Unclear/ Not reported	62% (146)	86% (43)	
	Not applicable	20% (47)	46% (23)	
Reported missing outcome data in trials				
Reported per arm		57% (134)	98% (49)	Handling missing data reported - 68% (34/ 50) Mentioned as risk of bias criterion in methods - 40% (20/50) Contact with authors - 8% (4/50) Sensitivity analyses (according to rate missing) - 4% (2/50) Single imputation Missing data mentioned - 56% (28/50) In results - 8% (4/50) In results and discussion
Reported per outcome	Yes	4% (9)	12% (6)	
	Complete/ no loss at trial level	11% (26)	28% (14)	
	Complete/ no loss at outcome level	3% (8)	12% (6)	
	No	84% (198)	100% (50)	
Handling	Excluded from analysis	18% (42)	42% (21)	
	Censored	11% (26)	34% (17)	
	Complete/ no loss at trial level	11% (25)	28% (14)	
	Single or multiple imputation	1% (2)	4% (2)	
	Unclear/ Not reported	59% (139)	92% (46)	
	No missing data	3% (8)	12% (6)	
Censoring in trials				
Handling	Sensitivity analysis (results not shown)	0% (1)	2% (1)	
Death as competing event in trials				
Handling reported	Yes[#]	3% (7)	10% (5)	Handling of deaths as competing events not reported nor discussed No outcomes with death as competing event assessed: 66% (33/50) of reviews
	No	25% (59)	54% (27)	
	Not applicable	86% (202)	92% (46)	
Treatment switching in trials				
Pre-specified	Reported as not planned or allowed	4% (10)	10% (5)	Handling treatment switching reported - 2% (1/50) Mentioned as risk of bias criterion in methods - 2% (1/50) Presence reported for each trial - 2% (1/50) Sensitivity analysis (e.g. according to rate) Treatment switching mentioned - 6% (3/50) In results - 4% (2/50) In discussion
	Reported as anticipated, e.g. protocol, sample size	3% (8)	12% (6)	
	Unclear/ Not reported	93% (218)	94% (47)	
	Not applicable	0% (1)	2% (1)	
Switching reasons	Course of disease (e.g. disease progression)	12% (29)	32% (16)	
	Participant (e.g. choose to switch)	9% (20)	20% (10)	
	Other[§]	11% (26)	28% (14)	
	Not reported	13% (30)	38% (19)	
	Not applicable	64% (151)	88% (44)	
Handling reported	Yes[$]	1% (3)	2% (1)	
	No	93% (219)	98% (49)	
	Not applicable	8% (19)	18% (9)	
Proportional hazards				

Assumption tested	Test (e.g., log-log, Schoenfeld residuals)	8% (19)	32% (16)	
	Visual inspection of curves	1% (2)	4% (2)	
	No	52% (124)	88% (44)	
	Not applicable (e.g., no HR)	29% (69)	52% (26)	Proportional hazards assessment not reported
Test results	Non-proportional	1% (3)	6% (3)	Handling non-proportional hazards not reported
	Reasonably proportional	1% (2)	4% (2)	
	Not reported	6% (16)	26% (13)	
	Not applicable	92% (216)	100% (50)	
Abbreviations: CI = confidence interval; IQR = Interquartile range; LTFU = Loss to follow-up; RPSFT = Rank preserving structural failure time				

Table 5: Handling of specific trial characteristics with relevance for time-to-event outcomes in the trials included in the reviews time-to-event outcome meta-analyses. (Other included, e.g., the reverse Kaplan-Meier method, median follow-up excluding censored individuals and mean follow-up; # Handling of included cumulative incidence curves alone or together with a Fine and Gray model; § Other included, e.g., administrative (e.g. interim analysis), pre-condition (e.g., allergies), intervention related (e.g., adverse events) and investigator/ physician (e.g., physicians decision); $ Handling included, e.g., rank preserving structural failure time models and sensitivity analyses, either treating cross-overs as outcome events or excluding cross-overs).*

A measure of trial follow-up was available for 82% (191/235 trials) of trials (table 5; appendix A10). Respective measures were most frequently reported as single measure across trial outcomes and only seldomly reported specifically for an individual trial outcome. Follow-up was predominately reported as median follow-up across outcomes and, although seldomly reported, calculated as median survival including surviving/event-free individuals only.

Missing outcome data was reported per trial arm for 57% (134/235 trials) of trials. About a third reported no information at all. The remaining either reported information across arms or for individual outcomes. If reported, median missing outcome data per trial arm was most frequently below 5% of the allocated population, in several cases, however, also substantially higher (appendix A10). Outcome specific missing outcome data was reported in few trials. Handling of missing outcome data consisted most frequently of entirely excluding or censoring respective individuals from the analysis. Regarding handling of potential informative censoring, one trial reported a sensitivity analysis, but did not show any results.

If death was as a potential competing event for an assessed trial outcome, only few trials reported the number of these potential competing events per arm (proportions in appendix A10). In response, trial authors presented survival time distributions as cumulative incidence curves in 21% (7/33 trials) of applicable trials. In two of these trials, authors used Fine and Gray regression to calculate HRs as well.

About a third of trials reported information regarding receipt of the comparator treatment in the intervention group or vice versa (treatment switching). Rates were most frequently below 10% of the allocated population, in some cases, however, they exceeded 20% and 50% (appendix A10). Six trials reported treatment switching as trial protocol specified, otherwise as anticipated, e.g., sample size calculations, or explicitly excluded the option. If treatment switching was reported, the most prominent reason was related to the course of disease, e.g.

disease progression. Additional analyses to deal with treatment switching were reported only in single trials.

According to trial reporting, proportionality of hazards for outcomes analyzed as HRs was assessed by statistical tests, e.g., log-log or Schoenfeld residuals, or by visual inspection of survival plots in 11% (19/166 trials) and 1% (2/166 trials) of trials. In only five trials, results of these assessments were reported, thrice as non-proportional and twice as reasonably proportional.

4. Discussion

4.1. Principal findings

The origin of included time-to-event data was determinable though our investigation in almost all reviews, but only rarely due to explicit reporting by review authors. Overall survival was the most commonly used time-to-event outcome in trials and reviews. Only around half of trials provided definitions for their assessed outcomes and few gave reasons for censoring. Available time-to-event summary data for individual trial outcomes consisted most frequently of a combination of HRs, log rank P-values, survival curves and either median or time-point specific survival times. Yet, data utilized for recalculation of summary data in reviews were only seldomly reported for individual trial outcomes.

Analyses included in reviews most frequently used individuals in their allocated trial arms, but only little over half were clearly performed in the complete allocated trial populations. Trial effect measures included in reviews were mostly unadjusted for covariates, and information on adjustment of available HRs was often not reported in trial publications.

Numerical missing outcome data was available per trial arm for little over half of trials and only rarely for individual outcomes. Trial conductors often handled it by excluding or censoring affected participants. For informative censoring, one trial indicated sensitivity analyses.

Treatment switching was reported for about a third of trials, with in some cases considerably high rates and most often due to the participants disease. Proportional hazards were assessed in 10% of trials, but results of such assessments were even more seldomly reported.

4.2. Comparison to review level handling of time-to-event data

Like their included trials, definitions of time-to-event outcomes of interest were provided only in half of reviews. Relevant follow-up information was infrequently defined. Even though reviews included predominately intention-to-treat analyses, eligible and included analysis types as well as details on adjustment of estimates were often not reported. Review methods to obtain time-to-event data varied substantially, most present was direct inclusion of the HR and complete sets of recalculation methods, and were only seldomly reported for an individual outcome. In reviews respectively, trial characteristics relevant to time-to-event analysis (for example variable follow-up, informative censoring, competing events, treatment switching and proportional hazards) were sporadically included in additional assessments (for example sensitivity analyses or certainty assessments) and scarcely mentioned in review texts.

4.3. Strengths and limitations

We ensured robustness of our extraction results through a priori developed forms and duplicate performance of relevant steps. Nevertheless, we must acknowledge potential limitations: first, we used random sampling to generate a representative, but manageable set of reviews. Secondly, we aimed to extend our exploration to reviews often considered the methodological gold-standard and based our sample on a fixed number of CR. We ensured relevance of the included reviews through selecting nCR published in Core Clinical Journals. Third, the limited number of included systematic reviews lead to imbalances between characteristics of CR and nCR. These appear, however, typical for comparisons of both and a comparison was not our primary intent (23, 24). Fourth, we imposed a restriction to primary and all-cause death-

including review outcomes which often constitute a subgroup of outcomes that is reported with greater rigor. Fifth, for feasibility, we limited our assessment to comparisons of 20 trials. But, because the total number of excluded reviews was small (2/74 (3%) CR and 22/401 (5%) nCR during full-text screening), we assume minimal impact.

4.4. Relation to other work

Previous studies support our findings of deficient reporting of time-to-event analysis-relevant information in trial publications, including the start and end points of observations, censoring and follow-up information, assumptions, such as proportional hazards in Cox models, and details on statistical modelling as well as numbers of events and censored observations (16-20, 25). Batson et al. (19) discuss implications of limited trial-level reporting to meta-analysis and particularly promote openness to alternative approaches when assumptions underlying the Cox model HR are in question. Our assessment focusses on trials that are included in time-to-event outcome meta-analyses and confirms their findings. In addition, we show that insufficient trial reporting is also transferred to review publications.

Kahale et al. (26) assessed the handling of missing outcome data in systematic reviews and trials included in meta-analyses of dichotomous data and found that the approach to missing outcome data was explained only in little more than a third of their assessed trials. Determining missing outcome data handling for time-to-event trial outcomes constitutes a particular hardship. Available reporting does not permit the distinction between loss to follow-up censoring and censoring for administrative causes (for example end of follow-up) so that trial participants with potential missing outcome data can be excluded without visibly reducing the analysis sample. We focused our extraction on explicitly reported handling of missing outcome data and assume that in many of the "not reported" cases, lost individuals were naively censored. Kahale and colleagues found that a minor proportion of their assessed reviews

consistently approached missing outcome data in included trials in their analyses, which agrees with our findings.

4.5. Explanations, implications, and further research

Limited reporting in trial publications imposes complications for all who rely on reported information to evaluate the credibility of time-to-event outcome effects from trials, e.g. for meta-analyses. With the recently published CONSORT (Consolidated Standards of Reporting Trials) extension to trial outcomes, in addition to general CONSORT guidance, some of the reporting issues we identified might improve, e.g. appropriate outcome definitions and details on statistical methods, handling of missing outcome data and specification of the analysis population (27). Overall, trial authors should adhere to available reporting guidelines and suggestions, both for general outcome reporting as well as for time-to-event outcome specific information, e.g., for survival curves (17, 18, 27-29).

In response to current reporting limitations on trial level, review authors are encouraged to rigorously follow available guidance, and to explicitly report deficiencies in trial publications they encounter (4, 6, 7). Still, additional guidance and further research on the optimal translation of time-to-event related trial issues to meta-analyses of aggregate data is needed.

5. Conclusions

The poor reporting of time-to-event outcomes and associated methods in trial publications limits not only the usefulness of these trials but also that of the systematic reviews and meta-analyses relying on them.

CRediT authorship contribution statement:

MG: Conceptualization, data curation, formal analysis, investigation, methodology, project administration, writing – original draft, writing - review & editing; CH: Formal analysis, investigation, writing - review & editing; CI: Formal analysis, investigation, writing - review & editing; AMB: Formal analysis, investigation, writing - review & editing; RB: Conceptualization, writing - review & editing; EvD: Conceptualization, writing - review & editing; LGH: Conceptualization, writing - review & editing; IM: systematic search; MT: Conceptualization, writing - review & editing; NK: Conceptualization, data curation, formal analysis, investigation, methodology, writing - review & editing; NS: Conceptualization, methodology, supervision, writing - review & editing

Declarations of interest:

One included non-Cochrane review was co-authored by a project participant (LGH), who did not appraise data or resolve conflicts for these reviews. All other authors declare no relevant conflicts of interest.

Funding:

Funded by the German Research Foundation (DFG) (project SK 146/2-1; https://gepris.dfg.de/gepris/projekt/437030575)

References

1. Leung K-M, Elashoff RM, Afifi AA. CENSORING ISSUES IN SURVIVAL ANALYSIS. Annual Review of Public Health. 1997;18(1):83-104.
2. Lagakos SW. General right censoring and its impact on the analysis of survival data. Biometrics. 1979;35(1):139-56.
3. Kaplan EL, Meier P. Nonparametric Estimation from Incomplete Observations. Journal of the American Statistical Association. 1958;53(282):457-81.
4. 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.
5. Kleinbaum DG, Klein M. Survival Analysis. 3 ed. New York: Springer-Verlag; 2012.
6. Parmar MKB, Torri V, Stewart L. Extracting summary statistics to perform meta-analyses of the published literature for survival endpoints. Statistics in Medicine. 1998;17(24):2815-34.
7. Higgins JPT, Thomas J, Chandler J, Cumpston M, Li T, Page MJ, et al. Cochrane Handbook for Systematic Reviews of Interventions: Cochrane; 2020. Available from: www.training.cochrane.org/handbook.
8. Hernán MA. The hazards of hazard ratios. Epidemiology (Cambridge, Mass). 2010;21(1):13-5.
9. Stensrud MJ, Hernán MA. Why Test for Proportional Hazards? JAMA. 2020;323(14):1401-2.
10. Rulli E, Ghilotti F, Biagioli E, Porcu L, Marabese M, D'Incalci M, et al. Assessment of proportional hazard assumption in aggregate data: a systematic review on statistical methodology in clinical trials using time-to-event endpoint. British Journal of Cancer. 2018;119(12):1456-63.
11. Austin PC, Fine JP. Accounting for competing risks in randomized controlled trials: a review and recommendations for improvement. Statistics in Medicine. 2017;36(8):1203-9.
12. Schumacher M, Ohneberg K, Beyersmann J. Competing risk bias was common in a prominent medical journal. Journal of Clinical Epidemiology. 2016;80:135-6.
13. Sullivan TR, Latimer NR, Gray J, Sorich MJ, Salter AB, Karnon J. Adjusting for Treatment Switching in Oncology Trials: A Systematic Review and Recommendations for Reporting. Value in Health. 2020;23(3):388-96.
14. Ishak KJ, Proskorovsky I, Korytowsky B, Sandin R, Faivre S, Valle J. Methods for adjusting for bias due to crossover in oncology trials. Pharmacoeconomics. 2014;32(6):533-46.
15. Goldkuhle M, Bender R, Akl EA, van Dalen EC, Nevitt S, Mustafa RA, et al. GRADE Guidelines: 29. Rating the certainty in time-to-event outcomes-Study limitations due to censoring of participants with missing data in intervention studies. Journal of Clinical Epidemiology. 2021;129:126-37.
16. Zhu X, Zhou X, Zhang Y, Sun X, Liu H, Zhang Y. Reporting and methodological quality of survival analysis in articles published in Chinese oncology journals. Medicine (Baltimore). 2017;96(50):e9204.
17. Abraira V, Muriel A, Emparanza JI, Pijoan JI, Royuela A, Plana MN, et al. Reporting quality of survival analyses in medical journals still needs improvement. A minimal requirements proposal. Journal of Clinical Epidemiology. 2013;66(12):1340-6.e5.
18. Altman DG, De Stavola BL, Love SB, Stepniewska KA. Review of survival analyses published in cancer journals. British Journal of Cancer. 1995;72(2):511-8.
19. Batson S, Greenall G, Hudson P. Review of the Reporting of Survival Analyses within Randomised Controlled Trials and the Implications for Meta-Analysis. PLOS ONE. 2016;11(5):e0154870.

