Journal of Clinical Epidemiology 105 (2019) 80−91

Journal of Clinical Epidemiology

ORIGINAL ARTICLE

No differences were found between effect estimates from conventional and registry-based randomized controlled trials

Tim Mathes*, Pauline Klaßen, Dawid Pieper

Institute for Research in Operative Medicine, Chair of Surgical Research, Faculty of Health, School of Medicine, Witten/Herdecke University, Ostmerheimer Street 200, 51109 Cologne, Germany

Accepted 10 September 2018; Published online 23 September 2018

Abstract

Objectives: The study aims to assess whether the results from registry-based randomized controlled trials (RRCTs) systematically differ from the results of conventional randomized controlled trials (CRCTs).

Study Design: The study was meta-epidemiological study. We identified RRCTs (February 2016) and subsequent systematic reviews (SRs) that included one RRCT (04/2017). We calculated pooled odds ratios for RRCTs and CRCTs for mortality and other incidence measures (e.g., cardiovascular events). We assessed the agreement between RRCTs and CRCTs using various measures with descriptive statistics and the odds ratio of pooled odds ratios (OROR) for RRCT(s) vs. CRCTs. An OROR of >1 indicates that the effect estimates from RRCTs were larger than the effects estimates from CRCTs.

Results: Overall, we compared 15 and 14 effect estimates for mortality and incidence measures, respectively. The 95% confidence interval (CI) overlap was 100% for any outcome. Conflicting effect directions were observed in 27% (4/15) of mortality and 7% (1/14) of incidence measures. The ORORs for RRCT(s) vs. CRCTs was 1.03 (95% CI 0.97−1.09) for mortality and 1.05 (95% CI 0.98−1.12) for other incidence measures.

Conclusions: Our analysis indicates that for objective outcomes, there is no systematic difference between effect estimates from RRCTs and CRCTs.

Keywords: Randomized controlled trials; Health registries; Patient registries; Study design

1. Introduction

Randomization is the only way to control unmeasured confounding, and consequently, a randomized controlled trial (RCT) is the only study design that allows attributing strong causation to an intervention. However, the low external validity of RCTs has often been criticized [1].

In registries, standardized patient data are routinely collected, but usually, the design of studies that are based on registry data is observational and thus prone to unmeasured confounding [2]. Registry-based RCTs (RRCTs) are RCTs using the data or infrastructure of a pre-existing registry. RRCTs have the potential to combine the advantages of randomization (balance of unmeasured confounding) with the assumed advantages of registries (higher external validity) and enable long-term follow-ups of large sample sizes for relatively little effort [3]. RRCTs are especially a promising future technology because the number of registries is increasing [4]. However, RRCTs also have new methodological difficulties [5]. One of the most important questions to be answered is whether registry data are of high enough quality to provide robust results [6,7].

Studies showed that the agreement of outcome information between different data sources might differ [8−10]. RRCTs are not at risk of unmeasured confounding, but there is the fear that results from RRCTs and conventional RCTs (CRCTs) differ systematically because of lower data quality in the registry data compared with actively collected data.

The main potential sources that might cause a systematic difference in effect estimates from RRCTs and CRCTs are missing outcome values (e.g., drop-outs) and misclassifications (e.g., wrong cause of death) that systematically differ between study arms. In large RCTs, patient groups can be expected to be homogeneous at baseline. Thus, the baseline risk for missing values and misclassifications is similar and can be assumed to be completely missing at random. Therefore, if analyzed according to the intention-to-treat principle, it could be assumed that the effect estimates are

* Corresponding author. Tel. +49 221x98957 43; fax: +49 221 98957 30.

E-mail address: Tim.Mathes@uni-wh.de (T. Mathes).

https://doi.org/10.1016/j.jclinepi.2018.09.011

What is new?

Key findings

- In our meta-epidemiological study, the effect estimates from registry-based randomized controlled trials and conventional randomized controlled trials were similar, and we found no systematic difference.

What this adds to what was known?

- This is the first study to indicate that the effect estimates from registry-based randomized controlled trials and conventional randomized controlled trials do not differ.

What is the implication and what should change now?

- Gathering objective outcomes for randomized controlled trials from registries should be considered more often, especially when designing pragmatic randomized controlled trials.

unbiased in CRCTs as well as RRCTs and consequently are not systematically different. However, in RCTs, differential missing values and misclassifications related to the study arms (e.g., small trials with unbalanced confounders, missing data related to the intervention or follow-up care) might occur. It is proposed that the impact of differential missing values, that is, data not missing (completely) at random but are related to the study arm, and misclassifications might be higher if the data quality is lower (e.g., fewer attempts are made to avoid missing values or misclassifications). Consequently, if the data quality of registries and actively collected data systematically differ, the effect estimates might also differ systematically.

