HHS Public Access

Author manuscript

J Clin Epidemiol. Author manuscript; available in PMC 2022 September 01.

Published in final edited form as:

J Clin Epidemiol. 2021 September ; 137: 126–132. doi:10.1016/j.jclinepi.2021.03.027.

A structured methodology review showed analyses of functional outcomes are frequently limited to "survivors only" in trials enrolling patients at high risk of death

Elizabeth Colantuoni,
Department of Biostatistics, Bloomberg School of Public Health, Johns Hopkins University, Baltimore MD; Outcomes After Critical Illness and Surgery (OACIS) Group, Johns Hopkins University, Baltimore, Maryland

Ximin Li,
Department of Biostatistics, Bloomberg School of Public Health, Johns Hopkins University, Baltimore MD

Mohamed D. Hashem,
Department of Medicine, Marshfield Clinic, Marshfield WI

Timothy D. Girard,
Clinical Research, Investigation, and Systems Modeling of Acute illness (CRISMA) Center, Department of Critical Care Medicine, University of Pittsburgh School of Medicine, Pittsburgh PA

Daniel O. Scharfstein,
Division of Biostatistics, Department of Population Health Sciences, University of Utah School of Medicine, Salt Lake City, Utah

Dale M. Needham
Outcomes After Critical Illness and Surgery (OACIS) Group, Johns Hopkins University, Baltimore, Maryland; Division of Pulmonary and Critical Care Medicine, Johns Hopkins University School of Medicine, Baltimore, Maryland

Corresponding author: Elizabeth Colantuoni, PhD, ejohnso2@jhmi.edu.

Credit Author Roles:
Elizabeth Colantuoni: Conceptualization, Data Curation, Formal analysis, Project administration, Methodology, Supervision, Writing-original draft, Writing-review & editing
Ximin Li: Data curation, Formal analysis, Writing – review & editing
Mohamed D Hashem: Data curation, Writing – review & editing
Timothy D Girard: Conceptualization, Methodology, Writing- review & editing
Daniel O Scharfstein: Conceptualization, Methodology, Writing – review & editing
Dale M Needham: Conceptualization, Methodology, Project administration, Funding acquisition, Writing – review & editing

5 CONTRIBUTORS
DMN obtained funding for the study and EC, TDG, DOS and DMN designed it. EC and XL set up the database. EC designed the code for the automated screening of the initial literature search. EC, XL, and MDH screened for eligibility and extracted data. EC and XL analysed and interpreted the data. EC and XL drafted the report. All authors critically reviewed the report for important intellectual content and approved the final submitted version.

Abstract

Objective: This structured methodology review evaluated statistical approaches used in RCTs enrolling patients at high risk of death and makes recommendations for reporting future RCTs.

Study design and setting: Using PubMed, we searched for RCTs published in five general medicine journals from January 2014 to August 2019 wherein mortality was ≥10% in at least one randomized group. We abstracted primary and secondary outcomes, statistical analysis methods, and patient samples evaluated (all randomized patients vs. "survivors only").

Results: Of 1947 RCTs identified, 434 met eligibility criteria. Of the eligible RCTs, 91 (21%) and 351 (81%) had a primary or secondary functional outcome, respectively, of which 36 (40%) and 263 (75%) evaluated treatment effects among "survivors only". In RCTs that analyzed all randomized patients, the most common methods included use of ordinal outcomes (e.g., modified Rankin Scale) or creating composite outcomes (primary: 41 of 91 [45%]; secondary: 57 of 351 [16%]).

Conclusion: In RCTs enrolling patients at high risk of death, statistical analyses of functional outcomes are frequently conducted among "survivors only," for which conclusions might be misleading. Given the growing number of RCTs conducted among patients hospitalized with COVID-19 and other critical illnesses, standards for reporting should be created.