20. Mathoulin-Pelissier S, Gourgou-Bourgade S, Bonnetain F, Kramar A. Survival End Point Reporting in Randomized Cancer Clinical Trials: A Review of Major Journals. Journal of Clinical Oncology. 2008;26(22):3721-6.
21. Salika T, Turner RM, Fisher D, Tierney JF, White IR. Implications of analysing time-to-event outcomes as binary in meta-analysis: empirical evidence from the Cochrane Database of Systematic Reviews. BMC Medical Research Methodology. 2022;22(1):73.
22. Murad MH, Wang Z. Guidelines for reporting meta-epidemiological methodology research. BMJ Evidence-based Medicine. 2017;22(4):139-42.
23. Page MJ, Shamseer L, Altman DG, Tetzlaff J, Sampson M, Tricco AC, et al. Epidemiology and Reporting Characteristics of Systematic Reviews of Biomedical Research: A Cross-Sectional Study. PLOS Medicine. 2016;13(5):e1002028.
24. Goldkuhle M, Narayan VM, Weigl A, Dahm P, Skoetz N. A systematic assessment of Cochrane reviews and systematic reviews published in high-impact medical journals related to cancer. BMJ Open. 2018;8(3):e020869.
25. Vervölgyi E, Kromp M, Skipka G, Bender R, Kaiser T. Reporting of loss to follow-up information in randomised controlled trials with time-to-event outcomes: a literature survey. BMC Med Res Methodol. 2011;11:130.
26. Kahale LA, Khamis AM, Diab B, Chang Y, Lopes LC, Agarwal A, et al. Meta-Analyses Proved Inconsistent in How Missing Data Were Handled Across Their Included Primary Trials: A Methodological Survey. Clin Epidemiol. 2020;12:527-35.
27. Butcher NJ, Monsour A, Mew EJ, Chan A-W, Moher D, Mayo-Wilson E, et al. Guidelines for Reporting Outcomes in Trial Reports: The CONSORT-Outcomes 2022 Extension. JAMA. 2022;328(22):2252-64.
28. Morris TP, Jarvis CI, Cragg W, Phillips PPJ, Choodari-Oskooei B, Sydes MR. Proposals on Kaplan-Meier plots in medical research and a survey of stakeholder views: KMunicate. BMJ Open. 2019;9(9):e030215.
29. Schulz KF, Altman DG, Moher D, the CG. CONSORT 2010 Statement: updated guidelines for reporting parallel group randomised trials. BMC Medicine. 2010;8(1):18.

Appendix

Appendix A1: Complete search strategy for non-Cochrane reviews

Medline on February 8th, 2021

#	Searches
1	"time-to-event".tw,kf.
2	"log rank".tw,kf.
3	survival.tw,kf.
4	hazard.tw,kf.
5	Kaplan-meier estimate/
6	kaplan-meier.tw,kf.
7	(method* adj1 (product* or limit*)).tw,kf.
8	(cumulative* adj1 incidence*).tw,kf.
9	outcome expectation.tw,kf.
10	(cox adj2 (model* or proportional*)).tw,kf.
11	proportional hazards models/
12	or/1-11
13	(randomi?ed or placebo or randomly).ab.
14	meta analysis.mp,pt.
15	12 and 13 and 14
16	limit 15 to dt=20170101-20200801

Appendix A2: List of extraction items

For a complete list of extraction items extracted for reviews and their assessed time-to-event outcomes, please see the appendix of Goldkuhle et al. 2023 (*submitted*).

Extraction items for trials

#	Item	Option	Description
1	Review ID		Individual ID (number) of review corresponding to the overall review sheet
2	Trial ID		Individual ID (number) of trial (following alphabetical order from review)
3	Review description of trial		Description of trial in review (if only a reference is used, please choose last name of fist author and publication data of that reference)
4	Other than primary publication	Yes; No; Not applicable (e.g. no primary publication defined)	Do you use another than the primary publication in the review for the assessment on overall trial level (e.g. because the review included time-to-event data is only reported in another publication)?
5	Comment on publications of this trial		Field to comment on the publications of this trial in the review
6	First author		Last name of the first author of the trial publication at hand (primary publication or, if applicable, most recent referenced full text publication including relevant outcome data) Primary publication must include results data that was used in the review (If no result data for eligible outcomes included in publication that was labeled in review as "primary publication", please choose most current full-text publication with utilized result data)
7	Publication year		Date of trial publication at hand (primary publication or, if applicable, most recent referenced full text publication including relevant outcome data) Primary publication must include results data that was used in the review (If no result data for eligible outcomes included in publication that was labeled in review as "primary publication", please choose most current full-text publication with utilized result data)
8	Journal		Full title of the journal
9	Publication format	Journal publication (first full publication or not otherwise reported); Journal publication (updated analysis); Journal publication (final analysis); Abstract (e.g. conference presentation); Registry entry; Clinical trial report; Other	
10	Trial PICO		Please enter the PICO of the trial as complete as possible
11	Trial design	Superiority trial or not otherwise specified; Non-inferiority trial; Equivalency trial; Other	Was this trial designed as a non-inferiority trial, equivalence trial or any other design except superiority? If not explicitly reported choose "no"
12	Experimental treatment type	Biologics/drug; Surgical procedure; Medical devices; Radiotherapy; Behavioral intervention; Exercise intervention; Screening; Other (please specify)	Which type of experimental treatment did the trial participants receive?
13	Control treatment type	Placebo; No treatment; Observation; Usual or best-supportive care; Biologics/drug; Surgical procedure; Medical devices; Radiotherapy; Behavioral intervention; Exercise intervention; Screening; Other (please specify)	Which type of control treatment did the trial participants receive?
14	Type of follow-up	Median; Mean;	Which type of follow-up measure(s) where reported for the overall trial population.

#	Item	Option	Description
		Minimum follow-up; Maximum follow-up; IQR/ lower and upper range of IQR; 95% CI of median; 95% CI of mean; Standard deviation Fixed time-point of outcome measurement only; No indicator of follow-up reported; Follow-up reported for outcomes;	Irrespective of whether in total or per arm
15	Follow-up calculation		How was follow-up time calculated?
16	Overall follow-up reported		Was a measure of duration of follow-up for the entire analyzed population reported in the trial publication?
17	Median overall follow-up		Median overall trial follow-up time (in months) Empty field = "Not reported"
18	Follow-up reported per arm		Was a measure of follow-up for the population analyzed in each of the compared arms reported in the trial publication?
19	Median experimental follow-up		Median follow-up time in experimental group (in months) Empty field = "Not reported"
20	Median control follow-up		Median follow-up time in control group (in months) Empty field = "Not reported"
21	Field for commenting on the PICO or follow-up in the assessed trial		
22	Total number of TTE outcomes		What was the total number of outcomes compared in the trial as time-to-event outcomes? (quantitatively compared outcomes with relative effect measure only) If necessary type "Unclear" "Not reported"
23	List of TTE outcomes		List of outcomes examined as TTE in this trial according to the assessed trial publication
24	Safety data as TTE outcomes	Yes; No; Unclear; Not reported; Not applicable (no safety outcomes reported)	Where any adverse events (safety data) assessed with time-to-event methodology?
25	Experimental randomization ratio		Randomization ratio experimental arm (1:1 ="1"; 2:1 ="2", ...)
26	Control randomization ratio		Randomization ratio control arm (1:1 ="1"; 1:2 ="2", ...)
27	Randomized experimental participants		What is the total number of participants randomized to the experimental group? Empty field = "Not reported"
28	Randomized control participants		What is the total number of participants randomized to the control group? Empty field = "Not reported"
29	Total randomized participants		
30	Any missing outcome data per arm reported		Was any missing outcome data reported for this trial per arm? (Including that there was NO missing data or similar statements!)
31	Missing outcome data per arm reported	Yes; No; Explicitly reported complete follow-up; Reported for individual outcomes; Reported across arms only; Explicitly reported no LTFU	
32	Total missing outcome data in experimental arm		We extract individuals with missing outcome data (MOD) (individuals clearly have MOD based on RCT reporting), this includes but is not limited to: 1.a: Individuals reported as with "Explained/ Unexplained LTFU", "Outcome not assessable", "Data not available", etc. 1.b: All other reasons, if explicitly reported as not followed-up, excluded, withdrawn, explicitly imputed, etc. 1.c: Otherwise clear that outcome data collection (assessment of outcome events) was not possible for individuals for a given reason (CAVE: Censoring in TTE analysis allows to include individuals with MOD for some duration into the trial and thus the "denominator" (e.g. the first number of individuals at risk under a survival curve)) In case a number of individuals who discontinued treatment is reported (e.g. in Lancet flow-diagrams) - Please only extract numbers of participants for which it is clear that they could not contribute outcome data (e.g. reported as lost to follow-up) If necessary type:

#	Item	Option	Description
			"unclear" "not reported"
33	Total missing outcome data in control arm		We extract individuals with missing outcome data (MOD) (individuals clearly have MOD based on RCT reporting), this includes but is not limited to: 1.a: Individuals reported as with "Explained/ Unexplained LTFU", "Outcome not assessable", "Data not available", etc. 1.b: All other reasons, if explicitly reported as not followed-up, excluded, withdrawn, explicitly imputed, etc. 1.c: Otherwise clear that outcome data collection (assessment of outcome events) was not possible for individuals for a given reason (CAVE: Censoring in TTE analysis allows to include individuals with MOD for some duration into the trial and thus the "denominator" (e.g. the first number of individuals at risk under a survival curve)) In case a number of individuals who discontinued treatment is reported (e.g. in Lancet flow-diagrams) - Please only extract numbers of participants for which it is clear that they could not contribute outcome data (e.g. reported as lost to follow-up) If necessary type: "unclear" "not reported"
34	Number: Control treatment in experimental group		How many individuals in the experimental group received the treatment assigned to the control group? Add number If necessary type: "Unclear" "Not reported"
35	Number: Experimental treatment in control group		How many individuals in the control group received the treatment assigned to the experimental group? Add number If necessary type: "Unclear" "Not reported"
36	Comparator treatments protocol specified	Yes, reported as protocol specified; Yes, reported as protocol amendment (after trial start); Otherwise reported as anticipated; Reported as not planned or allowed Other (specify); Unclear; Not reported; Not applicable;	Was it explicitly reported that the reception of comparator treatments was protocol specified? Should be clear from the publication (information only reported in the protocol (not in appendix or publication) should not be counted)
37	Comparator treatments in sample size	Yes; Unclear; Not reported; Not applicable	Was it reported that the reception of comparator treatments was included in sample size calculations?
38	Comments on comparator treatments		Comments on the reception of comparator treatments in this trial
39	Reason for comparator treatments	Course of disease - related (e.g. disease progression); Pre-condition related (e.g. too obese, allergies); Intervention related (e.g. adverse events); Participant related (e.g. choose to switch); Administrative (e.g. interim analysis); Other (please specify); Unclear; Not reported; Not applicable	For what reason did participants in the trial arms receive comparator treatments?
40	Field to comment on the number of participants who received comparator treatments		
41	TTE specific RoB items considered	Yes (please specify); No; Unclear (please specify); Not applicable	Did the review authors consider any time-to-event specific or related items in their risk of bias assessment for this trial?
42	Comments on the risk of bias assessment		Comments on the risk of bias assessment of the authors on trial outcome level

Abbreviations: CI = confidence interval; IQR = interquartile range; LTFU = loss-to-follow-up; MOD = missing outcome data; PICO = people-intervention-comparator-outcome; RCT = randomized controlled trial; RoB = risk of bias; TTE = time-to-event

Extraction items for trial outcomes

#	Item	Option	Description
1	Outcome ID		Individual number of this outcome (for a hierarchy, see Excel sheet; e.g. OS = 1, PFS = 2, ...) Refers to the outcome in the review (e.g. if the review authors named their outcome progression-free survival, but included data on relapse-free survival from this trial in the very same meta-analysis, please use the outcome ID for PFS)
2	Trial ID		Individual ID (number) of trial (following alphabetical order from review)
3	Review ID		Individual ID (number) of review corresponding to the overall review sheet (assessable in Excel sheet)
4	Trial outcome		Trial outcome assessed in this column Please shorten: overall survival = OS; progression-free survival = PFS; disease free survival = DFS; etc
5	TTE data used	HR and confidence intervals; HR together with other information (e.g. events in each arm, total events, etc.); Observed and expected events or hazard rates on research and control arm; Observed - expected events together with log-rank V; P-value together with additional information (e.g. events, total events, etc.); Survival curves; Median survival times; Time-point specific survival times; IPD (recalculated or from publication); Reported, but method not clear; Other (specify); Not specified for this trial outcome; Unclear	What type of time-to-event data were used for this outcome from the trial to include it in the review publication (according to the review authors)?
6	Specification of recalculated TTE data		Specification on how review authors recalculated time-to-event data for this outcome from the trial to include it in the review publication
7	HR from review		HR of this trial from the review (e.g. in Forest Plot)
8	Lower 95% CI from review		Lower bound 95% CI of this trial from the review (e.g. in Forest Plot)
9	Upper 95% CI from review		Upper bound 95% CI of this trial from the review (e.g. in Forest Plot)
10	TTE specific RoB assessment?	Yes (please specify); No; Unclear (please specify)	Did the review authors consider any TTE specific trial characteristics in their outcome-specific rob assessment?
11	Comments on the risk of bias assessment		Comments on the risk of bias assessment of the authors on trial outcome level
12	Trial level information: Please make sure that you extract the following data using a publication of this trial that includes the time-to-event data eligible for the review! If this publication is not the primary publication in the review it must be: -referenced in the review -include TTE outcome date applicable for the population and follow-up included in the review You might have to compare the data included in the review with the data in the referenced trial publications before data extraction. If you are unsure which publication to use, please discuss with the extraction team.		
13	Relevant TTE data in primary publication?	Yes; No; Only single publication referenced in review; No primary publication highlighted	Is there time-to-event data available in the primary trial publication that is applicable for the review? Refers to methodological and result data. Available result data should refer to the follow-up time-point in the review If no eligible time-to-event data is available, the extraction for trial_overall and trial_outcome must be performed using the publication with applicable time-to-event data
14	Outcome data in referenced publications?	Yes; No	Result time-to-event (!) data for this outcome (e.g. HR, survival curves or any source used for recalculation in the review) available in the assessed trial publications referenced in the review? If no time-to-event result data for this outcome is available in the trial publications referenced in the review, the extraction STOPS for trial level outcome data
15	Comment when data is not available		Please make a comment when and why data is not available in the primary publication/ publication at hand
16	Was it clear where TTE data for this trial outcome was used from?	Yes, reported by review authors for this trial outcome; Yes, HR corresponds to HR directly available in trial publication (slight deviations e.g. of upper CI due to statistical software should be considered); Yes, because only single source of TTE data in cited publication(s); Unclear, HR was recalculated but source not reported in review; Unclear, because publication where TTE event data for this trial outcome could be reported could not be identified (e.g. among the cited publications);	

#	Item	Option	Description
		No extraction possible, no TTE data in cited publications; No extraction possible, full text or publication where TTE data is reported is not accessible; No extraction possible, data received from secondary source (e.g. contact with authors); No extraction possible, completely unclear which/ whether data was included in review	
17	Primary trial outcome	Yes; No; No primary/ secondary outcomes defined	Was this outcome one of the primary outcomes of the trial?
18	Outcome definition		Please provide the complete outcome definition from the trial publication
19	Composite outcome	Yes; No; Unclear; Not reported	Was the outcome a combined outcome including several events of interests e.g. progression-free survival - progressive disease, overall mortality
20	Composite outcomes described?	Yes; No; Not applicable ("Composite outcome" unclear or not reported)	Were the outcome events composing this composite outcome described?
21	Outcome events consistent with review?	Yes; No; Unclear; Not applicable	Were the outcome events in the definition used in the trial consistent with the outcome definition in the review?
22	Start of outcome assessment reported	Yes; No; Unclear; Not applicable (e.g. outcome not defined)	Was the start time-point of outcome assessment for this outcome reported (e.g. in the outcome definition or in the statistical methods section)?
23	Outcome assessment start	Randomization; Enrollment; Allocated treatment; Previous treatment (e.g. surgery); Other (specify); Not applicable (e.g. start of follow-up not reported); Unclear	What was the defined or otherwise reported start time-point of outcome assessment for this outcome?
24	Competing events possible?	Yes; No; Unclear	Are competing events possible by definition of the outcome? Choose yes, e.g. if overall mortality was not part of the defined outcome
25	Censoring reasons reported?	Yes; Unclear; Not reported	Were any reasons for censoring individuals for this outcome reported? Censoring reasons are sometimes reported together with the definition of the outcome and sometimes in the statistical analysis section.
26	Censoring reasons	Participants last known to be event-free; End of follow-up; Loss to follow-up; Inadequate outcome assessment; Participant withdrawal or consent withdrawal; Alternative treatment; Treatment discontinuation; Other (specify); Unclear; Not applicable (no details on censoring reported for this outcome)	What were the reasons for censoring for this outcome, if reported? Please choose the most applicable.
27	Field for commenting on the outcome definition		Field for commenting on the outcome definition
28	"Which and what kind of time-to-event data was available in the trial publication?" Refers to all time-to-event data in trial publication		
29	Available time-to-event data	HR or log(HR); Observed and expected events (log-rank) or hazard rates; P-value (log-rank); Survival curves; Restricted Mean Survival Time; Median survival times (per arm); Time-point specific survival rates (per arm); Median cumulative incidence (per arm); Time-point specific cumulative incidence (per arm); Greys Test; Wilcoxon-Gehan test; Mean and SD per arm; Other (specify); Type of test unclear or not reported	What types of time-to-event data were available in the assessed trial publications for this outcome (excl. time-point specific or median survival times)?
30	Methods for HRs	Cox model; Fine and Gray; Parametric model (specify); Log-rank; Other (specify); No HR calculated;	If hazard ratios (HR, log(HR), etc.) were available, which methods were used to calculate them?