A previous study found higher effect estimates in a propensity score analysis based on routinely collected health data than in subsequent RCTs. This difference might be caused by unmeasured confounding as well as the low quality of routine data and their interaction (e.g., confounders also influence drop-out) [11].

The objective of this meta-epidemiological study was to assess whether the results from RRCTs systematically differ from the results of CRCTs. For this purpose, we compared the effect estimates of RRCTs with the effect estimates of CRCTs for the same research question. We compared the same study design, such that we considered whether there is an isolated effect of different data sources on the results.

2. Material and methods

There was no a priori protocol for this study.

2.1. Literature search

We systematically searched all PubMed databases for RRCTs combining search terms for RCTs and registries in February 2016 [3,12]. In brief, we considered all studies that used a random allocation for the intervention that gathered at least one outcome from a registry. We considered two types of registries, patient registries, and registries in a national health information system (for definitions of registries applied for study selection, see Supplement 1). We did not apply any exclusion criteria for the population, intervention, or control.

The search for RRCTs can be found in Supplement 2 [3].

For each RRCT, we searched systematic reviews (SRs) on the same research question (same patients, intervention, comparison [PIC]). We used two approaches to identify SRs (April 2017). First, we searched for Cochrane SRs in the Cochrane database of SRs via the Cochrane Library. We searched for each research question separately using search terms for the intervention and synonyms (e.g., thrombolysis OR fibrinolysis) in title/abstract/keywords. Second, we searched Google scholar via the "cited by" function for non-Cochrane SRs using a standardized procedure (see Supplement 1). We obtained full-text articles for all potentially relevant SRs.

We applied the following inclusion criteria to all SRs:

- Performed a meta-analysis, which included an outcome of at least one of the previously identified RRCTs and at least one CRCT and
- Provided sufficient data for recalculation of the meta-analysis

The data collection method is usually not clearly labeled as registry-based or conventional. Therefore, we checked the method section of each RRCT and each SR to assess whether the data source used for the outcome assessment in RCTs satisfies our definition of a registry, and the trial can thus be classified as an RRCT.

We compared effect estimates for all-cause mortality and other incidental findings (e.g., cancer diagnosis, cardiovascular events) because these are known to be the outcomes that are most prevalent in registries [7]. We excluded studies on organization and coordination of care (e.g., quality improvement interventions) because such interventions are usually too clinically heterogeneous to be pooled in a meta-analysis. We included non-Cochrane SRs only if we could not identify a Cochrane SR or if the literature search of the non-Cochrane SR included additional new studies (i.e., studies published after the last search of the Cochrane SR). We preferred Cochrane SR because these are usually of higher methodological quality than non-Cochrane SRs [13,14]. If we identified a non-Cochrane SR that included new studies as well as a Cochrane SR on the same research question, we included both in the primary analysis.

Table 1. Characteristics of included systematic reviews

Comparison label	Patients	Intervention/control	Outcome
Breast cancer screening (Cochrane) [30]	Women	Screening for breast cancer with mammography/no screening	Mortality[c]
Colorectal cancer screening (Cochrane) [25]	Adults	Colorectal cancer screening/no screening/no screening	Mortality
Exercise-based cardiac rehabilitation (Cochrane) [16]	Myocardial infarction	Cardiac rehabilitation/usual care	Mortality
			Myocardial infarction
			Hospital admission
LCPUFA supplementation (Cochrane) [24]	Pregnant women	LCPUFA in pregnancy	Incidence of any allergy in children
Lung cancer screening [28]	Asymptomatic adults	Lung cancer screening/no screening	Mortality
Off-pump (Cochrane) [19]	Ischemic heart disease treated with CBAG	Off-pump/on-pump CBAG	Mortality
			Myocardial infarction
Off-pump (Dieberg 2016) [20]			Mortality
			Myocardial infarction
PCI [23]	ST-Segment–Elevation Myocardial Infarction	PCI/fibrinolytic	Mortality
Prostate cancer screening (Cochrane) [26]	Men	Prostate cancer screening/no screening	Mortality
			Prostate cancer diagnosis
Prostate cancer screening (Djulbegovic) [27]			Prostate cancer diagnosis
Radial access [32]	Women	Radial access/femoral access	Mortality
Radial access [32]	Women	Radial access/femoral access	Incidence of bleeding
Remote ischemic conditioning [31]	Risk of myocardial injury	Remote conditioning/no conditioning	Mortality
Remote ischemic conditioning [31]	Risk of myocardial injury	Remote conditioning/no conditioning	Cardiovascular and cerebral events
Routine invasive strategy (Cochrane) [21]	Non-ST-elevation acute coronary syndromes	Routine invasive strategy/selective invasive strategy	Mortality
Routine invasive strategy (Cochrane) [21]	Non-ST-elevation acute coronary syndromes	Routine invasive strategy/selective invasive strategy	Myocardial infarction
Routine invasive strategy (Elgendy) [22]			Mortality
Routine invasive strategy (Elgendy) [22]	Adults without cardiovascular diseases	Statins/placebo or usual care	Myocardial infarction
Statin for prevention of cardiovascular disease (Cochrane) [17]			Mortality
Statin for prevention of cardiovascular disease (Cochrane) [17]	Adults without cardiovascular diseases	Statins/placebo or usual care	Cardiovascular events
Statin for prevention of cardiovascular disease (USPSTF) [18]			Cancer incidence
			Cardiovascular Mortality
			Mortality
Thrombus Aspiration [29]	Patients undergoing primary PCI	Routine aspiration thrombectomy and conventional PCI	Mortality
Thrombus Aspiration [29]	Patients undergoing primary PCI	Routine aspiration thrombectomy and conventional PCI	Major adverse cardiac events