Keywords

Randomized controlled trials; patient mortality; functional outcomes; "truncation due to death"; "survivors only" analysis; reporting guidelines

1 Introduction

Though mortality is often the primary outcome in randomized controlled trials (RCTs) enrolling patients that are at high risk of death, such as patients in an intensive care unit or with an advanced cancer, functional outcomes (e.g., cognition, physical disability and quality of life) are frequently evaluated as key secondary outcomes or, increasingly, as primary outcomes.[1,2] Within the subset of trial participants, however, who die before follow-up is completed, functional outcomes cannot be assessed; this "truncation due to death," as it is known in the statistical literature, complicates comparisons across randomized groups.[3,4,5] To adhere to the intention-to-treat principle, for example, all randomized participants should be analyzed, but functional outcomes are undefined for those who die prior to assessment. Additionally, a "survivors-only" comparison of functional outcomes may be biased if mortality differs across randomized groups such that measured or unmeasured baseline characteristics of the survivors for whom functional outcomes can be obtained are not balanced.[6]

Even though these issues have been highlighted in both the statistical and medical literature,[6,7,8] we hypothesize that the issue of "truncation due to death" is not widely considered during the design and analysis of RCTs that enroll patients at high risk of death, leading to uncertainty in conclusions regarding functional outcomes. Further, given the growing number of RCTs that enroll hospitalized patients with COVID-19 and other critical

illnesses whom experience high mortality,[9,10,11,12] it is urgent that reports of analyses and results regarding functional outcomes that may be "truncated due to death" be standardized. We therefore present findings of a structured methodology review of RCTs that enrolled patients at high risk of death published in 5 high-impact general medical journals. Our primary goals were to identify functional outcomes, to describe the statistical methodology applied when analyzing these functional outcomes, and to make recommendations for reporting.

2 Methods

2.1 Search Strategy and Selection Criteria

This structured methodology review was conducted in accordance with the PRISMA 2009 Checklist for reporting a systematic review, with one exception. We did not assess risk of bias since this assessment was not relevant to the objective of our review (i.e., we focused on the outcomes assessed and statistical methods used rather than on the findings/conclusions of the RCTs).[13]

A medical librarian searched PubMed to identify all RCTs published from January 1, 2014 through August 13, 2019 in the following five high-impact, general medicine journals: *Annals of Internal Medicine, British Medical Journal, Journal of the American Medical Association, Lancet,* and *New England Journal of Medicine.* These journals were selected based on both the breadth of patient populations evaluated in RCTs typically published in these journals and their requirement that authors adhere to the CONSORT 2010 guidelines when reporting the design, analysis and results of RCTs.[14] Search terms were used to limit results to the relevant dates, journals, and publication types specified, e.g., ("Lancet" [Journal]) AND ("2014"[Date-Publication]) AND (randomized controlled trial [pt] OR controlled clinical trial [pt])

Initially, the identified citations' titles and abstracts were screened for the presence of the words: "death", "mortality" or "survival", using the R statistical software (qrep function). Subsequently, two investigators (among EC, MDH, KA and XL) independently reviewed each selected citation's abstract to determine whether the following eligibility criteria were met: RCT, reported mortality ≥10% in at least one treatment group. This threshold for reported mortality was used to define patient populations at high risk of death.

From each eligible article (and their supplemental appendices, if provided), two investigators (among EC, MDH, KA and XL) independently extracted the following information: a) primary outcome, b) one secondary functional outcome, if available, and c) statistical analysis methods used to analyze the primary outcome and secondary functional outcome, if applicable. For RCTs with more than one secondary functional outcome, we selected the outcome with a clearly defined statistical analysis method that was first reported in the Methods section.

Throughout the process, discrepancies were resolved through consensus, with adjudication by EC if consensus could not be reached among the two reviewers.

2.2 Data Analysis

As described below, for each outcome, we recorded the type of variable, statistical analysis method for making comparisons across the randomized groups (e.g., two-sample t-test), and the sample population upon which the comparisons were made (e.g., all randomized patients or survivors).

We classified the type of variable into one of 5 categories: a) mortality: including overall survival, disease-specific survival, or landmark mortality status at a fixed point in time (e.g., alive at 6 months post-randomization); b) composite: mortality combined with another patient outcome, such as disease progression (e.g., progression-free survival) or hospitalization (e.g., hospital-free survival); c) continuous: an outcome or characteristic of the patient taking any value within a range of real numbers (e.g., distance walked within a 6-Minute Walk Test), d) ordinal: ranked categories of health states from worst to best (e.g., modified Rankin Scale), and e) binary: the presence vs. absence of a patient condition or event (e.g., occurrence of an adverse event).