#	Item	Option	Description
		Unclear; Not reported	
31	Available types of analyses (e.g. ITT, PP)	ITT; Modified ITT; Per-protocol; As treated; Unclear; not reported	Which types of analyses are available (e.g. ITT or PP) in the trial publications for this outcome? Please use description by trial authors (e.g. if an analysis was labeled "ITT" and there were post randomization exclusions still use "ITT" and not "mITT". Whether it was "a real ITT analysis" is assessed in the next item.) Please choose all available analyses in the publication at hand (the type of analysis that was used in the review will be specified in the following)
32	ITT analysis in complete population?	Yes; No; Unclear; Not applicable (no ITT analysis mentioned, e.g. only mITT)	If an analysis according to the ITT principle was described by the authors, was this analysis performed in the complete allocated trial population or were there post-randomization exclusions (e.g. participants did not receive the intended treatment, were mistakenly enrolled, did withdraw consent, died or developed the outcome of interest before treatment)
33	(Un)adjusted/ (un)stratified HRs available?	Unadjusted (univariate including treatment variables only); Adjusted, baseline characteristics; Adjusted, post-baseline exposure; Adjusted, but factors unclear/ not reported; Stratified, but factors unclear; Stratified, randomization stratification factors; Stratified, baseline characteristics; Other (please specify); Unclear; Not reported; Not applicable (no HR directly reported)	Were unadjusted, adjusted and/ or stratified HRs available in the trial publications for this outcome and if adjusted for which factors? Please choose all available HRs in the publication at hand (the type of analysis that was used in the review will be specified in the following)
34	(Un)adjusted/ (un)stratified P-values available?	Unadjusted (univariate including treatment variables only); Adjusted, baseline characteristics; Adjusted, post-baseline exposure; Adjusted, but factors unclear/ not reported; Stratified, but factors unclear; Stratified, randomization stratification factors; Stratified, baseline characteristics; Other (please specify); Unclear; Not reported; Not applicable (no log-rank P-value directly reported)	
35	Field to comment on methods (e.g. specify method to calculate relative effect measures)		
36	Methods - Limited to outcome analysis included in meta-analysis "What was reported in the trial publication for the data included in the meta-analysis?" Refers to the outcome analysis included in the review meta-analysis only		
37	Type of analysis included in meta-analysis	ITT; modified ITT; per-protocol; as treated; unclear; not reported;	Please indicate the analysis producing the estimate (e.g. HR, log-rank results, survival curves) included in the review meta-analysis as labeled by the trial authors. Please use description by trial authors (e.g. if the analysis was labeled "ITT" and there were post randomization exclusions still use "ITT" and not "mITT". Whether it was "a real ITT analysis" is assessed in the next item.)
38	Selected analysis in complete population?	Yes; No; Unclear; Not reported; Not applicable (e.g. only subgroup analysis included in review)	Was this analysis performed in the complete allocated trial population?
39	Population analyzed in allocated arm?	Yes; No; Unclear; Not reported	Were individuals analyzed in the arm they were allocated too for this outcome analysis (except those who were excluded from the sample, e.g. because of mITT)?
40	Pooled estimate unadjusted, adjusted or stratified?	Unadjusted; Adjusted; Stratified; Unclear; Not reported	Was the effect estimate that was pooled in the meta-analysis for this trial an unadjusted or adjusted estimate (applicable to any type of available effect measure, e.g. HR, Observed - expected, log-rank results, survival curves)? Survival curves and median/ time-point specific survival probabilities calculated with Kaplan-Meier or cumulative incidence are expected to be unadjusted - if explicitly reported otherwise, please indicate by comment
41	Survival plots presented?	Yes, Kaplan-Meier; Yes, cumulative incidence; Yes, other (please specify); No, no graphs were presented	Were survival plots presented for the assessed analysis?

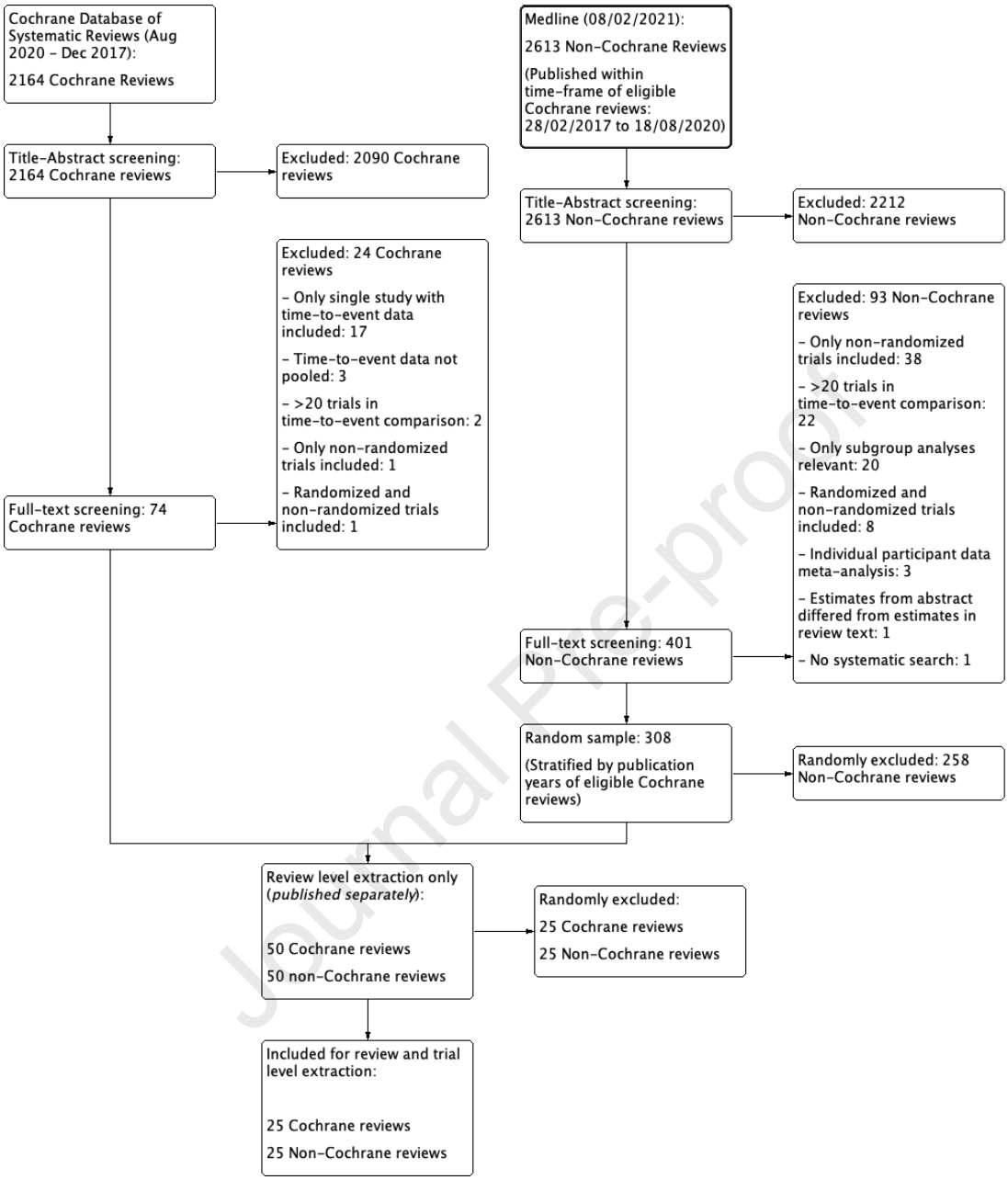
#	Item	Option	Description
42	Individuals at risk reported?	Yes; No; Not applicable	Was the number of individuals at risk over time reported along the survival curve for the assessed analysis
43	Censored observations presented?	Yes, marked on the survival curve; Yes, reported together with the number of individuals at risk; No; Not applicable	Were censored observations presented for the assessed analysis?
44	Censoring balanced?	Yes; No; Unclear; Not applicable	Please make a judgement whether censoring was balanced between arms or pattern in the trial groups differed over time to a degree that is not corresponding to event rates No = More individuals censored in one trial arm compared to the other or pattern in groups differing over time to a degree that is not corresponding to event rates (e.g. early censoring in one group compared to late censoring in the other
45	Proportional hazards tested?	Yes, visual inspection of curves; Yes, statistical test (e.g. Log-log, Schoenfeld Residuals)); No; Not applicable (e.g. no HRs calculated)	Was the proportional hazards assumption tested for the assessed analysis by the trial conductors?
46	Outcome of proportional hazards assessment	Reasonably proportional; Non-proportional; Not applicable; Not reported	What was the outcome of the authors assessment of proportional hazards for the assessed analysis?
47	If analyzed population differs: Experimental individuals		If the analyzed population differs from allocated population (e.g. "mITT", "PP or "as treated" analysis, separate adjusted analysis, exclusion for missing outcome data, ...): number of individuals analyzed in experimental arm If necessary type: "Unclear" "Not reported" "Not applicable"
48	If analyzed population differs: Control individuals		If the analyzed population differs from allocated population (e.g. "mITT", "PP or "as treated" analysis, separate adjusted analysis, exclusion for missing outcome data, ...): number of individuals analyzed in control arm If necessary type: "Unclear" "Not reported" "Not applicable"
49	MOD specifically reported?	Yes; No; Unclear; Complete follow-up/ no LTFU reported at trial level; Complete follow-up/ no LTFU visible on trial outcome level	Missing outcome data specifically reported for this outcome analysis?
50	MOD differing from "MOD of allocated population"	Yes; No	Does missing outcome data for this analysis differ from "missing outcome data for allocated population" (e.g. because of mITT, PP or as treated analysis, separate adjusted analysis, ...): -Data not already excluded in analysis set -Irrespective of whether explicitly reported or not Use explicitly reported data before using data reported for "missing outcome data for allocated population" and subtracting the individuals excluded (e.g. mITT - individuals excluded before treatment)
51	Comparator interventions specifically reported?	Yes; No	Reception of comparator interventions specifically reported for this outcome?
52	Competing events in experimental group		How many patients experienced a competing event in the experimental group? If necessary type: "Unclear" "Not reported" "Not applicable"
53	Competing events in control group		How many patients experienced a competing event in the control group? If necessary type: "Unclear" "Not reported" "Not applicable"
54	Comments on sample size and numbers		
55	Events in experimental arm		Number of events in experimental arm If necessary type: "Not reported" "Unclear"
56	Events in control arm		Number of events in control arm If necessary type: "Not reported" "Unclear"

#	Item	Option	Description
57	Final experimental number at risk		Final number at risk at last follow-up in experimental arm from curve that is applicable to analysis If necessary type: "Unclear" "Not applicable" (if not curves or no number at risk for this analysis are reported)
58	Final control number at risk		Final number at risk at last follow-up in control arm from curve that is applicable to analysis If necessary type: "Unclear" "Not applicable" (if not curves or no number at risk for this analysis are reported)
59	Comments regarding the sample size		
60	Applicable HR		Hazard ratio applicable to meta-analysis as reported in the trial publication (if HR reported as effect measure) Empty field = "Not reported/ unclear" NA = "Not applicable"
61	Applicable lower 95% CI		Lower 95% CI for the assessed analysis (if HR reported as effect measure) Empty field = "Not reported/ unclear" NA = "Not applicable"
62	Applicable upper 95% CI		Upper 95% CI for the assessed analysis (if HR reported as effect measure) Empty field = "Not reported/ unclear" NA = "Not applicable"
63	HR <1 increased or decreased risk of event in experimental group	Decreased risk; Increased risk; Unclear; Not applicable (e.g. no HR calculated)	Does a HR <1 indicate an increased or decreased risk of the outcome in the group that is assessed as experimental group in the trial publication?
64	HR for events or absence of events?	Event; Absence of event; Unclear; Not applicable (e.g. no HR calculated)	Is the respective HR from this trial publication applicable for events (e.g. death, relapse) or absence of events (e.g. overall survival, all cause mortality)?
65	Applicable HR directly available	Was inverted; Was pooled from multiple study arms (as reported in review); Other (specify); Not applicable (e.g. no HR calculated); HR(s) from study differ from HR in MA (HR likely recalculated)	How did the review authors include this HR into the meta-analysis? Is this HR directly available from the assessed trial publication or was it altered in any way by the review authors? (e.g. inverted, pooled from more than one experimental arm, etc.)
66	Standard error		Standard error (if HR not reported or not the appropriate effect measure) If necessary type: "Not reported" "Not applicable"
67	Variance		Variance (log-rank; if HR not reported or not the appropriate effect measure)
68	P-value		P-value of statistical test of group comparison (logrank if not otherwise (e.g. Mantel-Hanezel, Cox) reported in comment field; if HR not reported or not the appropriate effect measure) Empty field = "Not reported/ unclear" NA = "Not applicable"
69	Field for comments on effect measures		Field for comments on effect measures (e.g. type of test, where necessary: Log-rank observed minus-expected events, etc.)
70	Follow-up time specifically reported?	Yes; No	Information on duration of follow-up specifically reported for this outcome? "Specifically" could be outcome specific in primary publication/ publication at hand or in separate publication that includes data for this outcome
71	Comments on trial outcome follow-up		Comments on trial outcome follow-up if reported specifically
72	Missing data handling	Censored; Excluded from analysis; Single imputation; Multiple imputation; Sensitivity analyses; No missing data; other (please specify); Unclear; Not reported; Not applicable (e.g. no effect measure calculated); Complete follow-up/ no LTFU reported at trial level	How was missing data handled?
73	Comment on MOD analyses		E.g. did the advanced methods to assess the robustness of results for this outcome towards MOD change interpretation?
74	Advanced methods competing events	Fine and Gray and cumulative incidence curves; Cumulative incidence curves; Other;	Were any advanced methods to assess the robustness of results for this outcome towards competing events used?

