ELSEVIER

Journal of Clinical Epidemiology 67 (2014) 410−415

Journal of Clinical Epidemiology

Subgroup analysis in randomized controlled trials appeared to be dependent on whether relative or absolute effect measures were used

Roderick P. Venekamp[a,b,*], Maroeska M. Rovers[b,c,d], Arno W. Hoes[b], Mirjam J. Knol[b]

[a]*Department of Otorhinolaryngology, University Medical Centre Utrecht, PO Box 85500, 3508 GA Utrecht, The Netherlands*
[b]*Julius Center for Health Sciences and Primary Care, University Medical Centre Utrecht, PO Box 85500, 3508 GA Utrecht, The Netherlands*
[c]*Department of Health Evidence, Radboud University Medical Center Nijmegen, PO Box 9101, 6500 HB Nijmegen, The Netherlands*
[d]*Department of Operating Rooms, Radboud University Medical Center Nijmegen, PO Box 9101, 6500 HB Nijmegen, The Netherlands*

Accepted 11 November 2013; Published online 5 February 2014

Abstract

Objectives: To assess whether relative or absolute effect measures were used in subgroup analyses of randomized controlled trials (RCTs) and study whether conclusions would change if subgroup effects were calculated on a different scale than reported.

Study Design and Setting: We studied all 327 RCTs published in 2010 in five major medical journals. For trials with a dichotomous primary outcome, we extracted reported main and subgroup effect measures. If crude subgrouping data were reported, we calculated the subgroup effects on both relative and absolute scales.

Results: Of the 229 RCTs with a dichotomous primary outcome, 120 (52%) performed subgroup analyses. In 106 of these 120 (88%) RCTs, relative effect measures were used for subgroup analyses, whereas an absolute scale was used in 9 (8%) trials. Two (2%) RCTs reported both relative and absolute subgroup effects. Crude data of the subgroups could be extracted in 41 of the 120 (34%) RCTs. Calculating subgroup effects on a different scale than reported lead to a change in conclusion in 17% of the 41 trials.

Conclusion: Almost all RCTs used relative effect measures for subgroup analyses. Interpretation of subgroup effects, however, appeared to be dependent on whether relative or absolute effect measures were used.

Keywords: Subgroup analysis; Randomized controlled trials; Treatment effects; Relative risk reduction; Absolute risk reduction; Epidemiology

1. Introduction

Randomized controlled trials (RCTs) are widely regarded as providing the most reliable evidence on the benefits and harms of interventions. In addition to main analyses, RCTs frequently perform subgroup analyses to identify specific subgroups of patients who do (or do not) benefit from the intervention [1−3]. Clinical guidelines often incorporate results of subgroup analyses, and such findings can therefore influence clinical decisions considerably.

Previous studies demonstrated that interpretation of trial results may be influenced by the use of either relative [eg, relative risk (RR), odds ratio (OR), hazard ratio (HR)] or absolute [eg, risk difference (RD)] effect measures in outcome reporting as benefits of interventions are often perceived larger if outcomes were reported with relative effect measures than if the same trial results were presented with absolute effect measures [4−8]. Consequently, reporting both relative and absolute effect measures for primary and secondary outcomes in RCTs is, nowadays, strongly recommended by the Consolidated Standards of Reporting Trials (CONSORT) statement [9]. Opposite to these explicit recommendations for the main analyses, the current CONSORT statement does not include clear recommendations on the use of specific effect measures for subgroup analyses. This, however, is remarkable as it has been acknowledged that subgroup analyses can lead to different results and conclusions with regard to statistical significance depending on whether relative or absolute effect measures are used [10]. To illustrate this phenomenon, we provide numerical examples based on RCTs performed by Dondorp et al. [11] (Appendix A at www.jclinepi.com) and Decousus et al. [12] (Appendix B at www.jclinepi.com).

As far as we are aware, no previous studies have been performed to investigate whether subgroup analyses are reported with relative or absolute effect measures and what the impact of such choices may be. We therefore systematically reviewed RCTs that were published in five major

Funding: None.
Conflicts of interest: None.
* Corresponding author. Tel.: +31-88-7555100; fax: +31-88-7568099.
E-mail address: R.P.Venekamp@umcutrecht.nl (R.P. Venekamp).