For each primary and secondary outcome, we classified the statistical analysis method(s) used into 8 categories: a) survival: a time-to-event analysis (e.g., Log-rank test or Cox proportional hazards model), b) competing risk: the Fine and Gray competing risk survival model with death defined as the competing risk, c) parametric: parametric methods including two-sample t-tests, linear regression, linear mixed models, two-sample tests for proportions, or logistic regression models, d) non-parametric: non-parametric methods including Wilcoxon rank-sum test, Chi-square test), e) incidence rate: number of events divided by total exposure time for each randomized group, compared by Poisson test, f) worst-rank analysis: creation of a new outcome variable where patients who die are assigned a worst state to the continuous, ordinal or binary outcome with group comparisons made using a non-parametric method (e.g., for patients who die, assign a score of −1 for the 6-Minute Walk Test distance and use the Wilcoxon rank sum test),[15,16] g) survivor average causal effect: causal inference approach comparing the average functional outcome among patients who would have survived regardless of the intervention they receive,[17,18,19,20] or h) pattern mixture model: approach applied to functional outcomes measured repeatedly over the course of the RCT where the average functional outcomes measured over time are compared among subsets of patients with similar survival experiences.[21,22] We recorded a single statistical analysis method for each outcome, but we anticipated that multiple statistical analyses may be reported for an outcome. In such cases, we used the analysis reporting data on all randomized patients (e.g., we recorded the worst-rank analysis if the two-sample t-test was reported as the primary analysis and the worst-rank analysis was conducted as a sensitivity analysis).

When identifying the sample population upon which randomized group comparisons were made, we relied on the text of the statistical methods and results along with the accompanying tables and/or figures (reported in the main text or supplement).

The extracted data were summarized using counts and proportions. Statistical analyses were conducting using R.[23] The protocol for this structured review was registered in PROSPERO (CRD42018102656).[24]

2.3 Role of funding source

This work was supported through a grant from the National Heart, Lung and Blood Institute (NHLBI R24HL111895). The funder of the study had no role in study design, data collection, data analysis, data interpretation, or writing of the report. The corresponding author had full access to all the data in the study and had final responsibility for the decision to submit for publication.

3 Results

3.1 Data synthesis

Our search identified 1947 published RCTs, of which 773 (40%) were selected for inclusion based on the initial automated screen for the keywords: “death”, “mortality” or “survival” in the title or abstract (Figure 1, Table 1). Of these 773, 434 (22%) met eligibility criteria with 91 (20%) reporting a primary functional outcome and the remaining 343 (80%) reporting mortality or a composite outcome as the primary outcome. A secondary functional outcome was identified in 81% (n=351) of 434 eligible published RCTs.

All identified ordinal outcomes included mortality as a ranked health state. Specifically, the modified Rankin Scale and Glasgow Outcome Scale ranked death as the worst health state, and the EQ-5D utility score includes death as a numeric score of 0 and includes negative values to represent health states worse than death. One RCT reported estimating the survivor average causal effect and one study reported the use of a pattern mixture model.

The sample population used for statistical analysis of the functional outcomes was seldom explicitly stated such that we had to deduce this information from the statistical analysis methods, reported results, and/or supplemental appendices (e.g., the number of surviving patients was provided as the sample size in the results table accompanying the mean (SD) of a continuous functional outcome). Parametric and non-parametric analyses of continuous and binary variables were applied to only surviving patients; otherwise, all randomized patients were included in the analysis of the functional outcomes. Note that although the survivor average causal effect is comparing the average functional outcome among a subset of patients who would survive regardless of which randomized group they were assigned, we classified this approach as one that includes all randomized patients since to estimate the survivor average causal effect, a model for the probability of surviving is developed using data from all randomized patients.[18]

3.2 Primary functional outcomes

Among the 91 eligible RCTs with a primary functional outcome, 64 (70%) were continuous or binary outcomes, and the remaining 27 (30%) were ordinal outcomes (modified Rankin Scale: Glasgow Outcome Scale: 6, study-specific defined ordinal outcome: 2) (Table 2). Forty percent (n = 36) of the 91 primary functional outcomes were analyzed using data from only surviving patients. The remaining 55 (60%) of 91 primary functional outcomes were analyzed using data from all randomized patients; these included 27 ordinal outcomes. Among the 64 RCTs with continuous or binary primary functional outcomes, 14 (22%), 5 (8%) and 9 (14%) were analyzed using the worst-rank, incident type, and competing risk

approach, respectively, and the remaining were analyzed using parametric or non-parametric methods applied to only surviving patients (Supplemental Table).