#	Item	Option	Description
		No; Not applicable	
75	Comment competing event methods		E.g. did the advanced methods to assess the robustness of results for this outcome towards competing events change interpretation?
76	Advanced methods informative censoring	Rank preserving structural failure time; Inverse probability (censoring) weighting; Iterative parameter estimation; Multiple; Other; No; Not applicable	Where any advanced methods to assess the robustness of results for this outcome towards informative censoring (non-administrative censoring) used?
77	Comment informative censoring methods		E.g. did the advanced methods to assess the robustness of results for this outcome towards informative censoring (non-administrative censoring) change interpretation?
78	Advanced methods comparator treatments	Rank preserving structural failure time; Inverse probability (censoring) weighting; Iterative parameter estimation; Multiple; Other; No; Not applicable	Were any advanced methods to assess the robustness of results towards the reception of comparator treatments in trial participants used?
79	Comment advanced methods comparator treatments		E.g. did the advanced methods to assess the robustness of results towards the reception of comparator treatments in trial participants change interpretation?
80	Comments on time-to-event specific methods or alternative TTE analytic methods		General comments on advanced time-to-event specific methods or alternative time-to-event analytic methods, that were included in the trial report (e.g. Inverse Probability (Censoring) Weighting applied to adjust for specific event, such as the reception of a relevant third intervention)

Abbreviations: CI = confidence interval; HR = hazard ratio; ITT = intention-to-treat; IQR = interquartile range; IPD = individual participant data; LTFU = loss-to-follow-up; mITT = modified intention-to-treat; MOD = missing outcome data; PP = per protocol; RoB = risk of bias; SD = standard deviation; TTE = time-to-event

Appendix A3: Flow-diagram



Appendix A4: Characteristics of trials and trial outcomes that data could not be extracted for

Domain		Trial (N = 13)	Trial outcome (N = 18)	Review (N = 50)
Review type	Cochrane	69% (9)	67% (12)	12% (6)
	Non-Cochrane	31% (4)	33% (6)	8% (4)
Multiple publications of trial referenced in review	Yes	54% (7)		10% (5)
Other than primary trial publication used for extraction	Yes	8% (1)		2% (1)
	No	23% (3)		6% (3)
	Not applicable (e.g. no primary publication defined)	69% (9)		14% (7)
Relevant TTE data in primary publication for particular outcome	No		39% (7)	6% (3)
	No primary publication highlighted		17% (3)	4% (2)
	Only single publication referenced in review		44% (8)	12% (6)
Reasons why extraction was not possible	No TTE data in cited publications		56% (10)	10% (5)
	Completely unclear which/ whether data was included in review		22% (4)	6% (3)
	Full text or publication where TTE data is reported is not accessible		11% (2)	4% (2)
	Data received from secondary source (e.g. contact with authors)		11% (2)	2% (1)
Review outcomes adopted from trials	All-cause mortality/ Overall survival		67% (12)	18% (9)
	Disease-free survival		11% (2)	4% (2)
	Composite of all-cause death, myocardial infarction or stroke		6% (1)	2% (1)
	Progression-free survival		6% (1)	2% (1)
	Thrombolysis in myocardial infarction (TIMI) major bleeding		6% (1)	2% (1)
	Time to death from prostate cancer		6% (1)	2% (1)
Methods to recalculate TTE data reported by review authors	HR and confidence intervals		22% (4)	2% (1)
	Time-point specific survival times		11% (2)	2% (1)
	P-value together with additional information (e.g. events)		6% (1)	2% (1)
	Not specified for this trial outcome		61% (11)	16% (8)
Relative effect measures included in review	Favorable, statistically significant		11% (2)	4% (2)
	Favorable, statistically non-significant (CI crosses 1)		28% (5)	8% (4)
	Unfavorable, statistically significant		6% (1)	2% (1)
	Unfavorable, statistically significant non-significant (CI crosses 1)		50% (9)	10% (5)
	Direction of effect unclear (HR = 1)		6% (1)	2% (1)
TTE specific risk of bias rating at trial level	No		100% (18)	

Abbreviations: CI = confidence interval, HR = hazard ratio, TTE = time-to-event

Appendix A5: Characteristics of included reviews

Domain		Review			Review outcome		
		Overall (N = 50)	**Cochrane (n = 25)**	**Non-Cochrane (n = 25)**	**Overall (N = 70)**	**Cochrane (n = 33)**	**Non-Cochrane (n = 37)**
Publication							
Publication year	2017	26% (13)	24% (6)	28% (7)			
	2018	26% (13)	24% (6)	28% (7)			
	2019	28% (14)	28% (7)	28% (7)			
	2020	20% (10)	24% (6)	16% (4)			
	Median (IQR)	2018 (2017.25 - 2019)	2019 (2018 - 2019)	2018 (2017 - 2019)			
	Mean (range)	2018.42 (2017 - 2020)	2018.52 (2017 - 2020)	2018.32 (2017 - 2020)			
Journal Impact Factor (2021)	Median (IQR)		12.008	4,501 (3.372 – 6.883)			
	Mean (range)		12.008	8,06 (1.817 – 35.855)			
Review updates		28% (14)	48% (12)	8% (2)			
Multiple review comparisons		30% (15)	44% (11)	16% (4)			
Population							
Medical field	Neoplasms	82% (41)	88% (22)	76% (19)			
	Diseases of the circulatory system	14% (7)	8% (2)	20% (5)			
	Diseases of the skin and subcutaneous tissue	6% (3)	12% (3)	0% (0)			
	Diseases of blood, blood-forming organs, immune mechanism	4% (2)	4% (1)	4% (1)			
	Other	4% (2)	4% (1)	4% (1)			
Medical condition	Breast cancer	14% (7)	12% (3)	16% (4)			
	Colorectal cancer	8% (4)	12% (3)	4% (1)			
	Prostate cancer	6% (3)	12% (3)	0% (0)			
	Biliary tract cancer	4% (2)	0% (0)	8% (2)			
	Gastric cancer	4% (2)	0% (0)	8% (2)			
	Non-ischaemic cardiomyopathy	4% (2)	0% (0)	8% (2)			
	Non-small cell lung cancer	4% (2)	4% (1)	4% (1)			
	Ovarian cancer	4% (2)	8% (2)	0% (0)			
	Other	52% (26)	52% (13)	52% (13)			
Clinical stage	Early/ First line	34% (17)	44% (11)	24% (6)			
	Advanced/ Second or third line	30% (15)	24% (6)	36% (9)			
	No restriction	20% (10)	24% (6)	16% (4)			
	Not reported	2% (1)	0% (0)	4% (1)			
	Not applicable	14% (7)	8% (2)	20% (5)			
Age group	Adults	92% (46)	96% (24)	88% (22)			
	Both	2% (1)	4% (1)	0% (0)			

Domain		Review			Review outcome		
		Overall (N = 50)	Cochrane (n = 25)	Non-Cochrane (n = 25)	Overall (N = 70)	Cochrane (n = 33)	Non-Cochrane (n = 37)
	Not reported	6% (3)	0% (0)	12% (3)			
Interventions							
Comparisons	Biologics/ drug vs. Biologic/ drug	32% (16)	20% (5)	44% (11)			
	Surgical procedure vs. Surgical procedure	8% (4)	12% (3)	4% (1)			
	Biologics/ drug vs. Observation	6% (3)	0% (0)	12% (3)			
	Biologics/ drug (schedule alteration)	4% (2)	4% (1)	4% (1)			
	Biologics/ drug vs. Placebo	4% (2)	0% (0)	8% (2)			
	Follow-up strategy vs. Follow-up strategy	4% (2)	8% (2)	0% (0)			
	Other	42% (21)	56% (14)	28% (7)			
Comparator treatment considered?	No	100% (50)	100% (25)	100% (25)			
Outcomes - Planned							
Planned outcome number	Median (IQR)	6 (4 - 8)	7 (6 - 8)	4 (2 - 5)			
	Mean (range)	5.89 (1 - 15)	6.72 (3 - 10)	4.95 (1 - 15)			
Planned TTE outcome number	Median (IQR)	2 (2 - 2)	2 (2 - 2)	2 (1.75 - 2)			
	Mean (range)	2.35 (1 - 12)	2.00 (1 - 3)	2.71 (1 - 12)			
Number of outcomes analyzed	Median (IQR)	5 (3 - 6)	5 (5 - 6)	4 (2 - 5)			
	Mean (range)	5.24 (1 - 18)	5.44 (1 - 10)	5.04 (1 - 18)			
Number of TTE outcomes analyzed	Median (IQR)	2 (1 - 2)	2 (1 - 2)	2 (1 - 2)			
	Mean (range)	2.22 (1 - 12)	1.80 (1 - 3)	2.64 (1 - 12)			
Outcomes - Definition							
Outcome reporting per review	Absence of event only	56% (28)	40% (10)	72% (18)			
	Event only	26% (13)	24% (6)	28% (7)			
	Both (with reasoning) only	6% (3)	12% (3)	0% (0)			
	Mixed	4% (2)	8% (2)	0% (0)			
	At least one unclear	8% (4)	16% (4)	0% (0)			
Outcome reporting per individual outcome	Absence of event				59% (41)	45% (15)	70% (26)
	Event				26% (18)	21% (7)	30% (11)
	Both (with reasoning)				10% (7)	21% (7)	0% (0)
	Unclear				6% (4)	12% (4)	0% (0)
Reviews including follow-up start in outcome definitions	Randomization	38% (19)	60% (15)	16% (4)			
	Allocated treatment	4% (2)	4% (1)	4% (1)			
	Enrollment	2% (1)	4% (1)	0% (0)			
	At least one not applicable	56% (28)	32% (8)	80% (20)			
Follow-up start included in outcome definition	Randomization				43% (30)	67% (22)	22% (8)
	Allocated treatment				4% (3)	3% (1)	5% (2)
	Enrollment				1% (1)	3% (1)	0% (0)
	Not reported				51% (36)	27% (9)	73% (27)
Reviews mentioning heterogeneous TTE outcome definitions	In discussion	4% (2)	0% (0)	8% (2)			
	In results	2% (1)	4% (1)	0% (0)			
	Not reported	94% (47)	96% (24)	92% (23)			
Heterogeneous outcome definitions discussed	Yes	2% (1)	0% (0)	4% (1)	1% (1)	0% (0)	3% (1)
	No	100% (50)	100% (25)	100% (25)	99% (69)	100% (33)	97% (36)
Follow-up							

Domain		Review			Review outcome		
		Overall (N = 50)	Cochrane (n = 25)	Non-Cochrane (n = 25)	Overall (N = 70)	Cochrane (n = 33)	Non-Cochrane (n = 37)
Reviews reporting a planned follow-up duration	Minimum duration of follow-up required	4% (2)	4% (1)	4% (1)			
	Longest follow-up	2% (1)	4% (1)	0% (0)			
	Maximum duration of follow-up specified	2% (1)	4% (1)	0% (0)			
	Not reported	92% (46)	88% (22)	96% (24)			
Follow-up time specification for TTE outcomes	Longest follow-up	6% (3)	8% (2)	4% (1)	6% (4)	6% (2)	5% (2)
	Minimum duration of follow-up required	2% (1)	4% (1)	0% (0)	3% (2)	6% (2)	0% (0)
	Maximum duration of follow-up specified	2% (1)	4% (1)	0% (0)	1% (1)	3% (1)	0% (0)
	Not reported	90% (45)	84% (21)	96% (24)	90% (63)	85% (28)	95% (35)
Sample size							
Number of included studies in reviews and meta-analyses	Median (IQR)	5 (4 - 8)	5 (4 - 9)	5 (4 - 7)	4 (2.25 - 5)	3 (2 - 5)	4 (4 - 5)
	Mean (range)	7 (2 - 21)	7 (2 - 21)	5 (2 - 13)	5 (2 - 15)	4 (2 - 15)	5 (2 - 12)
Total population in review or meta-analysis	Median (IQR)	1697 (957 - 3838)	1184 (505 - 4190)	1728 (1370 - 3252)	1521 (571 - 4580.5)	571 (351 - 1741)	1948 (1455 - 5093)
	Mean (range)	3621 (307 - 38723)	3229 (307 - 13216)	3962 (343 - 38723)	4133 (181 - 38723)	2369 (181 - 12528)	6117 (623 - 38723)
	Not reported	14% (7)	20% (5)	8% (2)	27% (19)	18% (6)	35% (13)
Analyses - Comparative effect measures							
HR type eligible in reviews	HR/ log(HR) not further specified	88% (44)	92% (23)	84% (21)			
	HR/ log(HR) from Cox model	2% (1)	4% (1)	0% (0)			
	Cox model HR/ log HR, log-rank and Kaplan Meier Curve	2% (1)	4% (1)	0% (0)			
	Not reported	8% (4)	0% (0)	16% (4)			
HR types eligible per outcome	HR/ log(HR) from Cox model	2% (1)	4% (1)	0% (0)	1% (1)	3% (1)	0% (0)
	Not reported	98% (49)	96% (24)	100% (25)	99% (69)	97% (32)	100% (37)
Methods to obtain TTE data per review	HR and confidence intervals	56% (28)	64% (16)	48% (12)			
	Specified set of methods (e.g. Tierney 2008 (4))	44% (22)	76% (19)	12% (3)			
	log(HR) and standard error	20% (10)	32% (8)	8% (2)			
	Survival curves	18% (9)	20% (5)	16% (4)			
	HR with other information (e.g. events)	8% (4)	16% (4)	0% (0)			
	IPD (recalculated or from publication)	4% (2)	4% (1)	4% (1)			
	P-value with additional information (e.g. events)	4% (2)	4% (1)	4% (1)			
	Median survival times	2% (1)	4% (1)	0% (0)			
	Reported, but method not clear	2% (1)	4% (1)	0% (0)			
	Risk ratio	2% (1)	0% (0)	4% (1)			
	Unclear	4% (2)	4% (1)	4% (1)			
	Not reported	20% (10)	0% (0)	40% (10)			
Recalculation of TTE data reported for an outcome	Yes	36% (18)	28% (7)	4% (1)			
	Not reported	164% (82)	132% (33)	196% (49)			
Methods to obtain TTE data for an outcome	HR and confidence intervals	6% (3)	12% (3)	0% (0)	7% (5)	15% (5)	0% (0)
	P-value with additional information (e.g. events)	4% (2)	8% (2)	0% (0)	3% (2)	6% (2)	0% (0)
	Survival curves	4% (2)	4% (1)	4% (1)	3% (2)	3% (1)	3% (1)
	IPD (recalculated or from publication)	2% (1)	4% (1)	0% (0)	1% (1)	3% (1)	0% (0)

Domain		Review			Review outcome		
		Overall (N = 50)	Cochrane (n = 25)	Non-Cochrane (n = 25)	Overall (N = 70)	Cochrane (n = 33)	Non-Cochrane (n = 37)
	Time point specific survival times	2% (1)	4% (1)	0% (0)	3% (2)	6% (2)	0% (0)
	Unclear	6% (3)	12% (3)	0% (0)	6% (4)	12% (4)	0% (0)
Analyses - ITT/ PP							
Types of analyses eligible in reviews	ITT	42% (21)	76% (19)	8% (2)			
	Not reported	58% (29)	24% (6)	92% (23)			
Types of analyses eligible for outcome analyses	Not reported	100% (50)	100% (25)	100% (25)	100% (70)	100% (33)	100% (37)
Types of analyses included in reviews	ITT	18% (9)	20% (5)	16% (4)			
	Included trial(s) did not report type of analysis	6% (3)	8% (2)	4% (1)			
	Not reported for all trials	16% (8)	24% (6)	8% (2)			
	Not reported for any trial	66% (33)	56% (14)	76% (19)			
Types of analyses included in outcome analyses	ITT	2% (1)	4% (1)	0% (0)	3% (2)	6% (2)	0% (0)
	Not reported for all trials	2% (1)	4% (1)	0% (0)	1% (1)	3% (1)	0% (0)
	Not reported	96% (48)	92% (23)	100% (25)	96% (67)	91% (30)	100% (37)
Unadjusted/ adjusted HRs eligible in reviews	Hierachical (adjusted before unadjusted)	4% (2)	8% (2)	0% (0)			
	Both	2% (1)	4% (1)	0% (0)			
	Unadjusted only	2% (1)	4% (1)	0% (0)			
	Adjusted only	2% (1)	4% (1)	0% (0)			
	Hierachical (unadjusted before adjusted)	2% (1)	4% (1)	0% (0)			
	Unclear	6% (3)	12% (3)	0% (0)			
	Not reported	82% (41)	64% (16)	100% (25)			
Dealing with unadjusted/ adjusted HRs	Unclear	8% (4)	16% (4)	0% (0)			
	Not reported	10% (5)	20% (5)	0% (0)			
	Not applicable	82% (41)	64% (16)	100% (25)			
Stratified HRs eligible in reviews	Yes	2% (1)	4% (1)	0% (0)			
	No	98% (49)	96% (24)	100% (25)			
Unadjusted/ adjusted HRs eligible in outcome analyses	Unadjusted only	2% (1)	4% (1)	0% (0)	1% (1)	3% (1)	0% (0)
	Adjusted only	2% (1)	4% (1)	0% (0)	1% (1)	3% (1)	0% (0)
	Not reported	96% (48)	92% (23)	100% (25)	97% (68)	94% (31)	100% (37)
Dealing with unadjusted/ adjusted HRs	Only adjusted HRs included, others likely excluded	2% (1)	4% (1)	0% (0)	1% (1)	3% (1)	0% (0)
	Unadjusted HRs recalculated	2% (1)	4% (1)	0% (0)	1% (1)	3% (1)	0% (0)
	Not applicable	96% (48)	92% (23)	100% (25)	97% (68)	94% (31)	100% (37)
Unadjusted/ adjusted HRs discussed in reviews	In results	2% (1)	4% (1)	0% (0)			
	Not reported	96% (48)	92% (23)	100% (25)			
Unadjusted/ adjusted analyzed discussed for individual outcomes	No	100% (50)	100% (25)	100% (25)	100% (70)	100% (33)	100% (37)
Analyses - Varying follow-up between included trials							
Dealing with varying follow-up in reviews	Sensitivity analyses (e.g. shorter/longer)	10% (5)	16% (4)	4% (1)			
	Included in meta-regression	4% (2)	0% (0)	8% (2)			
	Exclusion of studies with divergent follow-up time	2% (1)	0% (0)	4% (1)			
	Included in interpretation of heterogeneity	2% (1)	4% (1)	0% (0)			
	Mentioned as RoB criterion in methods	2% (1)	4% (1)	0% (0)			
	Unclear	2% (1)	0% (0)	4% (1)			
	Not reported	78% (39)	76% (19)	80% (20)			