Abbreviations: AAA, abdominal aortic aneurysm; CBAG, coronary artery bypass grafting; CG, control group; IG, intervention group; LCPUFA, *n*-3 long-chain polyunsaturated fatty acids; PCI, percutaneous coronary intervention.

[a] Number includes registry-based randomized controlled trial.

[b] In original meta-analysis.

[c] No meta-analysis of whole population (only for subgroups), data from own meta-analysis.

[d] Weights from own meta-analysis (not reported in primary study).

Registry	Studies in meta-analysis[a]	Number of patients in meta-analysis (IG/CG)	Number of patients from RRCT (IG/CG)	Study weight[b]
Canadian Cancer Registry, Canadian National Mortality Database (Canada a and b) Swedish Cancer Registers, Swedish Cause of Death Registry (Malmö I, Göteborg, Östergöland, Kopparberg 1977) NHS Central Register (UK age trial)	8	248,646/289,180	25,214/25,216 (Canada a) 19,711/19,694 (Canada b) 21,088/21,195 (Malmö I) 53,884/106,956 (UK age trial) 21,650/29,961 (Göteborg) 38,942/37,675 (Östergöland) 38,568/18,479 (Kopparberg)	5.7%[d] (Canada a) 8.1%[d] (Canada b) 14.5%[d] (Malmö I) 11.5%[d] (UK age trial) 12.8%[d] (Göteborg) 17.3%[d] (Östergöland) 16.2%[d] (Kopparberg)
Funen Patient Database, National Board of Health Registry, National Cancer Register, as well as the Danish National Register of Patients (Funen) Register at the Department of Pathology and the National Cancer Register (Göteborg) NHS Central Register (Nottingham)	4	172,734/156,908	30,967/30,966 (Funen) 34,144/34,164 (Göteborg) 76,466/76,384 (Nottingham)	27.4% (Funen) 20.6% (Göteborg) 37.7% (Nottingham)
Civil Registration System, National Patient Registry	13	3,495/3,328	227/219	5.2%
	11	2,877/2,767	227/219	5.0%
	6	984/932	227/219	31.9%
National Patient Registry	3	891/874	263/265	6.1%
Danish Lung Cancer Registry, Danish Cancer Registry, Danish Causes of Death Registry, National Patient Registry	3	4,518/4,971	2,052/2,052	36.4%[d]
Danish National Patient Registry	74	2,485/2,465	450/450 (Doors) 176/163 (BBS)	4.2% (Doors) 21.7% (BBS)
Danish National Patient Registry	54	2,485/2,465	450/450 (Doors) 176/163 (BBS)	23.1% (Doors) 13.3% (BBS)
Danish National Patient Registry	43	7,754/7,776	450/450 (Doors) 176/163 (BBS)	4.6% (Doors) 6.4% (BBS)
Danish National Patient Registry	34	6,036/6,082	450/450 (Doors) 176/163 (BBS)	8.4% (Doors) 5.4% (BBS)
Danish Central Person Register (DANAMI) Merged data from National Patient Registry, National Cause of Death Registry, Swedish Coronary Angiography and Angioplasty Registry National Registry of Thoracic Surgery (SWEDES)	24	4,068/4,072	223/222 (DANAMI II) 101/104 (SWEDES)	4.3%[d] (DANAMI II) 1.4%[d] (SWEDES)
Swedish Cancer Register, National Cause of Death Register (Stockholm) National registries (ERSPC) Regional prostate cancer register (Norrkoping)	4	125,024/169,832	82,816/99,183 (ERSPC) 1,494/7,532 (Norrkoping) 2,374/24,772 (Stockholm)	41.9% (ERSPC) 1.8% (Norrkoping) 24.5% (Stockholm)
Swedish Cancer Register, National Cause of Death Register (Stockholm) National registries (ERSPC) Regional Prostate Cancer Register (Norrkoping)	4	125,024/169,832	82,816/99,183 (ERSPC) 1,494/7,532 (Norrkoping) 2,374/24,772 (Stockholm)	26.9% (ERSPC) 21.4% (Norrkoping) 24.9% (Stockholm)
National registries (ERSPC) Regional Prostate Cancer Register (Norrkoping) Swedish Regional Cancer Registry West Regional Cancer Registries, cause of death certificate (Gothenburg)	5	159,372/181,428	72,890/89,353 (ERSPC) 1,494/7,532 (Norrkoping) 4,055/4,002 (Gothenburg)	22% (ERSPC) 21% (Norrkoping) 19% (Gothenburg)
CathPCI Registry	9	6,653/6,319	104/116	5.0%[d]
CathPCI Registry	8	6,578/6,241	299/338	1.8%[d]
UK National Registry for deaths (Davies) Danish Registry of Causes of Death, Danish National Registry of Patient (Sloth)	3	383/389	126/125 (Botker) 95/97 (Hoole)	48.2% (Botker) 16.3% (Hoole)
	3	383/389	126/125 (Botker) 95/97 (Hoole)	38.1% (Botker) 37.0% (Hoole)
Swedish, Danish, and Norwegian National Registries	8	4,545/4,370	1,222/1,235	18.0%
Swedish, Danish, and Norwegian National Registries	8	4,545/4,370	1,222/1,235	25.7%
Swedish, Danish, and Norwegian National registries	12	4,916/4,734	1,222/1,235	27.0%
Swedish, Danish, and Norwegian National registries	12	4,916/4,734	1,222/1,235	34.3%
Scotland death registry	13	24,408/23,652	3,302/3,293	50.6%
Database of routine data of the National Health System of Scotland records; Database of psychiatric hospital admissions and cardiac surgeries	14	24,217/23,832	3,302/3,293	45.7%
Database of routine data of the National Health System of Scotland records	11	19,789/18,950	3,291/3,293	36.9%
Scotland death registry	10	23,019/23,075	3,302/3,293	67.0%
Scotland death registry	15	35,967/35,164	3,302/3,293	14.6%
Swedish national health registry (SCAAR)	18	10,277/10,310	3,621/3,623	41.4%
Swedish national health registry (SCAAR)	17	10,256/10,290	3,621/3,623	26.8%