3.3 Secondary Functional Outcomes

Among the 351 eligible RCTs for which we identified at least one functional secondary outcome, 76% (n = 267) had hypothesized a difference in mortality between treatment groups based on having either a primary outcome of mortality (35%, n = 124) or a composite primary outcome including mortality (41%, n = 143). Of the 351 identified secondary functional outcomes, 320 (91%) were continuous or binary. Regardless of the variable type, only 25% (88 of 351) of the secondary functional outcomes were analyzed using an approach that included all randomized patients; 31 of the 88 secondary functional outcomes analyzed in this way were ordinal. Of the 320 continuous and binary secondary functional outcomes, 24 (8%) were analyzed using the worst-rank analysis, 21 (7%) were analyzed as incident type, 10 (3%) were analyzed using a competing risk approach (e.g., time to improved quality of life with death as competing risk), 1 (<1%) was analyzed using the survivor average causal effect, and 1 (<1%) was analyzed using the pattern mixture model (Supplemental Table). The majority (82%) of the continuous and binary secondary functional outcomes were analyzed using a parametric or non-parametric method applied only to surviving patients.

4 Discussion

In this structured methodology review of 434 RCTs enrolling patients at high risk of death, published between January 1, 2014 and August 13, 2019 in 5 high impact general medical journals, we found that 76% of the RCTs had mortality or a composite endpoint (including mortality) as the primary outcome and 21% and 81% of the RCTs had a primary or secondary functional outcome, respectively. As previously reported,[8] there is no perfect solution for the analysis of functional outcomes "truncated due to death;" however, three key findings from this review support recommendations for the design and reporting of RCTs enrolling patients at high risk of death (Table 3).

First, the sample population used in the analysis of functional outcomes was seldom explicitly stated. We were able to deduce the sample population using the description of the statistical method and data presented in tables or supplementary appendices, an approach consistent with CONSORT 2010 guidance (see Guideline 16).[14] However, the potential challenges to interpreting results in these RCTs warrant more explicit reporting by study authors. More specifically, when reporting RCTs conducted among patients at high risk of death, authors should clearly state in the Methods section the population of participants included in the analysis of functional outcomes (Table 3).[25,26] This recommendation is a necessary part of comprehensive and transparent reporting of missing data and missing data analyses, both of which are currently part of the author instructions for *New England Journal of Medicine* and *Annals of Internal Medicine.*

Second, of the RCTs in our review that included a primary or secondary functional outcome, 40% and 75%, respectively, compared the functional outcome across randomized groups using data only from surviving patients. Findings from such comparisons based only on

survivors should be interpreted with caution as the results from the "survivors-only" analysis can only be interpreted as the causal effect of treatment on the functional outcome when the randomized groups have no effect on mortality. If the trial intervention does effect mortality, survivors assessed for functional outcomes may be systematically different across the randomized groups; such differential follow-up may bias results. Therefore, in addition to clearly defining the sample population used in the analysis, when reporting "survivors-only" analyses, authors should report on measured baseline characteristics across survivors in the randomized groups and describe limitations of the analysis in the Discussion section with explicit mention of how measured and unmeasured confounders may impact their conclusions.

Lastly, of the RCTs that included a primary or secondary functional outcome, 60% and 25%, respectively, compared the functional outcome across randomized groups using data from all randomized patients. These comparisons adhere to the intention-to-treat principle; however, the methodology available to make such comparisons require additional assumptions. The RCTs that selected an ordinal variable as the functional outcome or that applied the worst-rank analysis to a continuous or binary variable require the ability to rank mortality among a set of possible health states. The choice of where to rank mortality may differ based on the perspectives of patients, families, clinicians and researchers. Authors should describe and support their choice of ranking in the Methods section. Moreover, authors should describe in the Discussion section how their ranking may impact their findings. Alternatively, the survivor average causal effect compares the average functional outcome among a special subset of patients—those who would survive regardless of which group they are randomized to—which cannot be identified directly from the data. Assumptions are required to estimate the survivor average causal effect. For example, many of the proposed approaches for estimating the survivor average causal effect requires the monotonicity assumption, which states that there are no patients who would die when receiving the active treatment but survive when receiving the usual care/placebo.[17,18,19,20] Further, the effect applies to a subset of the population that is not readily identifiable at the time of treatment decision making. In the Methods section, authors should clearly state and justify the assumptions made to estimate the survivor average causal effect. In the Discussion section, they should discuss the implications of potential violations of these assumptions on their findings. For both statistical analysis approaches, sensitivity analyses for the required assumptions should be presented.[18, 25, 26]