Domain		Review			Review outcome		
		Overall (N = 50)	Cochrane (n = 25)	Non-Cochrane (n = 25)	Overall (N = 70)	Cochrane (n = 33)	Non-Cochrane (n = 37)
Dealing with varying follow-up for individual outcomes	Results reported for multiple time-points	2% (1)	4% (1)	0% (0)	1% (1)	3% (1)	0% (0)
	Not reported per outcome	98% (49)	96% (24)	100% (25)	99% (69)	97% (32)	100% (37)
Varying follow-up discussed	In discussion	18% (9)	16% (4)	20% (5)			
	In results	8% (4)	8% (2)	8% (2)			
	In results and in discussion	6% (3)	8% (2)	4% (1)			
	Not reported	68% (34)	68% (17)	68% (17)			
Varying follow-up discussed for individual outcomes	Yes	6% (3)	4% (1)	8% (2)	4% (3)	3% (1)	5% (2)
	No	96% (48)	96% (24)	96% (24)	96% (67)	97% (32)	95% (35)
Analyses - Missing outcome data							
Dealing with missing outcome data in reviews	Mentioned as RoB criterion in methods	68% (34)	92% (23)	44% (11)			
	Contact with authors	40% (20)	76% (19)	4% (1)			
	Sensitivity analyses (rate of missing values)	8% (4)	16% (4)	0% (0)			
	Single imputation	4% (2)	8% (2)	0% (0)			
	Not reported	28% (14)	0% (0)	56% (14)			
Dealing with missing data in individual outcomes	Not reported per outcome	100% (50)	100% (25)	100% (25)	100% (70)	100% (33)	100% (37)
Missing outcome data discussed	In results	56% (28)	80% (20)	32% (8)			
	In results and discussion	8% (4)	16% (4)	0% (0)			
	Not reported	36% (18)	4% (1)	68% (17)			
Missing outcome data discussed for individual outcomes	No	100% (50)	100% (25)	100% (25)	100% (70)	100% (33)	100% (37)
Analyses - Informative censoring							
Dealing with informative censoring in reviews	Mentioned as RoB criterion in methods	2% (1)	4% (1)	0% (0)			
	Not reported	98% (49)	96% (24)	100% (25)			
Dealing with informative censoring for individual outcomes	Not reported per outcome	100% (50)	100% (25)	100% (25)	100% (70)	100% (33)	100% (37)
Informative censoring discussed	Not reported	100% (50)	100% (25)	100% (25)			
Informative censoring discussed for individual outcomes	No	100% (50)	100% (25)	100% (25)	100% (70)	100% (33)	100% (37)
Analyses - Competing events							
Dealing with deaths as competing event in reviews	Not reported	34% (17)	32% (8)	36% (9)			
	No outcomes with potential competing events	66% (33)	68% (17)	64% (16)			
Dealing with deaths as competing events in individual outcomes	Not reported per outcome	46% (23)	48% (12)	44% (11)	37% (26)	39% (13)	35% (13)
	Not applicable	70% (35)	64% (16)	76% (19)	63% (44)	61% (20)	65% (24)
Deaths as competing events discussed	Not reported	34% (17)	32% (8)	36% (9)			
	No outcomes with potential competing events	66% (33)	68% (17)	64% (16)			
Deaths as competing events discussed for individual outcomes	No	46% (23)	48% (12)	44% (11)	37% (26)	39% (13)	35% (13)
	Not applicable	70% (35)	64% (16)	76% (19)	63% (44)	61% (20)	65% (24)
Analyses - Treatment switching							
Dealing with treatment switching in reviews	RoB criterion in methods	2% (1)	4% (1)	0% (0)			
	Presence reported for each trial	2% (1)	4% (1)	0% (0)			
	Sensitivity analysis (e.g. rate of participants), RoB criterion	2% (1)	0% (0)	4% (1)			
	Not reported	94% (47)	92% (23)	96% (24)			
Dealing with treatment switching in individual outcomes	Not reported per outcome	100% (50)	100% (25)	100% (25)	100% (70)	100% (33)	100% (37)
Treatment switching discussed	In results	6% (3)	8% (2)	4% (1)			
	In discussion	4% (2)	8% (2)	0% (0)			

Domain		Review			Review outcome		
		Overall (N = 50)	Cochrane (n = 25)	Non-Cochrane (n = 25)	Overall (N = 70)	Cochrane (n = 33)	Non-Cochrane (n = 37)
	Not reported	90% (45)	84% (21)	96% (24)			
Treatment switching discussed for individual outcomes	No	100% (50)	100% (25)	100% (25)	100% (70)	100% (33)	100% (37)
Analyses - Proportional hazards							
Proportional hazards assessed in reviews	Not reported	100% (50)	100% (25)	100% (25)			
Proportional hazards assessed in individual outcomes	Not reported per outcome	100% (50)	100% (25)	100% (25)	100% (70)	100% (33)	100% (37)
Dealing with (non-)proportional hazards	Not applicable	100% (50)	100% (25)	100% (25)			
Test for proportionality for individual outcomes	Not applicable	100% (50)	100% (25)	100% (25)	100% (70)	100% (33)	100% (37)
Non-proportionality of hazards indicated	Not applicable	100% (50)	100% (25)	100% (25)	100% (70)	100% (33)	100% (37)
Dealing with (non-)proportional hazards	Not applicable	100% (50)	100% (25)	100% (25)	100% (70)	100% (33)	100% (37)
Results							
Interpretation of pooled HR	Favourable, statistically significant	44% (22)	20% (5)	68% (17)	37% (26)	18% (6)	54% (20)
	Favourable, statistically non-significant	46% (23)	48% (12)	44% (11)	39% (27)	42% (14)	35% (13)
	Unfavourable, statistically significant	4% (2)	4% (1)	4% (1)	3% (2)	3% (1)	3% (1)
	Unfavourable, statistically significant non-significant	26% (13)	44% (11)	8% (2)	20% (14)	33% (11)	8% (3)
	Direction of effect unclear (HR = 1)	4% (2)	4% (1)	4% (1)	1% (1)	3% (1)	0% (0)
HR for events or non events	Event	96% (48)	96% (24)	96% (24)	97% (68)	97% (32)	97% (36)
	Unclear	4% (2)	4% (1)	4% (1)	3% (2)	3% (1)	3% (1)
Interpretation of HR<1	Decreased risk	92% (46)	92% (23)	92% (23)	94% (66)	94% (31)	95% (35)
	Increased risk	4% (2)	4% (1)	4% (1)	3% (2)	3% (1)	3% (1)
	Unclear	4% (2)	4% (1)	4% (1)	3% (2)	3% (1)	3% (1)
Trial HRs inverted	Not reported	100% (50)	100% (25)	100% (25)	100% (70)	100% (33)	100% (37)
Risk of Bias							
Risk of bias tools specified	RoB 1, study level	58% (29)	64% (16)	52% (13)			
	RoB 1, outcome level	18% (9)	36% (9)	0% (0)			
	Other (e.g. CONSORT)	8% (4)	0% (0)	16% (4)			
	Jadad scale	4% (2)	0% (0)	8% (2)			
	RoB 2	4% (2)	0% (0)	8% (2)			
	No RoB assessment	8% (4)	0% (0)	16% (4)			
TTE specific risk of bias criteria used	Yes (e.g. "risk of bias related to censoring")	2% (1)	4% (1)	0% (0)			
	No	90% (45)	96% (24)	84% (21)			
	Not applicable	8% (4)	0% (0)	16% (4)			

Abbreviations: HR = hazard ratio; IPD = individual participant data, IQR = interquartile range; ITT = intention to treat; PP = per protocol; RoB = risk of bias; TTE = time-to-event

Figure A6: Frequency and combinations of available time-to-event summary data items per individual trial time-to-event outcome

	Which individual time-to-event data items (✓) were available for individual trial outcomes?											*Total*
Survival curves	✓	✓	✓	✓	✓	✓	✓	✓				83% (263)
P-value (log-rank)	✓	✓	✓	✓	✓	✓	✓			✓		76% (240)
HR or log(HR)	✓	✓	✓	✓					✓	✓		72% (226)
Time-point specific survival rates (per arm)	✓		✓		✓							46% (145)
Median survival times (per arm)		✓	✓				✓					40% (125)
*Other**											✓	33% (105)
Total	16% (50)	15% (47)	10% (33)	7% (22)	6% (18)	3% (10)	3% (9)	3% (8)	2% (7)	2% (7)	33% (104)	

*Appendix-Figure 1: Frequency (rows) and combinations (colomns) of available time-to-event summary data items per individual time-to-event outcome available in trial publications. Numbers in coloms represent the available individual items (e.g. surival curves), numbers in the bottom rows represent the frequency of available combinations (e.g. survival curves together with a P-value, HR/ log(HR) and time-point specific survival times). * Other included, e.g., median or time-point specific cumulative incidence per arm, observed and expected events, event times per participant and restricted mean survival time (RMST) (Abbreviations: HR = hazard ration).*

Table A7: *Extended time-to-event specific methodological characteristics of trials included in review time-to-event outcome meta-analyses*

Domain		Trial (N = 235)	Trial outcome						Review		
			Overall (N = 315)	Cochrane (n = 131)	Non-Cochrane (n = 184)	All-cause mortality/ Overall survival (n = 198)	Combined, including all-cause mortality (n = 77)	Not including all-cause mortality (n = 40)	Overall (N = 50)	Cochrane (n = 25)	Non-Cochrane (n = 25)
Analyses - Comparative effect measures											
Available TTE data	HR/log(HR), P-value (log-rank), Survival curves, Time-point specific survival rates (per arm)	15% (36)	16% (50)	20% (26)	13% (24)	14% (28)	25% (19)	8% (3)	38% (19)	32% (8)	44% (11)
	HR or log(HR), P-value (log-rank), Survival curves, Median survival times (per arm)	18% (42)	15% (47)	8% (10)	20% (37)	10% (20)	31% (24)	8% (3)	36% (18)	28% (7)	44% (11)
	HR or log(HR), P-value (log-rank), Survival curves, Median survival times (per arm), Time-point specific survival rates (per arm)	12% (29)	10% (33)	4% (5)	15% (28)	13% (25)	10% (8)	0% (0)	28% (14)	16% (4)	40% (10)
	HR or log(HR), P-value (log-rank), Survival curves	9% (21)	7% (22)	7% (9)	7% (13)	9% (18)	5% (4)	0% (0)	30% (15)	24% (6)	36% (9)
	P-value (log-rank), Survival curves, Time-point specific survival rates (per arm)	7% (16)	6% (18)	4% (5)	7% (13)	6% (11)	4% (3)	10% (4)	16% (8)	12% (3)	20% (5)
	P-value (log-rank), Survival curves	4% (10)	3% (10)	2% (3)	4% (7)	4% (7)	0% (0)	8% (3)	10% (5)	8% (2)	12% (3)
	P-value (log-rank), Survival curves, Median survival times (per arm)	4% (9)	3% (9)	2% (2)	4% (7)	4% (8)	0% (0)	3% (1)	10% (5)	8% (2)	12% (3)
	Survival curves	3% (6)	3% (8)	5% (7)	1% (1)	3% (5)	0% (0)	8% (3)	10% (5)	16% (4)	4% (1)
	HR or log(HR)	2% (5)	2% (7)	1% (1)	3% (6)	3% (5)	1% (1)	3% (1)	10% (5)	4% (1)	16% (4)
	HR or log(HR), P-value (log-rank)	3% (6)	2% (7)	4% (5)	1% (2)	2% (4)	3% (2)	3% (1)	12% (6)	16% (4)	8% (2)
	Other	41% (96)	33% (104)	44% (58)	25% (46)	34% (67)	21% (16)	53% (21)	172% (86)	192% (48)	152% (38)
TTE data	Survival curves	83% (195)	83% (263)	76% (100)	89% (163)	85% (168)	90% (69)	65% (26)	92% (46)	84% (21)	100% (25)
	P-value (log-rank)	78% (183)	76% (240)	71% (93)	80% (147)	75% (148)	87% (67)	63% (25)	94% (47)	92% (23)	96% (24)
	HR or log(HR)	71% (166)	72% (226)	64% (84)	77% (142)	68% (135)	95% (73)	45% (18)	90% (45)	80% (20)	100% (25)
	Time-point specific survival rates (per arm)	49% (115)	46% (145)	50% (66)	43% (79)	48% (95)	49% (38)	30% (12)	82% (41)	76% (19)	88% (22)
	Median survival times (per arm)	43% (100)	40% (125)	29% (38)	47% (87)	39% (78)	51% (39)	20% (8)	58% (29)	56% (14)	60% (15)
	Type of test unclear or not reported	8% (18)	6% (20)	7% (9)	6% (11)	6% (12)	5% (4)	10% (4)	26% (13)	20% (5)	32% (8)
	Median cumulative incidence (per arm)	2% (5)	2% (6)	2% (2)	2% (4)	1% (2)	4% (3)	3% (1)	8% (4)	8% (2)	8% (2)
	Mean and standard deviation per arm	2% (4)	1% (4)	2% (3)	1% (1)	2% (4)	0% (0)	0% (0)	8% (4)	12% (3)	4% (1)
	Observed and expected events (log-rank) or hazard rates	2% (4)	1% (4)	3% (4)	0% (0)	2% (4)	0% (0)	0% (0)	4% (2)	8% (2)	0% (0)
	Wilcoxon-Gehan test	1% (3)	1% (3)	2% (3)	0% (0)	2% (3)	0% (0)	0% (0)	6% (3)	12% (3)	0% (0)
	Time-point specific cumulative incidence	1% (2)	1% (2)	1% (1)	1% (1)	1% (2)	0% (0)	5% (2)	6% (3)	8% (2)	4% (1)
	Event times per participant	1% (3)	1% (3)	2% (3)	0% (0)	1% (2)	0% (0)	3% (1)	4% (2)	8% (2)	0% (0)
	Test results not numerically reported	1% (3)	1% (3)	0% (0)	2% (3)	2% (3)	0% (0)	0% (0)	4% (2)	0% (0)	8% (2)
	Cox model coefficients and/or P-values	1% (3)	1% (3)	1% (1)	1% (2)	1% (1)	0% (0)	5% (2)	4% (2)	4% (1)	4% (1)
	Restricted mean survival time (RMST)	1% (2)	1% (2)	1% (1)	1% (1)	1% (2)	0% (0)	0% (0)	4% (2)	4% (1)	4% (1)
	Greys Test	0% (1)	0% (1)	1% (1)	0% (0)	0% (0)	0% (0)	3% (1)	2% (1)	4% (1)	0% (0)
	Absolute risk reduction (Andersen and Altman methodology)	0% (1)	0% (1)	0% (0)	1% (1)	0% (0)	0% (0)	3% (1)	2% (1)	0% (0)	4% (1)