All searches on matching SRs and study selection were performed by one reviewer (T.M.).

2.2. Data collection

We extracted data on patients, intervention, comparison, outcome (PICO) the name of the used registry, the number of studies included in the meta-analysis, and the number of patients included in the RRCTs/CRCTs in standardized pilot tables. Furthermore, we extracted the weights of the RRCT in the original meta-analysis to assess the influence of the RRCTs on the pooled effect estimate. In the case multiple follow-up time points were reported, we extracted data only for the last follow-up time point.

We extracted outcome data (events and number of patients per group) directly into the analysis software (RevMan). If possible, all data were extracted from the SRs. If the necessary data of the individual RRCT was not provided in the SR (e.g., events per group), we extracted data from the primary studies that were referenced in the SR. Furthermore, we extracted the number of included patients (RRCT and meta-analysis) and number of studies included in the meta-analysis.

Data were extracted by one reviewer (T.M.), and a second reviewer verified the correctness of entries (P.K.). Discrepancies were discussed until consensus.

2.3. Analysis

We compared the effects of CRCTs with the effects of the RRCT(s). We expressed all effect estimates as intervention vs. comparator. For harmonization of the effect direction, we defined the comparator as no treatment, placebo, or usual care/established treatment. It was not necessary to switch groups for this task because we only considered negative outcomes, and all identified SRs reported meta-analysis results in this direction (lower ORs indicate an effect in favor of the intervention). For this task, we calculated the pooled odds ratios (ORs) of all CRCTs, excluding RRCTs as well as ORs for each RRCT. If there was more than one RRCT, we also performed a meta-analysis for the RRCTs, excluding data from the CRCT. We used the DerSimonian and Laird method for the meta-analysis of each comparison because this is the standard random effects meta-analytic approach.

We used different measures to compare the effect estimates from RRCTs and CRCTs. We counted the number of conflicting effect directions for pairs of meta-analyses, the number of nonoverlapping 95% confidence intervals (CIs), and the number of the CRCT point estimators for ORs not included in the 95% CI of the RRCTs.

For the calculation of all 95% CIs of the metadata, we estimated the standard errors from the 95% CIs of the meta-analysis for each research question.