Our structured methodology review has potential limitations. First, we limited the review to articles published in one of five high-impact general medical journals, a constraint that may impact the generalizability of our results. For instance, cancer trials, for which mortality is commonly the primary outcome with secondary outcomes often including quality of life and other patient-reported outcomes, represented 48% of the eligible trials in the review. If we had included subspecialty journals in our search, our findings may have differed. However, selection of the five high-impact general medicine journals was based, in part, on the broad range of topics evaluated and on these journals' mandatory reporting standards using the CONSORT 2010 guidelines. Hence, the use of statistical methods for addressing "truncation due to death" may be even less frequent in other medical journals that may not mandate such reporting. Second, it is possible that our literature search or screening process

did not identify some eligible RCTs published in the 5 target journals. If so, however, these omissions are unlikely to have caused a systematic bias to these results. Third, we only reported on the analysis of functional outcomes with respect to the statistical analysis and sample population based on the statistical methods and results reported for the RCTs, we did not evaluate whether the authors communicated their findings with appropriate qualifiers (e.g., on average, *among survivors,* the intervention improved the reported quality of life compared to the placebo). Nevertheless, though qualifiers may inform readers about the limitations of an analysis, they do not address the fact that that results may be biased if differential follow-up occurred. Lastly, comparing functional outcomes among survivors only *may*, but does not always, produce misleading findings. For example, in RCTs where the randomized intervention has no effect on mortality or follow-up rates, survivors generally represent a random sample of the randomized patients and comparisons across the survivors are expected to be unbiased. However, the effect of randomized interventions on mortality is unknown prior to conducting the trial and thus, potential biases should be considered when developing the statistical analysis plan and reporting approach.

In conclusion, functional outcomes are commonly reported in RCTs conducted in patients at high risk of death, including those with advanced cancers or critical illness. Comparisons of these functional outcomes between the RCTs' randomized groups are most often conducted among survivors only, an approach that may lead to misleading study results. We have provided recommendations for reporting of statistical analyses applied to functional outcomes "truncated due to death," which could help improve the conduct and interpretation of findings from trials conducted among patients with COVID-19 and other critical illnesses.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

We would like to thank Carrie Price, MLS for her expertise in developing and implementing the literature search.

7 Declaration of Interests

EC, XL and DMH received funding from a grant from the National Heart, Lung and Blood Institute (NHLBI R24HL111895). TDG, DOS and MDH declare no competing interests.

8 References

1. Fried TR, Bradley EH, Towle VR, Allore H. Understanding the treatment preferences of seriously ill patients. N Engl J Med. 200244; 346(14): 1061–6. [PubMed: 11932474]
2. Turnbull AE, Rabiee A, Davis WE, et al.Outcome measurement in ICU survivorship research from 1970-2013: a scoping review of 425 publications. Crit Care Med2016;44:1267–77. 10.1097/CCM.0000000000001651 [PubMed: 26992067]
3. Rubin DB. Discussion of "Causal Inference with Counterfactuals" by A.P. Dawid. Journal of the American Statistical Association. 2000; 95:435–437.
4. Rubin DB. Estimating Causal Effects of Treatments in Randomized and Nonrandomized Studies. J Educ Psychol1974;66:688–70110.1037/h0037350
5. Frangakis CE, Rubin DB. Principal stratification in causal inference. Biometrics2002;58:21–9. 10.1111/j.0006-341X.2002.00021.x [PubMed: 11890317]