Domain		Trial (N = 235)	Trial outcome						Review		
			Overall (N = 315)	Cochrane (n = 131)	Non-Cochrane (n = 184)	All-cause mortality/ Overall survival (n = 198)	Combined, including all-cause mortality (n = 77)	Not including all-cause mortality (n = 40)	Overall (N = 50)	Cochrane (n = 25)	Non-Cochrane (n = 25)
HR calculation	Cox model	59% (138)	60% (188)	50% (66)	66% (122)	57% (113)	75% (58)	43% (17)	86% (43)	72% (18)	100% (25)
	Log rank	1% (2)	1% (2)	0% (0)	1% (2)	1% (2)	0% (0)	0% (0)	4% (2)	0% (0)	8% (2)
	Cox model and RPSFT model	0% (1)	0% (1)	0% (0)	1% (1)	1% (1)	0% (0)	0% (0)	2% (1)	0% (0)	4% (1)
	Cox model and Cox model with time dependent variable(s)	0% (1)	0% (1)	1% (1)	0% (0)	1% (1)	0% (0)	0% (0)	2% (1)	4% (1)	0% (0)
	Cox Markov model	0% (1)	0% (1)	1% (1)	0% (0)	1% (1)	0% (0)	0% (0)	2% (1)	4% (1)	0% (0)
	Cox model and Fine and Gray model	0% (1)	0% (1)	0% (0)	1% (1)	0% (0)	0% (0)	3% (1)	2% (1)	0% (0)	4% (1)
	Andersen-Gill regression model	0% (1)	0% (1)	1% (1)	0% (0)	1% (1)	0% (0)	0% (0)	2% (1)	4% (1)	0% (0)
	Cox model and Log rank method	0% (1)	0% (1)	0% (0)	1% (1)	0% (0)	1% (1)	0% (0)	2% (1)	0% (0)	4% (1)
	Fine and Gray	0% (1)	0% (1)	1% (1)	0% (0)	0% (0)	0% (0)	3% (1)	2% (1)	4% (1)	0% (0)
	Unclear	0% (1)	1% (2)	0% (0)	1% (2)	1% (1)	1% (1)	0% (0)	2% (1)	0% (0)	4% (1)
	Not reported	12% (29)	11% (34)	13% (17)	9% (17)	10% (19)	18% (14)	3% (1)	40% (20)	40% (10)	40% (10)
	No HR calculated	31% (72)	26% (82)	34% (44)	21% (38)	30% (59)	4% (3)	50% (20)	54% (27)	64% (16)	44% (11)
Survival plots											
Survival plots available	Kaplan-Meier	80% (188)	79% (249)	74% (97)	83% (152)	81% (161)	88% (68)	50% (20)	92% (46)	84% (21)	100% (25)
	Cumulative incidence	3% (6)	3% (8)	2% (2)	3% (6)	1% (2)	1% (1)	13% (5)	10% (5)	8% (2)	12% (3)
	Type of curve not reported	2% (4)	2% (5)	2% (2)	2% (3)	2% (3)	0% (0)	5% (2)	6% (3)	8% (2)	4% (1)
	Adjusted Kaplan-Meier	0% (1)	0% (1)	1% (1)	0% (0)	1% (1)	0% (0)	0% (0)	2% (1)	4% (1)	0% (0)
	No, but for other analyses of this outcome	3% (6)	2% (7)	4% (5)	1% (2)	3% (6)	0% (0)	0% (0)	8% (4)	12% (3)	4% (1)
	No, no graphs were presented	17% (40)	14% (45)	18% (24)	11% (21)	13% (25)	9% (7)	33% (13)	56% (28)	56% (14)	56% (14)
Number of individuals at risk reported	Yes	57% (134)	58% (184)	44% (58)	68% (126)	55% (108)	78% (60)	40% (16)	88% (44)	76% (19)	100% (25)
	No, but for other analyses of this outcome	2% (4)	2% (5)	2% (3)	1% (2)	2% (4)	1% (1)	0% (0)	6% (3)	8% (2)	4% (1)
	No	28% (66)	26% (81)	35% (46)	19% (35)	31% (61)	12% (9)	28% (11)	54% (27)	60% (15)	48% (12)
	Not applicable	17% (40)	14% (45)	18% (24)	11% (21)	13% (25)	9% (7)	33% (13)	56% (28)	56% (14)	56% (14)
Censoring events reported	On survival curve	38% (89)	38% (119)	24% (32)	47% (87)	37% (74)	49% (38)	18% (7)	68% (34)	64% (16)	72% (18)
	On survival curve and reported with individuals at risk	3% (8)	3% (11)	3% (4)	4% (7)	3% (5)	8% (6)	0% (0)	14% (7)	12% (3)	16% (4)
	No	43% (102)	43% (136)	51% (67)	38% (69)	45% (90)	34% (26)	50% (20)	80% (40)	80% (20)	80% (20)
	Not applicable	18% (43)	16% (49)	21% (28)	11% (21)	15% (29)	9% (7)	33% (13)	62% (31)	68% (17)	56% (14)
Censoring balanced	Yes	32% (75)	30% (96)	21% (28)	37% (68)	31% (61)	40% (31)	10% (4)	66% (33)	64% (16)	68% (17)
	No	9% (22)	8% (24)	5% (7)	9% (17)	6% (12)	14% (11)	3% (1)	28% (14)	20% (5)	36% (9)
	Unclear	4% (9)	3% (9)	1% (1)	4% (8)	3% (5)	3% (2)	5% (2)	14% (7)	4% (1)	24% (6)
	Not applicable	61% (143)	59% (186)	73% (95)	49% (91)	61% (121)	43% (33)	80% (32)	88% (44)	92% (23)	84% (21)
TTE data recalculation reported in revies per outcome											
TTE data recalculation	HR together with other infomation (e.g. events)	3% (8)	5% (15)	2% (2)	7% (13)	4% (8)	1% (1)	15% (6)	4% (2)	4% (1)	4% (1)
	P-value together with additional information (e.g. events)	6% (13)	5% (15)	11% (14)	1% (1)	6% (12)	3% (2)	3% (1)	8% (4)	12% (3)	4% (1)
	HR and confidence intervals	3% (8)	3% (10)	3% (4)	3% (6)	2% (4)	8% (6)	0% (0)	6% (3)	8% (2)	4% (1)

Domain		Trial (N = 235)	Trial outcome						Review		
			Overall (N = 315)	Cochrane (n = 131)	Non-Cochrane (n = 184)	All-cause mortality/ Overall survival (n = 198)	Combined, including all-cause mortality (n = 77)	Not including all-cause mortality (n = 40)	Overall (N = 50)	Cochrane (n = 25)	Non-Cochrane (n = 25)
	IPD (recalculated or from publication)	3% (6)	2% (6)	2% (3)	2% (3)	3% (5)	0% (0)	3% (1)	6% (3)	8% (2)	4% (1)
	Survival curves	2% (5)	2% (5)	2% (2)	2% (3)	2% (4)	0% (0)	3% (1)	8% (4)	8% (2)	8% (2)
	Only specified to be recalculated or obtained from authors	1% (2)	1% (2)	2% (2)	0% (0)	1% (2)	0% (0)	0% (0)	2% (1)	4% (1)	0% (0)
	Time-point specific survival times	0% (1)	1% (2)	2% (2)	0% (0)	1% (1)	0% (0)	3% (1)	2% (1)	4% (1)	0% (0)
	Not specified for this trial outcome	83% (196)	83% (260)	78% (102)	86% (158)	82% (162)	88% (68)	75% (30)	86% (43)	76% (19)	96% (24)

Abbreviations: HR = hazard ratio; IPD = individual participant data; RPSFT = Rank Preserving Structural Failure Time; TTE = time-to-event

Table A8: *General methodological characteristics of trials included in review time-to-event outcome meta-analyses*

Domain		Trial (N = 235)	Trial outcome						Review		
			Overall (N = 315)	Cochrane (n = 131)	Non-Cochrane (n = 184)	All-cause mortality/ Overall survival (n = 198)	Combined, including all-cause mortality (n = 77)	Not including all-cause mortality (n = 40)	Overall (N = 50)	Cochrane (n = 25)	Non-Cochrane (n = 25)
Analyses available in trial publication											
Available types of analyses	ITT	63% (147)	63% (198)	58% (76)	66% (122)	61% (121)	73% (56)	53% (21)	92% (46)	92% (23)	92% (23)
	ITT, Per-protocol	6% (14)	5% (15)	8% (10)	3% (5)	6% (11)	4% (3)	3% (1)	26% (13)	32% (8)	20% (5)
	Modified ITT	4% (9)	4% (12)	1% (1)	6% (11)	4% (8)	5% (4)	0% (0)	14% (7)	4% (1)	24% (6)
	Per-protocol	2% (5)	2% (6)	3% (4)	1% (2)	2% (4)	0% (0)	5% (2)	10% (5)	12% (3)	8% (2)
	ITT, As treated	2% (4)	2% (5)	0% (0)	3% (5)	2% (3)	1% (1)	3% (1)	8% (4)	0% (0)	16% (4)
	Modified ITT, Per-protocol	1% (2)	1% (3)	1% (1)	1% (2)	1% (1)	3% (2)	0% (0)	4% (2)	4% (1)	4% (1)
	ITT, Modified ITT	0% (1)	0% (1)	0% (0)	1% (1)	0% (0)	0% (0)	3% (1)	2% (1)	0% (0)	4% (1)
	As treated	0% (1)	0% (1)	0% (0)	1% (1)	0% (0)	0% (0)	3% (1)	2% (1)	0% (0)	4% (1)
	ITT, Per-protocol, As treated	0% (1)	0% (1)	0% (0)	1% (1)	0% (0)	1% (1)	0% (0)	2% (1)	0% (0)	4% (1)
	Unclear	0% (1)	0% (1)	1% (1)	0% (0)	1% (1)	0% (0)	0% (0)	2% (1)	4% (1)	0% (0)
	Not reported	26% (60)	23% (72)	29% (38)	18% (34)	25% (49)	13% (10)	33% (13)	58% (29)	68% (17)	48% (12)
ITT analysis available	ITT	69% (161)	70% (220)	66% (86)	73% (134)	68% (135)	79% (61)	60% (24)	96% (48)	96% (24)	96% (24)
	Modified ITT	5% (11)	5% (15)	2% (2)	7% (13)	5% (9)	8% (6)	0% (0)	18% (9)	8% (2)	28% (7)
	No	3% (6)	2% (7)	3% (4)	2% (3)	2% (4)	0% (0)	8% (3)	12% (6)	12% (3)	12% (3)
	Unclear	0% (1)	0% (1)	1% (1)	0% (0)	1% (1)	0% (0)	0% (0)	2% (1)	4% (1)	0% (0)
	Not reported	26% (60)	23% (72)	29% (38)	18% (34)	25% (49)	13% (10)	33% (13)	58% (29)	68% (17)	48% (12)
Available ITT analysis in complete population	Yes	43% (102)	44% (139)	34% (44)	52% (95)	43% (85)	49% (38)	40% (16)	92% (46)	88% (22)	96% (24)
	No	19% (45)	20% (62)	30% (39)	13% (23)	20% (39)	22% (17)	15% (6)	46% (23)	56% (14)	36% (9)
	Unclear	2% (4)	2% (5)	0% ()	3% (5)	1% (2)	3% (2)	3% (1)	8% (4)	0% ()	16% (4)
	Not applicable (no ITT, only mITT or subgroup))	38% (89)	35% (109)	37% (48)	33% (61)	36% (72)	26% (20)	43% (17)	78% (39)	80% (20)	76% (19)
Analyses included in review meta-analyses											
Type of analysis included	ITT	67% (158)	69% (216)	65% (85)	71% (131)	67% (133)	79% (61)	55% (22)	96% (48)	96% (24)	96% (24)
	mITT	5% (11)	5% (16)	2% (2)	8% (14)	5% (9)	6% (5)	5% (2)	18% (9)	8% (2)	28% (7)
	Per protocol	2% (5)	2% (7)	3% (4)	2% (3)	3% (5)	1% (1)	3% (1)	10% (5)	12% (3)	8% (2)
	As treated	1% (3)	0% (1)	0% (0)	1% (1)	0% (0)	0% (0)	3% (1)	6% (3)	8% (2)	4% (1)
	Unclear	0% (1)	1% (3)	2% (2)	1% (1)	1% (2)	0% (0)	3% (1)	2% (1)	0% (0)	4% (1)
	Not reported	26% (60)	23% (72)	29% (38)	18% (34)	25% (49)	13% (10)	33% (13)	58% (29)	68% (17)	48% (12)
Included analysis in complete population	Yes	55% (130)	55% (174)	45% (59)	63% (115)	56% (110)	55% (42)	55% (22)	96% (48)	96% (24)	96% (24)
	No	32% (75)	32% (100)	44% (57)	23% (43)	31% (62)	31% (24)	35% (14)	70% (35)	80% (20)	60% (15)
	Unclear	4% (9)	3% (10)	3% (4)	3% (6)	3% (5)	3% (2)	8% (3)	16% (8)	12% (3)	20% (5)
	Not reported	6% (15)	5% (17)	6% (8)	5% (9)	6% (12)	6% (5)	0% (0)	20% (10)	16% (4)	24% (6)
	Not applicable (e.g. subgroup included in review)	5% (12)	4% (14)	2% (3)	6% (11)	5% (9)	5% (4)	3% (1)	14% (7)	8% (2)	20% (5)

Domain		Trial (N = 235)	Trial outcome						Review		
			Overall (N = 315)	Cochrane (n = 131)	Non-Cochrane (n = 184)	All-cause mortality/ Overall survival (n = 198)	Combined, including all-cause mortality (n = 77)	Not including all-cause mortality (n = 40)	Overall (N = 50)	Cochrane (n = 25)	Non-Cochrane (n = 25)
Included analysis in allocated arm	Yes	87% (204)	88% (276)	79% (103)	94% (173)	88% (175)	91% (70)	78% (31)	98% (49)	96% (24)	100% (25)
	No	0% (1)	0% (1)	0% (0)	1% (1)	0% (0)	0% (0)	3% (1)	2% (1)	0% (0)	4% (1)
	Unclear	2% (5)	2% (5)	3% (4)	1% (1)	1% (2)	1% (1)	5% (2)	10% (5)	16% (4)	4% (1)
	Not reported	12% (28)	10% (33)	18% (24)	5% (9)	11% (21)	8% (6)	15% (6)	40% (20)	56% (14)	24% (6)
Covariate adjustment of available estimates in trial publication											
Available unadjusted and adjusted analyses	Adjusted, baseline characteristics	13% (30)	12% (39)	11% (15)	13% (24)	12% (24)	13% (10)	13% (5)	36% (18)	24% (6)	48% (12)
	Unadjusted; Adjusted, baseline characteristics	11% (26)	9% (29)	11% (14)	8% (15)	9% (17)	13% (10)	5% (2)	36% (18)	40% (10)	32% (8)
	Stratified, baseline characteristics	6% (15)	7% (23)	6% (8)	8% (15)	6% (12)	9% (7)	10% (4)	22% (11)	16% (4)	28% (7)
	Stratified, randomization stratification factors	5% (11)	5% (16)	2% (3)	7% (13)	6% (11)	6% (5)	0% (0)	14% (7)	8% (2)	20% (5)
	Unadjusted (univariate including treatment variables only)	5% (12)	5% (15)	5% (7)	4% (8)	5% (9)	1% (1)	13% (5)	22% (11)	16% (4)	28% (7)
	Stratified, factors unclear	3% (6)	3% (8)	1% (1)	4% (7)	2% (4)	5% (4)	0% (0)	10% (5)	4% (1)	16% (4)
	Unadjusted; Stratified, factors unclear	2% (5)	2% (7)	0% (0)	4% (7)	3% (5)	3% (2)	0% (0)	4% (2)	0% (0)	8% (2)
	Adjusted, factors unclear	3% (6)	2% (7)	4% (5)	1% (2)	3% (5)	1% (1)	3% (1)	10% (5)	16% (4)	4% (1)
	Adjusted, baseline characteristics; Not reported	1% (3)	2% (6)	3% (4)	1% (2)	2% (3)	4% (3)	0% (0)	4% (2)	4% (1)	4% (1)
	Unadjusted; Stratified, baseline characteristics	1% (2)	1% (2)	1% (1)	1% (1)	1% (2)	0% (0)	0% (0)	4% (2)	4% (1)	4% (1)
	Adjusted, factors unclear; Stratified, factors unclear	0% (1)	1% (2)	0% (0)	1% (2)	1% (1)	1% (1)	0% (0)	2% (1)	0% (0)	4% (1)
	Adjusted, baseline characteristics; Stratified, factors unclear	0% (1)	0% (1)	1% (1)	0% (0)	0% (0)	1% (1)	0% (0)	2% (1)	4% (1)	0% (0)
	Unadjusted; Stratified, randomization stratification factors	0% (1)	0% (1)	0% (0)	1% (1)	1% (1)	0% (0)	0% (0)	2% (1)	0% (0)	4% (1)
	Adjusted, factors unclear; Stratified, baseline characteristics	0% (1)	0% (1)	0% (0)	1% (1)	0% (0)	1% (1)	0% (0)	2% (1)	0% (0)	4% (1)
	Adjusted, baseline characteristics; Stratified, baseline characteristics	0% (1)	0% (1)	0% (0)	1% (1)	0% (0)	1% (1)	0% (0)	2% (1)	0% (0)	4% (1)
	Unclear	3% (8)	3% (11)	2% (3)	4% (8)	4% (7)	4% (3)	3% (1)	14% (7)	8% (2)	20% (5)
	Not reported	20% (48)	19% (59)	18% (24)	19% (35)	18% (35)	30% (23)	3% (1)	54% (27)	40% (10)	68% (17)
	Not applicable (no HR directly reported)	31% (73)	28% (87)	34% (45)	23% (42)	31% (62)	5% (4)	53% (21)	54% (27)	64% (16)	44% (11)
Unadjusted HR reported	Unadjusted (univariate including treatment variables only)	19% (44)	17% (54)	17% (22)	17% (32)	17% (34)	17% (13)	18% (7)	54% (27)	48% (12)	60% (15)
	No	51% (120)	52% (163)	47% (61)	55% (102)	48% (95)	74% (57)	28% (11)	78% (39)	56% (14)	100% (25)
	Unclear	3% (8)	3% (11)	2% (3)	4% (8)	4% (7)	4% (3)	3% (1)	14% (7)	8% (2)	20% (5)
	Not applicable (no HR directly reported)	31% (73)	28% (87)	34% (45)	23% (42)	31% (62)	5% (4)	53% (21)	54% (27)	64% (16)	44% (11)
Adjusted HR reported	Adjusted, baseline characteristics	26% (60)	24% (75)	26% (34)	22% (41)	22% (43)	32% (25)	18% (7)	58% (29)	52% (13)	64% (16)
	Adjusted, factors unclear	3% (8)	3% (10)	4% (5)	3% (5)	3% (6)	4% (3)	3% (1)	14% (7)	16% (4)	12% (3)
	No	42% (98)	42% (132)	34% (44)	48% (88)	40% (80)	55% (42)	25% (10)	70% (35)	52% (13)	88% (22)
	Unclear	3% (8)	3% (11)	2% (3)	4% (8)	4% (7)	4% (3)	3% (1)	14% (7)	8% (2)	20% (5)
	Not applicable (no HR directly reported)	31% (73)	28% (87)	34% (45)	23% (42)	31% (62)	5% (4)	53% (21)	54% (27)	64% (16)	44% (11)
	Stratified, baseline characteristics	8% (19)	9% (27)	7% (9)	10% (18)	7% (14)	12% (9)	10% (4)	24% (12)	16% (4)	32% (8)