We compared the number of observed 90% CI overlaps with the number of expected 90% CI overlaps assuming no difference between RRCTs and CRCTs exists (i.e., expected OR of RRCT vs. CRCT = 1). We used 90% CIs for this task because 90% CIs are narrower than 95% CIs and consequently, when using 90% CIs, the results are more conservative [15]. For the observed overlap of 90% CIs, we calculated the 95% CIs and assessed whether the 95% CIs include the expected 90% CI overlap. If the expected 90% CI overlap is not included in the 95% CIs of the observed 90% overlap, this would indicate a systematic difference between RRCT and CRCT.

In addition to the count data, we calculated the difference in the log ORs by subtracting the log OR of the RRCTs from the log OR of the CRCTs. We performed a meta-analysis of the difference of log ORs (pooled difference of pooled ORs) with an inverse variance (Paule and Mandel heterogeneity estimator) random effects model. We performed the meta-analysis of differences of log ORs for all included meta-analyses (Cochrane and non-Cochrane), such that for some research questions, two pairs of meta-analyses were included (Cochrane SR and non-Cochrane SR, which included new studies). Therefore, for the sensitivity analysis, we repeated the calculation of the pooled difference of log ORs excluding the Cochrane SR for the same research question (i.e., only the non-Cochrane SR was included for this research question) because otherwise, there is an overlap of included studies/patients, and consequently, this research question is overweighted (more than one comparison for a research question) in the pooled difference of the log OR. For the pooled log ORs, we calculated 95% CIs.

A meta-analysis of each comparison was performed with Review Manager (RevMan) Version 5.3 (The Nordic Cochrane Centre, The Cochrane Collaboration, 2014). All other analyses were performed with the R packages ‘‘meta’’ and ‘‘forest plots.’’ The R-Code is available from the first author upon request.

3. Results

3.1. Study selection

We identified 71 RRCTs. The study selection of the RRCTs is presented elsewhere [3]. In total, we identified 17 SRs that included at least one of these RRCTs [16−32]. These SRs were on 12 different PIC and included 30 RRCTs in total. For mortality, there were eight Cochrane SRs and 7 non-Cochrane SRs (15 comparisons in total) [16−23,25,26,28−32]. Of this 15 SRs, three (Cochrane SR and non-Cochrane SR, respectively) had exactly the same PICO. Consequently, we analyzed 12 different PIC.

For incidence measures, we found six Cochrane SRs and six non-Cochrane SRs [16,17,19−22,26,27,29−32]. Of this 12 SRs, two (Cochrane SR and non-Cochrane SR, respectively) had exactly the same PICO, and two Cochrane SRs reported more than one incidence measure (e.g., cancer incidence and cardiovascular events). Consequently, we analyzed 12 different PICO, 10 different PIC, and 14 comparisons in total.

3.2. Study characteristics

The description of PICO for each included SR and description of the meta-analysis are presented in Table 1. All SRs were about cancer (5/17 [30%]) or heart diseases (12/17 [70%]). Approximately, half the SRs considered a preventive intervention (7/17 [41%]), and most of these were on screening (4/17 [24%]).

In total, the meta-analyses on mortality included 1,012,187 participants from RRCTs and 369,197 participants from CRCTs. Standard deviations of RRCTs and CIs of the pooled ORs were mostly larger for RRCTs than CRCTs (see Figs. 1 and 2). The RRCT(s) had a median weight in the meta-analysis of 27% (range 5.0–86.1%). Six comparisons included more than one RRCT.

The meta-analyses of incidence measures included 512,593 participants from RRCTs and 224,821 participants from CRCTs. Standard deviations of RRCTs and CIs of the pooled ORs were mostly larger for RRCTs than CRCTs (see Figs. 3 and 4). The median weight of RRCTs in the meta-analysis was 31.3% (range 1.8–75.1%).

3.3. Comparison of effects from RRCTs and CRCTs

3.3.1. Mortality

The (pooled) effect estimates of RRCTs and CRCTs for mortality for each comparison are presented in Figure 3. The 95% CIs overlapped for all comparisons (100%). The point estimator for the ORs of CRCTs was in 87% (13/15) of comparisons included in the 95% CIs of RRCTs.