6. Rosenbaum PRThe consequences of adjustment for a concomitant variable that has been affected by the treatment. Journal of the Royal Statistical Society, Series A. 1984;147: 656–66.
7. Kurland BF, Johnson LL, Egleston BL, Diehr PH. Longitudinal Data with Follow-up Truncated by Death: Match the Analysis Method to Research Aims. Stat Sci. 2009;24(2):211. doi:10.1214/09-STS293 [PubMed: 20119502]
8. Colantuoni E, Scharfstein DO, Wang C, et al.Statistical methods to compare functional outcomes in randomized controlled trials with high mortality. BMJ. 2018;360:j5748. Published 2018 Jan 3. doi:10.1136/bmj.j5748 [PubMed: 29298779]
9. Karagiannidis C, Mostert C, Hentschker C, et al.Case characteristics, resource use, and outcomes of 10 021 patients with COVID-19 admitted to 920 German hospitals: an observational study. Lancet Respir Med. 2020; (published online July 28.)
10. Yehia BR, Winegar A, Fogel R, et al.Association of Race With Mortality Among Patients Hospitalized With Coronavirus Disease 2019 (COVID-19) at 92 US Hospitals. JAMA Netw Open. 2020;3(8):e2018039. doi:10.1001/jamanetworkopen.2020.18039 [PubMed: 32809033]
11. Richardson S, Hirsch JS, Narasimhan M, et al.Presenting Characteristics, Comorbidities, and Outcomes Among 5700 Patients Hospitalized With COVID-19 in the New York City Area. JAMA. 2020;323(20):2052–2059. doi:10.1001/jama.2020.6775 [PubMed: 32320003]
12. Turnbull AE, Rabiee A, Davis WE, et al.Outcome Measurement in ICU Survivorship Research From 1970 to 2013: A Scoping Review of 425 Publications. Crit Care Med. 2016;44(7): 1267–1277. [PubMed: 26992067]
13. Moher D, Liberati A, Tetzlaff J, Altman DG, The PRISMA Group (2009). Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement. BMJ 2009;339:b2535, doi: 10.1136/bmj.b2535 [PubMed: 19622551]
14. Schulz KF, Altman DG, Moher D, for the CONSORT Group. CONSORT 2010 Statement: updated guidelines for reporting parallel group randomised trials. BMJ2010;340:c332. [PubMed: 20332509]
15. Diehr P, Patrick DL,Spertus J,Kiefe CI,McDonell M,Andfihn SD (2001). Transforming self-rated health and the SF-36 scales to include death and improve interpretability. Medical Care39, 670–680 [PubMed: 11458132]
16. Lachin JM. Worst-rank score analysis with informatively missing observations in clinical trials. Control Clin Trials1999;20:408–22. 10.1016/S0197-2456(99)00022-7 [PubMed: 10503801]
17. Zhang JL, Rubin DB. Estimation of causal effects via principal stratification when some outcomes are truncated by “death”. Journal of Educational and Behavioral Statistics. 2003; 28 (4): 353–368.
18. Hayden D, Pauler DK, Schoenfeld D. An estimator for treatment comparisons among survivors in randomized trials. Biometrics2005;61:305–10. 10.1111/j.0006-341X.2005.030227.x [PubMed: 15737107]
19. Tchetgen Tchetgen EJ. Identification and estimation of survivor average causal effects. Stat Med2014;33:3601–28. 10.1002/sim.6181 [PubMed: 24889022]
20. Chiba Y, VanderWeele TJ. A simple method for principal strata effects when the outcome has been truncated due to death. Am J Epidemiol2011; 173:745–51. 10.1093/aje/kwq418 [PubMed: 21354986]
21. Ribaudo HJ, Thompson SG, Allen-Mersh TG (2001). A joint analysis of quality of life and survival using a random effect selection model. Statistics in Medicine 20, 1173–1184 [PubMed: 11304734]
22. Pauler DK,McCoy S and Moinpour C (2003). Pattern mixture models for longitudinal quality of life studies in advanced stage disease. Statistics in Medicine 22, 795–809 [PubMed: 12587106]
23. R Core Team (2013). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL http://www.R-proiect.org/
24. Colantuoni EA systematic review of the statistical methods used in randomized controlled trials evaluating treatments on long-term outcomes in patient populations with a high risk of mortality. PROSPERO: International prospective register of systematic reviews. 2018. CRD42018102656. Available from: https://www.crd.york.ac.uk/prospero/#recordDetails.
25. Mehrotra DV, Hemmings RJ, Russek-Cohen E; ICH E9/R1 Expert Working Group. Seeking harmony: estimands and sensitivity analyses for confirmatory clinical trials. Clinical Trials. 2016; 13(4): 456–8. doi: 10.1177/1740774516633115 [PubMed: 26908545]

26. Akacha M, Bretz F, Ruberg S. Estimands in clinical trials - broadening the perspective. Stat Med2017;36:5–19. doi:doi:10.1002/sim.7033 [PubMed: 27435045]