Domain		Trial (N = 235)	Trial outcome						Review		
			Overall (N = 315)	Cochrane (n = 131)	Non-Cochrane (n = 184)	All-cause mortality/ Overall survival (n = 198)	Combined, including all-cause mortality (n = 77)	Not including all-cause mortality (n = 40)	Overall (N = 50)	Cochrane (n = 25)	Non-Cochrane (n = 25)
Stratified HR reported	Stratified, factors unclear	5% (11)	6% (18)	2% (2)	9% (16)	5% (10)	10% (8)	0% (0)	14% (7)	4% (1)	24% (6)
	Stratified, randomization stratification factors	5% (12)	5% (17)	2% (3)	8% (14)	6% (12)	6% (5)	0% (0)	14% (7)	8% (2)	20% (5)
	No	51% (119)	49% (155)	53% (69)	47% (86)	47% (93)	62% (48)	35% (14)	84% (42)	80% (20)	88% (22)
	Unclear	3% (8)	3% (11)	2% (3)	4% (8)	4% (7)	4% (3)	3% (1)	14% (7)	8% (2)	20% (5)
	Not applicable (no HR directly reported)	31% (73)	28% (87)	34% (45)	23% (42)	31% (62)	5% (4)	53% (21)	54% (27)	64% (16)	44% (11)
Available unadjusted log-rank P-values	Stratified, baseline characteristics	7% (16)	8% (24)	5% (6)	10% (18)	5% (9)	19% (15)	0% (0)	14% (7)	8% (2)	20% (5)
	Unadjusted (univariate including treatment variables only)	7% (17)	6% (19)	5% (7)	7% (12)	7% (14)	4% (3)	5% (2)	26% (13)	20% (5)	32% (8)
	Stratified, randomization stratification factors	5% (11)	5% (16)	4% (5)	6% (11)	6% (11)	6% (5)	0% (0)	12% (6)	12% (3)	12% (3)
	Stratified, factors unclear	5% (11)	4% (14)	0% (0)	8% (14)	5% (9)	6% (5)	0% (0)	18% (9)	0% (0)	36% (9)
	Unadjusted; Adjusted, baseline characteristics	3% (8)	3% (10)	5% (6)	2% (4)	3% (6)	4% (3)	3% (1)	14% (7)	16% (4)	12% (3)
	Adjusted, baseline characteristics	3% (7)	3% (8)	0% (0)	4% (8)	4% (7)	1% (1)	0% (0)	14% (7)	0% (0)	28% (7)
	Adjusted, baseline characteristics; Stratified, randomization stratification factors	0% (1)	1% (2)	0% (0)	1% (2)	1% (1)	1% (1)	0% (0)	2% (1)	0% (0)	4% (1)
	Adjusted, baseline characteristics; Not reported	0% (1)	1% (2)	0% (0)	1% (2)	1% (1)	1% (1)	0% (0)	2% (1)	0% (0)	4% (1)
	Adjusted, factors unclear; Stratified, factors unclear	0% (1)	1% (2)	0% (0)	1% (2)	1% (1)	1% (1)	0% (0)	2% (1)	0% (0)	4% (1)
	Unadjusted; Stratified, baseline characteristics	0% (1)	0% (1)	0% (0)	1% (1)	0% (0)	1% (1)	0% (0)	32% (16)	60% (15)	4% (1)
	Not reported	49% (115)	47% (148)	53% (69)	43% (79)	48% (96)	40% (31)	53% (21)	70% (35)	80% (20)	60% (15)
	Not applicable (no log-rank P-value directly reported)	23% (53)	22% (69)	29% (38)	17% (31)	22% (43)	13% (10)	40% (16)	38% (19)	0% (0)	76% (19)
Unadjusted log-rank P-value reported	Unadjusted (univariate including treatment variables only)	11% (26)	10% (30)	10% (13)	9% (17)	10% (20)	9% (7)	8% (3)	38% (19)	32% (8)	44% (11)
	No	69% (163)	69% (216)	61% (80)	74% (136)	68% (135)	78% (60)	53% (21)	88% (44)	80% (20)	96% (24)
	Not applicable (no log-rank P-value directly reported)	23% (53)	22% (69)	29% (38)	17% (31)	22% (43)	13% (10)	40% (16)	60% (30)	60% (15)	60% (15)
Adjusted log-rank P-value reported	Adjusted, baseline characteristics	7% (17)	7% (22)	5% (6)	9% (16)	8% (15)	8% (6)	3% (1)	26% (13)	16% (4)	36% (9)
	Adjusted, but factors unclear	0% (1)	1% (2)	0% (0)	1% (2)	1% (1)	1% (1)	0% (0)	2% (1)	0% (0)	4% (1)
	No	72% (170)	70% (222)	66% (87)	73% (135)	70% (139)	78% (60)	58% (23)	90% (45)	88% (22)	92% (23)
	Not applicable (no log-rank P-value directly reported)	23% (53)	22% (69)	29% (38)	17% (31)	22% (43)	13% (10)	40% (16)	58% (29)	60% (15)	56% (14)
Stratified log-rank P-value	Stratified, baseline characteristics	7% (17)	8% (25)	5% (6)	10% (19)	5% (9)	21% (16)	0% (0)	16% (8)	8% (2)	24% (6)
	Stratified, randomization stratification factors	5% (12)	6% (18)	4% (5)	7% (13)	6% (12)	8% (6)	0% (0)	14% (7)	12% (3)	16% (4)
	Stratified, but factors unclear	5% (12)	5% (16)	0% (0)	9% (16)	5% (10)	8% (6)	0% (0)	20% (10)	0% (0)	40% (10)
	No	63% (148)	59% (187)	63% (82)	57% (105)	63% (124)	51% (39)	60% (24)	90% (45)	92% (23)	88% (22)
	Not applicable (no log-rank P-value directly reported)	23% (53)	22% (69)	29% (38)	17% (31)	22% (43)	13% (10)	40% (16)	60% (30)	60% (15)	60% (15)
Covariate adjustment of estimate included in review meta-analysis											
Included analysis unadjusted, adjusted, stratified	Unadjusted	28% (66)	25% (80)	36% (47)	18% (33)	28% (56)	16% (12)	30% (12)	60% (30)	68% (17)	52% (13)
	Stratified	15% (36)	18% (56)	8% (11)	24% (45)	16% (31)	27% (21)	10% (4)	34% (17)	20% (5)	48% (12)
	Adjusted	14% (32)	13% (41)	8% (11)	16% (30)	12% (24)	16% (12)	13% (5)	44% (22)	28% (7)	60% (15)
	Unclear	7% (17)	6% (20)	7% (9)	6% (11)	7% (14)	4% (3)	8% (3)	26% (13)	24% (6)	28% (7)

Domain		Trial (N = 235)	Trial outcome						Review		
			Overall (N = 315)	Cochrane (n = 131)	Non-Cochrane (n = 184)	All-cause mortality/ Overall survival (n = 198)	Combined, including all-cause mortality (n = 77)	Not including all-cause mortality (n = 40)	Overall (N = 50)	Cochrane (n = 25)	Non-Cochrane (n = 25)
	Not reported	40% (95)	37% (118)	40% (53)	35% (65)	37% (73)	38% (29)	40% (16)	74% (37)	68% (17)	80% (20)
Abbreviations: HR = hazard ratio; ITT = intention to treat; TTE = time-to-event											

Table A9: *Outcome results of trials included in review time-to-event outcome meta-analyses.*

Domain		Trial outcome				Review		
		Overall (N = 315)	**All-cause mortality/ Overall survival (n = 198)**	**Combined, including all-cause mortality (n = 77)**	**Not including all-cause mortality (n = 40)**	**Overall (N = 50)**	**Cochrane (n = 25)**	**Non-Cochrane (n = 25)**
Relative effect measures included in review according to point estimate and confidence interval								
*Effect in review**	Favourable, statistically significant	26% (81)	21% (42)	35% (27)	30% (12)	54% (27)	40% (10)	68% (17)
	Favourable, statistically non-sign.	51% (161)	57% (113)	49% (38)	25% (10)	78% (39)	68% (17)	88% (22)
	Unfavourable, statisitically significant	3% (10)	2% (4)	0% (0)	15% (6)	8% (4)	8% (2)	8% (2)
	Unfavourable, statisitically non-sign.	19% (60)	19% (38)	13% (10)	30% (12)	58% (29)	64% (16)	52% (13)
	Direction of effect unclear (HR = 1)	1% (3)	1% (1)	3% (2)	0% (0)	6% (3)	4% (1)	8% (2)
Relative effect measures from trial included in review								
*Effect of trial HR**	Favourable, statistically significant	18% (57)	15% (30)	29% (22)	13% (5)	44% (22)	24% (6)	64% (16)
	Favourable, statistically non-sign.	29% (91)	28% (55)	40% (31)	13% (5)	68% (34)	52% (13)	84% (21)
	Unfavourable, statistically significant	3% (8)	1% (1)	4% (3)	10% (4)	10% (5)	4% (1)	16% (4)
	Unfavourable, statistically non-sign.	12% (39)	14% (28)	12% (9)	5% (2)	46% (23)	44% (11)	48% (12)
	Confidence level unclear	3% (11)	3% (6)	5% (4)	3% (1)	18% (9)	16% (4)	20% (5)
	Direction of effect unclear	0% (1)	0% (0)	1% (1)	0% (0)	2% (1)	0% (0)	4% (1)
	Not applicable	34% (108)	39% (78)	9% (7)	58% (23)	64% (32)	76% (19)	52% (13)
HR <1 in experimental arm [#]	Decreased risk	57% (179)	54% (106)	77% (59)	35% (14)	84% (42)	68% (17)	100% (25)
	Increased risk	5% (16)	3% (6)	9% (7)	8% (3)	20% (10)	24% (6)	16% (4)
	Unclear	4% (13)	5% (9)	5% (4)	0% (0)	18% (9)	20% (5)	16% (4)
	Not applicable (e.g. no HR)	34% (107)	39% (77)	9% (7)	58% (23)	62% (31)	76% (19)	48% (12)
HR for events	Event	60% (190)	56% (110)	82% (63)	43% (17)	86% (43)	72% (18)	100% (25)
	Absence of event	2% (5)	1% (2)	4% (3)	0% (0)	6% (3)	4% (1)	8% (2)
	Unclear	4% (13)	5% (9)	5% (4)	0% (0)	18% (9)	20% (5)	16% (4)
	Not applicable (e.g. no HR)	34% (107)	39% (77)	9% (7)	58% (23)	62% (31)	76% (19)	48% (12)
HR directly available	Trial HR directly available	51% (162)	48% (96)	71% (55)	28% (11)	80% (40)	60% (15)	100% (25)
	Trial HR inverted	7% (23)	6% (11)	10% (8)	10% (4)	30% (15)	40% (10)	20% (5)
	Other [§]	10% (33)	11% (22)	12% (9)	5% (2)	30% (15)	40% (10)	20% (5)
	Not applicable (e.g. no HR calculated)	31% (97)	35% (69)	6% (5)	58% (23)	56% (28)	68% (17)	44% (11)

Abbreviations: CI = Confidence interval; HR = Hazard ratio; IQR = Interquartile range

**Favourable/ unfavourable corresponds to review intervention indicated, for example, in Summary of Findings tables or forest plots. The direction is based on the point estimate. [#]Decreased/increased risk of an event as HR < 1 is based on the review intervention and the review authors interpretation of the effect, i.e. when a HR<1 for overall survival that was interpreted as "beneficial for intervention", that HR clearly represented a decreased the risk of the event (death), irrespective of whether review authors named it as absence of the event (overall survival). [$]Other includes,*

e.g., difference between trial HR/ CI and HR/ CI in forest plot, unclear or different confidence levels (e.g. 99%, 80% or 97.5% CIs) or explicit reporting by review authors not to have included a given trial HR.

Table A10: *Extended specific trial characteristics with relevance for time-to-event outcomes in trials included in review time-to-event outcome meta-analyses*

Domain		Trial			Trial outcome	Review		
		Overall (N = 235)	Cochrane (n = 102)	Non-Cochrane (n = 133)	(N = 315)	Overall (N = 50)	Cochrane (n = 25)	Non-Cochrane (n = 25)
Follow-up								
Follow-up measures available	Follow-up reported for trial	79% (185)	72% (73)	84% (112)		96% (48)	92% (23)	100% (25)
	Follow-up reported for outcomes	3% (6)	4% (4)	2% (2)		10% (5)	12% (3)	8% (2)
	No indicator of follow-up reported	19% (44)	25% (25)	14% (19)		44% (22)	56% (14)	32% (8)
Follow-up measures	Median	66% (154)	57% (58)	72% (96)		92% (46)	84% (21)	100% (25)
	Minimum follow-up	25% (59)	23% (23)	27% (36)		56% (28)	44% (11)	68% (17)
	Maximum follow-up	23% (53)	19% (19)	26% (34)		54% (27)	40% (10)	68% (17)
	IQR/ lower and upper range of IQR	18% (43)	11% (11)	24% (32)		56% (28)	40% (10)	72% (18)
	Mean	3% (8)	3% (3)	4% (5)		12% (6)	8% (2)	16% (4)
	Fixed time-point of outcome measurement only	3% (8)	5% (5)	2% (3)		12% (6)	16% (4)	8% (2)
	Standard deviation	2% (5)	1% (1)	3% (4)		6% (3)	4% (1)	8% (2)
	95% CI of median	1% (2)	1% (1)	1% (1)		4% (2)	4% (1)	4% (1)
	Follow-up reported per outcome	3% (6)	4% (4)	2% (2)		10% (5)	12% (3)	8% (2)
Follow-up calculation	Median follow-up, surviving patients only	8% (19)	9% (9)	8% (10)		26% (13)	32% (8)	20% (5)
	Median follow-up, all patients	5% (11)	5% (5)	5% (6)		16% (8)	16% (4)	16% (4)
	Reverse Kaplan-Meier	3% (7)	4% (4)	2% (3)		14% (7)	16% (4)	12% (3)
	Median follow-up, multiple (e.g. all patients and surviving only)	1% (2)	2% (2)	0% (0)		4% (2)	8% (2)	0% ()
	Median follow-up, excluding censored	1% (2)	1% (1)	1% (1)		4% (2)	4% (1)	4% (1)
	Mean follow-up, multiple (e.g. all patients and surviving only)	0% (1)	1% (1)	0% (0)		2% (1)	4% (1)	0% ()
	Unclear	1% (3)	1% (1)	2% (2)		6% (3)	4% (1)	8% (2)
	Not reported	61% (143)	50% (51)	69% (92)		86% (43)	72% (18)	100% (25)
	Not applicable	20% (47)	27% (28)	14% (19)		46% (23)	56% (14)	36% (9)
Overall follow-up measure reported	Yes	59% (138)	46% (47)	68% (91)		84% (42)	72% (18)	96% (24)
Median overall follow-up	Median (IQR)	45 (22.8 - 67.6)	62.3 (44.5 - 98)	31.44 (15 - 48)				
	Mean (range)	52.87 (5 - 229.2)	75.68 (5 - 167)	38.94 (5.1 - 229.2)				
	Not reported/ unclear	47% (111)	54% (55)	42% (56)				
Analyses - Missing outcome data handling in included trials								
Missing outcome data handling	Excluded from analysis	18% (42)	25% (26)	12% (16)	17% (52)	42% (21)	48% (12)	36% (9)
	Censored	11% (26)	13% (13)	10% (13)	11% (36)	34% (17)	36% (9)	32% (8)