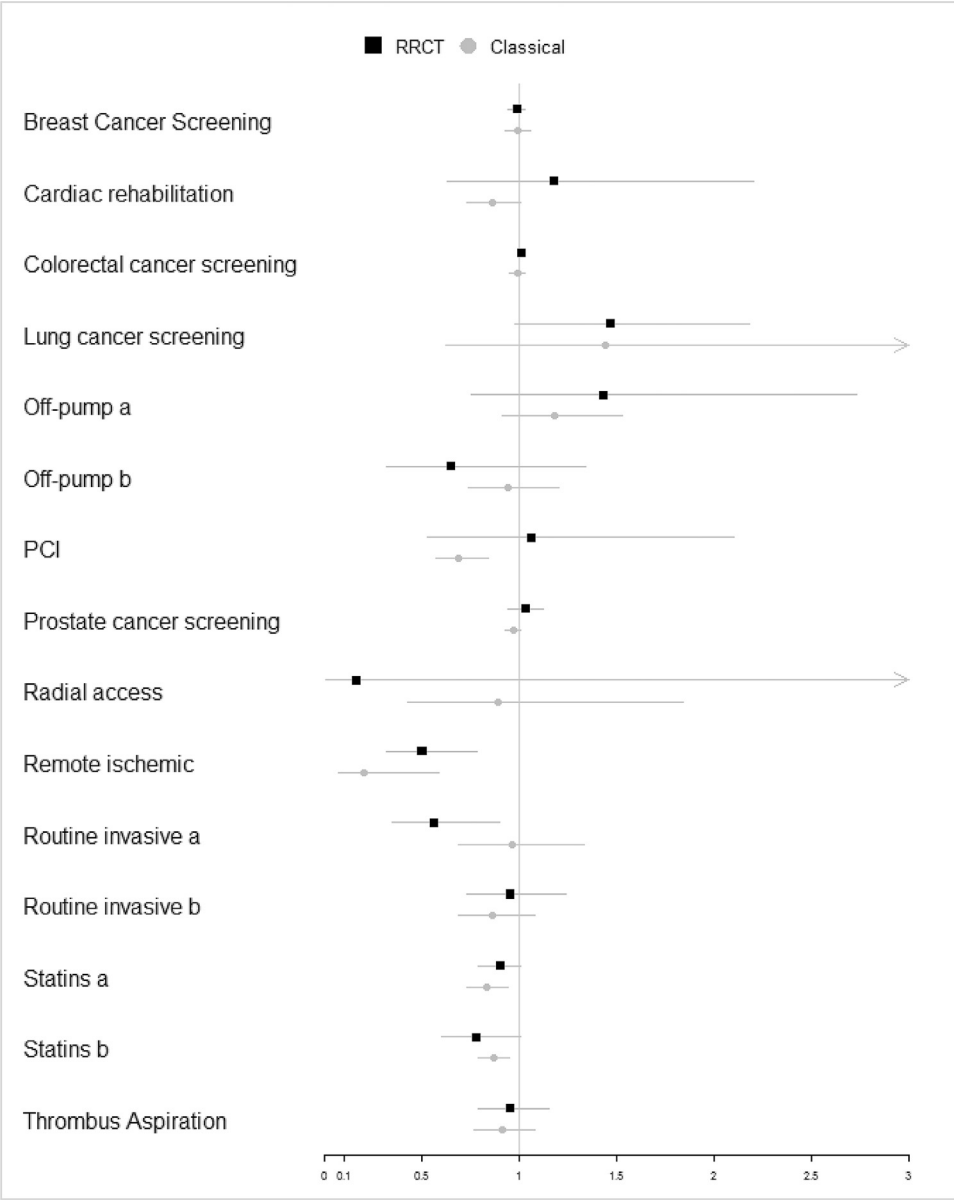

Fig. 1. RRCT vs. CRCT (odds ratio with 95% CIs for mortality). ‘‘a’’ Indicates a Cochrane SR, and ‘‘b’’ indicates a non-Cochrane SR.