Highlights

- Trials enrolling patients at high risk of death, including patients hospitalized with COVID-19 or other critical illnesses, evaluate functional patient outcomes
- Functional outcomes are “truncated due to death” making treatment comparisons challenging
- Authors seldom state whether patients who die are included in treatment comparisons
- Most often, analyses are conducted only on survivors, an approach that may be misleading.
- Recommendations are provided for reporting analyses conducted on functional outcomes “truncated due to death”

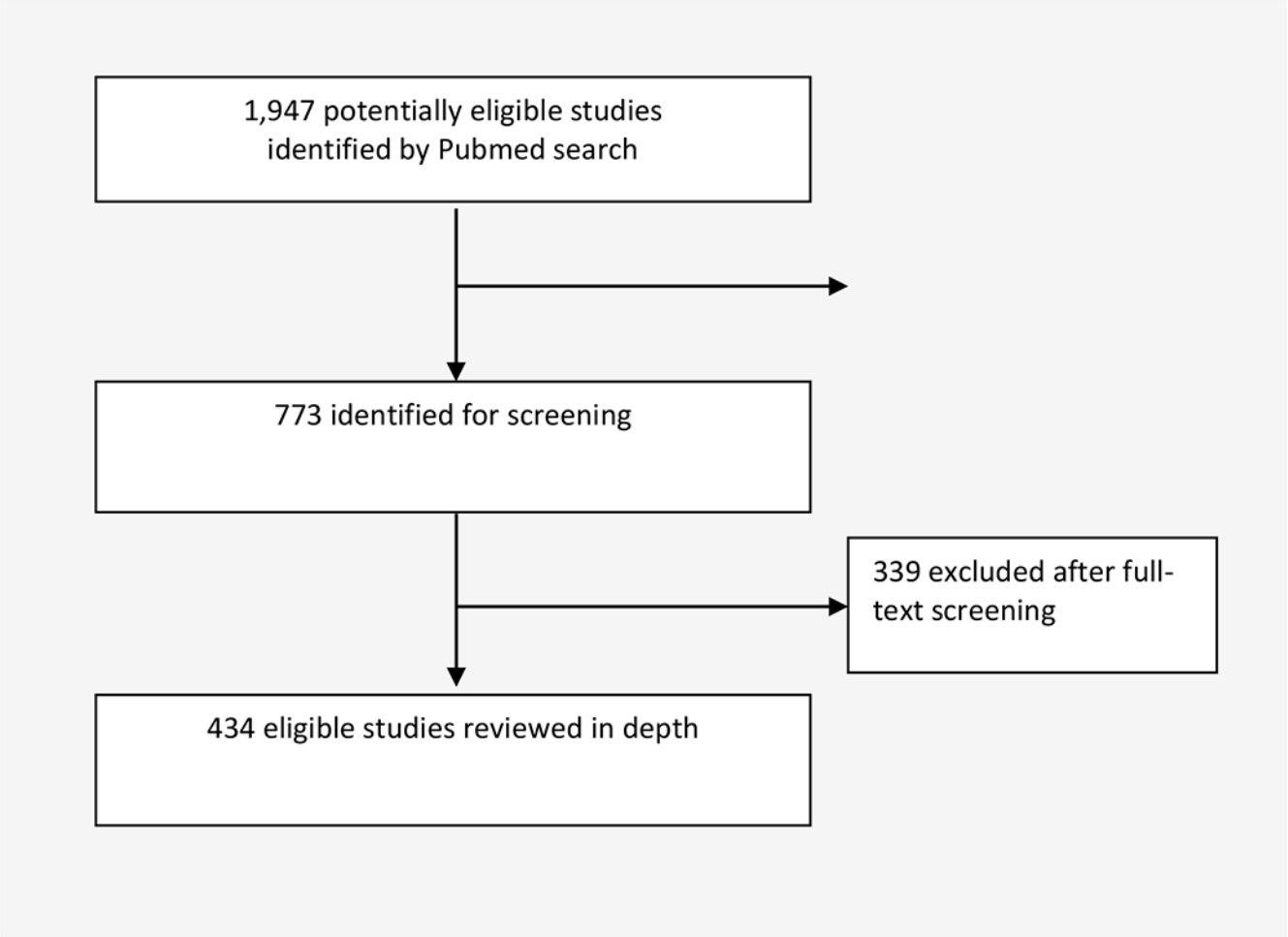

Figure 1:
Flow diagram

Table 1.