Domain		Trial			Trial outcome	Review		
		Overall (N = 235)	Cochrane (n = 102)	Non-Cochrane (n = 133)	(N = 315)	Overall (N = 50)	Cochrane (n = 25)	Non-Cochrane (n = 25)
	Complete follow-up/ no LTFU reported at trial level	11% (25)	6% (6)	14% (19)	10% (31)	28% (14)	24% (6)	32% (8)
	Single imputation	0% (1)	1% (1)	0% (0)	0% (1)	2% (1)	4% (1)	0% (0)
	Multiple imputation	0% (1)	0% (0)	1% (1)	0% (1)	2% (1)	0% (0)	4% (1)
	Unclear	1% (3)	3% (3)	0% (0)	1% (4)	2% (1)	4% (1)	0% (0)
	Not reported	58% (136)	52% (53)	62% (83)	57% (178)	92% (46)	84% (21)	100% (25)
	No missing data	3% (8)	4% (4)	3% (4)	4% (12)	12% (6)	12% (3)	12% (3)
Reported missing outcome data								
Missing outcome data reported	Yes	46% (108)	48% (49)	44% (59)		88% (44)	88% (22)	88% (22)
	Explicitly reported complete follow-up	6% (14)	4% (4)	8% (10)		22% (11)	16% (4)	28% (7)
	Explicitly reported no LTFU	5% (12)	3% (3)	7% (9)		14% (7)	12% (3)	16% (4)
	No	36% (84)	36% (37)	35% (47)		78% (39)	72% (18)	84% (21)
	Reported across arms only	5% (11)	9% (9)	2% (2)		18% (9)	28% (7)	8% (2)
	Reported for individual outcomes	3% (6)	0% (0)	5% (6)		10% (5)	0% (0)	20% (5)
Total missing outcome data in experimental arm	0	2% (4)	4% (4)	0% (0)		8% (4)	16% (4)	0% (0)
	<5%	26% (60)	23% (23)	28% (37)		64% (32)	48% (12)	80% (20)
	≥5%, <10%	6% (14)	5% (5)	7% (9)		26% (13)	20% (5)	32% (8)
	≥10%, <20%	6% (14)	9% (9)	4% (5)		24% (12)	32% (8)	16% (4)
	≥20%	6% (14)	7% (7)	5% (7)		22% (11)	24% (6)	20% (5)
	Not reported	1% (1)	2% (1)	0% (0)		2% (1)	4% (1)	0% (0)
	Not applicable	19% (25)	11% (6)	24% (19)		28% (14)	24% (6)	32% (8)
	Number randomly allocated not unclear	1% (2)	2% (1)	1% (1)		4% (2)	4% (1)	4% (1)
Total missing outcome data in control arm	0	4% (10)	4% (4)	5% (6)		12% (6)	8% (2)	16% (4)
	<5%	20% (47)	24% (24)	17% (23)		50% (25)	52% (13)	48% (12)
	≥5%, <10%	11% (25)	7% (7)	14% (18)		34% (17)	24% (6)	44% (11)
	≥10%, <20%	7% (16)	8% (8)	6% (8)		28% (14)	28% (7)	28% (7)
	≥20%	3% (7)	4% (4)	2% (3)		10% (5)	12% (3)	8% (2)
	Not reported	15% (20)	2% (1)	24% (19)		2% (1)	4% (1)	0% (0)
	Not applicable	19% (25)	11% (6)	24% (19)		28% (14)	24% (6)	32% (8)
	Number randomized not reported/ unclear	1% (2)	2% (1)	1% (1)		4% (2)	4% (1)	4% (1)
Outcome specific missing outcome data reported	Yes	4% (9)	1% (1)	6% (8)		12% (6)	4% (1)	20% (5)
	Complete follow-up/ no LTFU at trial level	11% (26)	7% (7)	14% (19)		28% (14)	24% (6)	32% (8)
	Complete follow-up/ no LTFU on trial outcome level	3% (8)	4% (4)	3% (4)		12% (6)	12% (3)	12% (3)
	No	84% (198)	90% (92)	80% (106)		100% (50)	100% (25)	100% (25)
Analysis - Censoring								
Advanced methods for censoring in trials	Sensitivity analysis (results not shown)	0% (1)	1% (1)	0% (0)		2% (1)	4% (1)	0% (0)
	No	100% (234)	99% (101)	100% (133)		100% (50)	100% (25)	100% (25)

Domain		Trial			Trial outcome	Review		
		Overall (N = 235)	Cochrane (n = 102)	Non-Cochrane (n = 133)	(N = 315)	Overall (N = 50)	Cochrane (n = 25)	Non-Cochrane (n = 25)
Analyses - (Death as) competing event								
Advanced methods for (death as) competing in trials	Cumulative incidence curves	2% (5)	1% (1)	3% (4)		10% (5)	4% (1)	16% (4)
	Fine and Gray and cumulative incidence curves	1% (2)	1% (1)	1% (1)		4% (2)	4% (1)	4% (1)
	No	25% (59)	25% (26)	25% (33)		54% (27)	52% (13)	56% (14)
	Not applicable	86% (202)	85% (87)	86% (115)		92% (46)	88% (22)	96% (24)
Number of (deaths) as competing event in experimental arm	<5%	1% (3)	0% (0)	2% (3)		4% (2)	0% (0)	8% (2)
	≥5%, <10%	1% (2)	0% (0)	2% (2)		2% (1)	0% (0)	4% (1)
	≥10%, <20%	0% (1)	0% (1)	0% (0)		2% (1)	4% (1)	0% (0)
	≥20%, <30%	0% (1)	1% (1)	0% (0)		0% (0)	0% (0)	0% (0)
	≥40%, <50%	0% (1)	1% (1)	0% (0)		2% (1)	4% (1)	0% (0)
	Unclear	3% (7)	2% (2)	4% (5)		8% (4)	8% (2)	8% (2)
	Not reported	11% (25)	13% (13)	9% (12)		20% (10)	20% (5)	20% (5)
	Not applicable	93% (218)	90% (92)	95% (126)		94% (47)	92% (23)	96% (24)
	Number randomized not reported/ unclear	0% (1)	0% (0)	1% (1)		2% (1)	0% (0)	4% (1)
Number of (deaths) as competing event in comparator arm	<5%	1% (3)	0% (0)	2% (3)		4% (2)	0% (0)	8% (2)
	≥5%, <10%	1% (3)	1% (1)	2% (2)		4% (2)	4% (1)	4% (1)
	≥30%, <40%	0% (1)	1% (1)	0% (0)		2% (1)	4% (1)	0% (0)
	≥40%, <50%	0% (0)	0% (0)	0% (0)		0% (0)	0% (0)	0% (0)
	Unclear	3% (7)	2% (2)	4% (5)		8% (4)	8% (2)	8% (2)
	Not reported	11% (25)	13% (13)	9% (12)		20% (10)	20% (5)	20% (5)
	Not applicable	93% (218)	90% (92)	95% (126)		94% (47)	92% (23)	96% (24)
	Number randomized not reported/ unclear	0% (1)	0% (0)	1% (1)		2% (1)	0% (0)	4% (1)
Analyses - Treatment switching								
Advanced methods for treatment switching in trials	Rank preserving structural failure time	0% (1)	0% (0)	1% (1)		2% (1)	0% (0)	4% (1)
	Sensitivity analysis (Cross-over as event)	0% (1)	0% (0)	1% (1)		2% (1)	0% (0)	4% (1)
	Sensitivity analysis (Excluding cross-overs)	0% (1)	0% (0)	1% (1)		2% (1)	0% (0)	4% (1)
	No	93% (219)	97% (99)	90% (120)		438% (219)	396% (99)	480% (120)
	Not applicable	7% (17)	3% (3)	11% (14)		34% (17)	12% (3)	56% (14)
Control treatment in experimental arm	0	7% (17)	7% (7)	8% (10)		20% (10)	20% (5)	20% (5)
	<5%	9% (21)	8% (8)	10% (13)		30% (15)	28% (7)	32% (8)
	≥5%, <10%	5% (12)	8% (8)	3% (4)		20% (10)	24% (6)	16% (4)
	≥10%, <20%	5% (11)	6% (6)	4% (5)		20% (10)	24% (6)	16% (4)
	≥20%, <30%	5% (12)	2% (2)	8% (10)		12% (6)	8% (2)	16% (4)
	≥30%, <40%	1% (3)	1% (1)	2% (2)		4% (2)	4% (1)	4% (1)
	Unclear	0% (1)	1% (1)	0% (0)		14% (7)	8% (2)	20% (5)
	Not reported	64% (151)	65% (66)	64% (85)		90% (45)	84% (21)	96% (24)
	Not applicable	0% (0)	0% (0)	0% (0)		0% (0)	0% (0)	0% (0)
	Number randomized not reported/ unclear	1% (3)	3% (3)	0% (0)		6% (3)	12% (3)	0% (0)

Domain		Trial			Trial outcome	Review		
		Overall (N = 235)	Cochrane (n = 102)	Non-Cochrane (n = 133)	(N = 315)	Overall (N = 50)	Cochrane (n = 25)	Non-Cochrane (n = 25)
Experimental treatment in control arm	0	9% (20)	8% (8)	9% (12)		26% (13)	24% (6)	28% (7)
	<5%	11% (27)	10% (10)	13% (17)		32% (16)	28% (7)	36% (9)
	≥5%, <10%	4% (10)	7% (7)	2% (3)		18% (9)	24% (6)	12% (3)
	≥10%, <20%	3% (7)	2% (2)	4% (5)		12% (6)	8% (2)	16% (4)
	≥20%, <30%	1% (2)	1% (1)	1% (1)		4% (2)	4% (1)	4% (1)
	≥40%, <50%	0% (1)	0% (0)	1% (1)		2% (1)	0% (0)	4% (1)
	≥50%, <60%	0% (1)	0% (0)	1% (1)		2% (1)	0% (0)	4% (1)
	≥70%, <80%	0% (1)	0% (0)	1% (1)		2% (1)	0% (0)	4% (1)
	≥80%, <90%	1% (2)	0% (0)	2% (2)		2% (1)	0% (0)	4% (1)
	Unclear	4% (10)	3% (3)	5% (7)		6% (3)	4% (1)	8% (2)
	Not reported	64% (151)	67% (68)	62% (83)		88% (44)	84% (21)	92% (23)
	Not applicable	0% (0)	0% (0)	0% (0)		0% (0)	0% (0)	0% (0)
	Number randomized not reported/ unclear	1% (3)	3% (3)	0% (0)		6% (3)	12% (3)	0% (0)
Treatment switching reported per outcome	No	100% (235)	100% (102)	100% (133)		100% (50)	100% (25)	100% (25)
Treatment switching pre-specified	Reported as protocol specified	0% (1)	0% (0)	1% (1)		2% (1)	0% (0)	4% (1)
	Reported as not planned or allowed	4% (10)	0% (0)	8% (10)		10% (5)	0% (0)	20% (5)
	Otherwise reported as anticipated	2% (5)	2% (2)	2% (3)		6% (3)	8% (2)	4% (1)
	Unclear	1% (2)	0% (0)	2% (2)		4% (2)	0% (0)	8% (2)
	Not reported	92% (216)	98% (100)	87% (116)		94% (47)	96% (24)	92% (23)
	Not applicable	0% (1)	0% (0)	1% (1)		2% (1)	0% (0)	4% (1)
Treatment switching reported as protocol specified	Yes	1% (2)	2% (2)	0% (0)		4% (2)	8% (2)	0% (0)
	Not reported	77% (181)	66% (67)	86% (114)		98% (49)	96% (24)	100% (25)
	Not applicable	22% (52)	32% (33)	14% (19)		54% (27)	64% (16)	44% (11)
Treatment switching reasons	Disease-related (e.g. disease progression)	12% (29)	8% (8)	16% (21)		32% (16)	24% (6)	40% (10)
	Participant related (e.g. choose to switch)	9% (20)	14% (14)	5% (7)		20% (10)	28% (7)	12% (3)
	Administrative (e.g. interim analysis)	4% (9)	3% (3)	5% (6)		8% (4)	8% (2)	8% (2)
	Pre-condition related (e.g. too obese, allergies)	4% (9)	7% (7)	2% (2)		16% (8)	24% (6)	8% (2)
	Intervention related (e.g. adverse events)	2% (5)	3% (3)	2% (2)		8% (4)	8% (2)	8% (2)
	Investigator/ physician related (e.g. treating physicians decision)	1% (3)	1% (1)	2% (2)		6% (3)	4% (1)	8% (2)
	Not reported	13% (30)	10% (10)	15% (20)		38% (19)	24% (6)	52% (13)
	Not applicable	64% (151)	67% (68)	62% (83)		88% (44)	88% (22)	88% (22)
Analysis - Proportional hazards								
Proportional hazards assumption tested	Yes, statistical test (e.g. Log-log, Schoenfeld Residuals)	8% (19)	8% (8)	8% (11)	7% (23)	32% (16)	32% (8)	32% (8)
	Yes, visual inspection of curves	1% (2)	1% (1)	1% (1)	1% (2)	4% (2)	4% (1)	4% (1)
	No, but for other analyses of this outcome	0% (1)	1% (1)	0% (0)	0% (1)	2% (1)	4% (1)	0% (0)
	No	52% (123)	53% (54)	52% (69)	65% (206)	88% (44)	76% (19)	100% (25)

Domain		Trial			Trial outcome	Review		
		Overall (N = 235)	**Cochrane (n = 102)**	**Non-Cochrane (n = 133)**	**(N = 315)**	**Overall (N = 50)**	**Cochrane (n = 25)**	**Non-Cochrane (n = 25)**
	Not applicable (e.g. no HRs calculated)	29% (69)	38% (39)	23% (30)	26% (83)	52% (26)	64% (16)	40% (10)
Results of proportional hazards tests	Non-proportional	1% (3)	2% (2)	1% (1)	1% (3)	6% (3)	8% (2)	4% (1)
	Reasonably proportional	1% (2)	1% (1)	1% (1)	1% (2)	4% (2)	4% (1)	4% (1)
	Not reported for this analysis, but reasonably for other analysis of this outcome	0% (1)	0% (0)	1% (1)	0% (1)	2% (1)	0% (0)	4% (1)
	Not reported	6% (15)	6% (6)	7% (9)	6% (19)	26% (13)	24% (6)	28% (7)
	Not applicable	92% (216)	92% (94)	92% (122)	92% (290)	100% (50)	100% (25)	100% (25)

Abbreviations: HR = hazard ratio; LTFU = Loss to follow-up; RoB = risk of bias; TTE = time-to-event

What is new?

Key findings

- We identified variable and often insufficient reporting of time-to-event outcomes and associated methods in publications of randomized trials included in aggregate data meta-analyses of current systematic reviews.
- Limited reporting included critical information such as outcome definitions, methods and trial characteristics relevant for assessing the certainty of time-to-event analyses, e.g. informative censoring and proportional hazards. Available time-to-event data varied substantially between trial publications.
- Limitations in trial reporting translate to review publications as well.

What this adds to what is known

- Previous methodological research suggested shortcomings in the reporting of time-to-event outcomes and analyses in study publications. Focusing on trials included in meta-analyses, we showed that these limitations have relevance for meta-analyses in current systematic reviews.

What are the implications and what should be changed?

- Trial authors should strictly adhere to available reporting guidelines for time-to-event analyses in randomized trial publications. Reporting standards for meta-analyses of time-to-event outcomes based on aggregate data are urgently needed.

CRediT authorship contribution statement:

MG: Conceptualization, data curation, formal analysis, investigation, methodology, project administration, writing – original draft, writing - review & editing; CH: Formal analysis, investigation, writing - review & editing; CI: Formal analysis, investigation, writing - review & editing; AMB: Formal analysis, investigation, writing - review & editing; RB: Conceptualization, writing - review & editing; EvD: Conceptualization, writing - review & editing; LGH: Conceptualization, writing - review & editing; IM: systematic search; MT: Conceptualization, writing - review & editing; NK: Conceptualization, data curation, formal analysis, investigation, methodology, writing - review & editing; NS: Conceptualization, methodology, supervision, writing - review & editing

Declaration of interests

☐ The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

☒ The authors declare the following financial interests/personal relationships which may be considered as potential competing interests:

Prof. Dr. Nicole Skoetz reports financial support was provided by German Research Foundation.