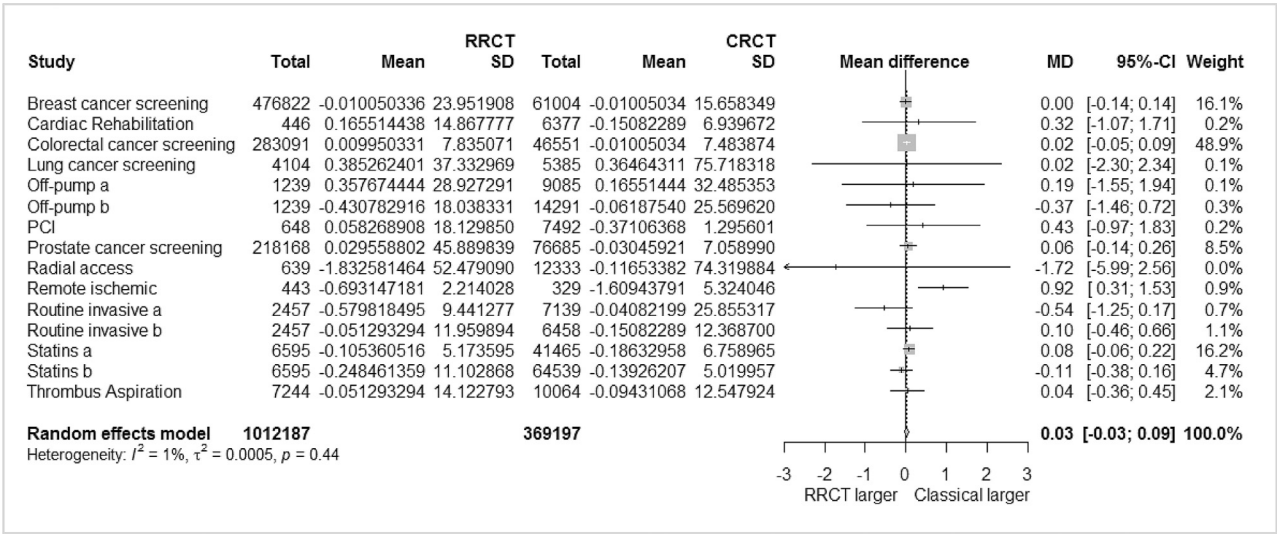

Fig. 2. Meta-analysis of differences between RRCTs and CRCTs (difference in log odds ratios with 95% CIs for mortality). ''a'' Indicates a Cochrane SR, and ''b'' indicates a non-Cochrane SR.

The point estimator for the ORs of RRCTs was in 53% (8/15) of comparisons included in the 95% CIs of CRCTs. A statistically significant effect was found in 26% (4/15) of comparisons for CRCTs and 13% (2/15) of RRCTs.

There were four conflicting effect directions. The 95% CI of the observed 90% CI overlap was 78–100%, which included the expected overlap of the 90% CIs (96%), indicating no difference between RRCTs and CRCTs.

Figure 2 shows the meta-analysis of the differences between RRCTs and CRCTS, including Cochrane SRs and non-Cochrane SRs (primary analysis) for mortality. The pooled difference in log ORs was 0.03 (95% CI −0.03 to 0.09, I^2:1%). This corresponds to a small OR of 1.03 (95% CI 0.97–1.09). The 95% CIs indicated no significant differences between RRCTs and CRCTs.

Results of the meta-analysis, including 12 comparisons, of only the most up-to-date SRs (sensitivity analysis), were similar (log OR 0.05; 95% CI −0.05 to 0.16; I^2: 12%).

3.3.2. *Incidence measures*

The (pooled) effect estimates of RRCTs and CRCTs for incidence measures for each research question are presented in Figure 1. In addition, the 95% CIs for incidence measures overlapped for all comparisons (100%). The point estimators for the OR of CRCTs were in 93% (13/14) of comparisons included in the 95% CIs of RRCTs. The estimators for the OR of RRCTs were in 64% (9/14) of comparisons included in the 95% CIs of CRCTs. A significant effect was found in 42% (6/14) and 36% (5/14) of comparisons for RRCTs and CRCTs, respectively. In six meta-analyses, more than one RCT was included.

There was one conflicting effect direction. The 95% CI of the observed 90% CI overlap was 77–100%, which included the expected overlap of the 90% CIs (96%), indicating no difference between RRCTs and CRCTs.

Figure 4 shows the meta-analysis of the differences between RRCTs and CRCTS, including the Cochrane SRs and non-Cochrane SRs (primary analysis) for incidence measures. The pooled difference in log ORs was 0.05 (95% CI −0.02 to 0.11, I^2: 0%). This corresponds to an OR of 1.05 (95% CI 0.98–1.12). The 95% CIs again indicated no significant differences between RRCTs and CRCTs.

Results for the pooled log OR, including 11 comparisons, of only the most up-to-date SRs (sensitivity analysis) were similar (log OR 0.05; 95% CI −0.02; 0.12; I^2: 0%).

4. Discussion

Our analysis provides no indication that the results from RRCTs and CRCTs differ systematically. We found no difference in the pooled ORs for either mortality or incidence measures. The OR of ORs for RRCTs vs. CRCTs was very small, and the 95% CI overlapped 1 (no effect). In addition, by comparing the expected and observed 90% CI overlap, we found no indication of a systematic difference for either outcome.