Summary of eligible randomized trials published in five general medicine journals during 2014 – 2019.

Journal Name	No. articles retrieved	No. (%) articles meeting eligibility criteria[1]	No. (%) of eligible articles with functional outcome(s)[2]			No. (%) of eligible articles with functional outcomes with comparisons based on only surviving patients[3]	
			Any	Primary	Secondary	Primary	Secondary
Annals of Internal Medicine	81	6 (7)	3 (50)	3 (50)	3 (50)	2 (67)	1 (33)
British Medical Journal	136	5 (4)	4 (80)	3 (60)	4 (80)	1 (33)	3 (75)
Journal of the American Medical Association	441	92 (21)	76 (83)	31 (34)	71 (77)	17 (55)	54 (76)
Lancet	515	102 (20)	86 (84)	22 (22)	84 (82)	9 (41)	64 (76)
New England Journal of Medicine	774	229 (30)	191 (83)	32 (14)	189 (83)	7 (22)	141 (75)
TOTAL	1947	434 (22)	360 (83)	91 (21)	351 (81)	36 (40)	263 (75)

[1] RCT, mortality reported for all treatment groups, and mortality ≥10% in at least one treatment group

[2] Functional outcomes include continuous, binary and ordinal outcomes

[3] Treatment comparisons of functional outcomes based only on surviving patients can be misleading if survivors differ in measured or unmeasured characteristics

Table 2.

Summary, count (%), of primary and secondary functional outcomes by variable type and statistical analysis method

Journal Name	Primary functional outcomes					Secondary functional outcome				
	No.	Variable type		Statistical analysis conducted among all randomized patients[1]		No.	Variable type		Statistical analysis conducted among all randomized patients[1]	
		Continuous or binary	Ordinal	All variable types	Continuous or binary[2]		Continuous or binary	Ordinal	All variable types	Continuous or binary[2]
Annals of Internal Medicine	3	3 (100)	0 (0)	1 (33)	1 (33)	3	3 (100)	0 (0)	2 (67)	2 (67)
British Medical Journal	3	3 (100)	0 (0)	2 (67)	2 (67)	4	3 (75)	1 (25)	1 (25)	0 (0)
Journal of the American Medical Association	31	26 (84)	5 (16)	14 (45)	9 (35)	71	64 (90)	7 (10)	17 (24)	10 (14)
Lancet	22	14 (64)	8 (36)	13 (59)	5 (36)	84	77 (92)	7 (8)	20 (24)	13 (17)
New England Journal of Medicine	32	18 (56)	12 (44)	23 (72)	11 (61)	189	173 (92)	16 (8)	48 (25)	32 (18)
TOTAL	91	64 (70)	27 (30)	55 (60)	28 (44)	351	320 (91)	31 (9)	88 (25)	57 (17)

[1] Statistical analyses conducted on all randomized patients include: a) ordinal outcomes with death as a possible health state analyzed using a non-parametric or parametric method, b) continuous or binary outcomes analyzed using the worst-rank analysis, competing risk analysis, pattern mixture model, or survivor average causal effect, or c) binary outcome analyzed using incidence type analysis.

[2] The number (%) of continuous or binary functional outcomes analyzed with a statistical method conducted on all randomized patients, denominator for the % is number of continuous or binary functional outcomes.

TABLE 3.

Recommendations for reporting statistical analyses applied to functional patient outcomes “truncated due to death”

Approach Used	Report in Methods	Report in Results	Acknowledge in Discussion
Survivors-only analysis	State that the analysis is conducted only among survivors	Across randomized groups, compare measured baseline characteristics among survivors whose functional outcomes were assessed	State that the results may be biased if measured or unmeasured confounders among survivors differ across randomized groups Describe how an unmeasured confounder may affect the results
Worst-rank analysis	State that the analysis is conducted among all randomized patients Define the ranking Provide justification of the ranking	Present results of sensitivity analyses for the ranking	State that the results may be sensitive to the ranking of death and the functional outcome Describe how changes in the ranking may affect the results
Survivor average causal effect	State that the target comparison group is the subset of patients that would survive regardless of the group to which they were randomized State and support with data (if possible) assumptions being made	Present results of sensitivity analyses for the assumptions	State that the results apply to only the subset of patients who would survive regardless of the group they were randomized to Describe whether violations of the assumptions are possible and the impact these violations may have on the results