0895-4356/$ - see front matter
http://dx.doi.org/10.1016/j.jclinepi.2013.11.003

What is new?

Key findings

- Almost all randomized controlled trials used relative effect measures for subgroup analyses.
- Interpretation of subgroup effects appeared to be dependent on whether relative or absolute effect measures were used.

What this adds to what was known?

- These findings are highly important as previous studies demonstrated that benefits of interventions are often perceived larger if outcomes were reported with relative effect measures than if the same trial results were presented with absolute effect measures.

What is the implication and what should change now?

- Reporting of relative risk reduction should therefore always be accompanied by presenting the absolute risk reduction.
- The Consolidated Standards of Reporting Trials statement should incorporate such recommendations not only for primary and secondary outcomes but also for subgroup analyses.

general medical journals to assess whether relative or absolute effect measures were used in subgroup analyses and whether these subgroup effect measures differed from the main effect measures. We also studied whether conclusions would change if subgroup effects were calculated on a different scale than reported.

2. Methods

2.1. Selection of trials

We included all RCTs that were published in 2010 in five major general medical journals: *Annals of Internal Medicine* (*AIM*), *British Medical Journal* (*BMJ*), *Journal of the American Medical Association* (*JAMA*), Lancet, and *New England Journal of Medicine* (*NEJM*). These RCTs were retrieved using a search filter for PubMed that combined the journal names with publication date [pd] "2010" and publication type [pt] "randomized controlled trials" (Fig. 1). We included all RCTs irrespective of design (eg, parallel, factorial, crossover), study type (eg, superiority, equivalence, noninferiority), method of randomization, or sample size. Trials that were published online in 2010 but in article in 2011 were excluded. We also excluded research letters, cost-effectiveness analyses, diagnostic accuracy studies, studies that were not RCTs, and secondary analyses of RCTs.

2.2. Data extraction

We used a standardized data extraction form to assess the RCTs. This data extraction form was designed based on the five RCTs that were published in article in 2011. Two reviewers (R.P.V. and M.J.K.) independently extracted data from the included trials. Discrepancies between the reviewers were resolved by discussion. For trials with a dichotomous primary outcome, we extracted the reported effect measure for the main effect [RR, OR, HR, incidence rate ratio (IRR), RD, and incidence rate difference (IRD)], and determined whether results were statistically significant ($P \leq 0.05$). Additionally, we assessed whether these RCTs performed subgroup analysis by reviewing the methods and results sections (including tables and supplementary appendix) of these trials. If so, we investigated the number of subgroup analyses performed and whether relative or absolute effect measures (or both) were used. In addition, we assessed whether these trials used the appropriate statistical method to test whether treatment effect varies across the subgroup of interest, that is, whether tests for interaction were performed [2,13]. If possible, we extracted the crude data of the different subgroups to determine whether results and conclusions would change.

2.3. Sample size and data analysis

The decision to include all RCTs of 2010 was based on pragmatic considerations rather than formal sample size calculations. Frequencies and summary statistics of the extracted items were calculated. We used SPSS version 17 (SPSS Inc., Chicago, IL, USA) for these analyses.

If crude data of subgroups with two categories were reported, we calculated subgroup effects on both relative (ratio of RRs or ratio of IRRs across strata and 95% confidence interval (CI) and *P*-value) and absolute (difference of RDs or difference of IRDs across the subgroup strata and 95% CI and *P*-value) scales [14]. For further explanation of these calculations, see numerical example based on the study by Dondorp et al. [11] (Appendix A at www.jclinepi.com). For trials that used HRs as effect measure for the subgroup analyses and which reported only events and absolute numbers of patients across subgroups with two categories (ie, they did not report person-time of follow-up across the subgroups), we calculated the RR and the RD of both subgroup strata. Additionally, we calculated both the ratio of RRs and the difference of RDs across strata with their 95% CIs and *P*-values. For further explanation of these calculations, see numerical example based on the study by Decousus et al. [12] (Appendix B at www.jclinepi.com). For subgroups with more than two categories, we used Rothman Episheet version June 11, 2008 (http://www.drugepi.org/dope-downloads/#Episheet) to derive the