One explanation for our observation that we found no systematic difference between RRCTs and CRCTS might be that the factors which potentially influence missing data and misclassification are equally distributed in the intervention and control groups at baseline in large RCTs. If missing values and misclassifications are (completely) at random, no bias of effect estimates exists, but missing values and misclassifications have an effect only on precision. However, in RCTs, there remains a risk of differential missing data and misclassifications. First, in RCT baselines, an imbalance of confounders (e.g., age) may exist that can increase the risk for missing values (e.g., higher drop-out

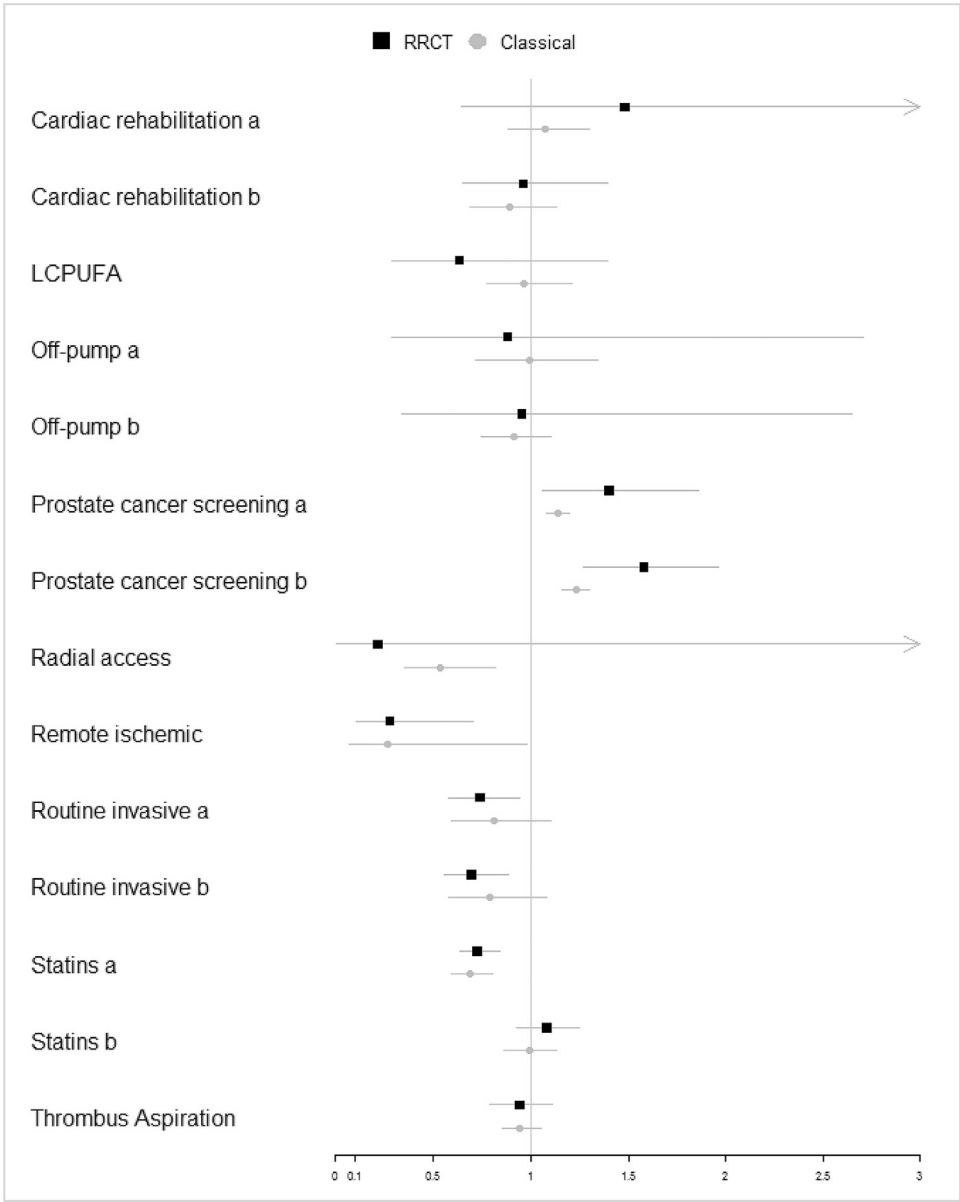

Fig. 3. RRCT vs. CRCT (odds ratio with 95% CIs for incidence measures). For off-pump, prostate cancer and routine invasive, ''a'' indicates a Cochrane SR, and ''b'' indicates a non-Cochrane SR. For cardiac rehabilitation and statins, ''a'' and ''b'' indicate different incidence measures.

rates of older people because of death) or misclassifications (e.g., more frequent assignment of natural death in higher age groups), especially in trials with small sample sizes [33]. Second, there can be intervention-related missing values (e.g., a diagnostic intervention might lead to more or less follow-up contact) and misclassifications (e.g., cancer screening and cancer-specific death) [34]. If it is assumed that the data quality of registries is lower than the data quality of actively collected data, then the bias caused by differential missing values and misclassifications might be more serious in RRCTs than CRCTs. In particular, little standardization (measurement time points and instruments) and/or no verification of outcome assessment (e.g., cause of death by biopsy) of registry data can be a source of bias. For example, a Cochrane SR on breast cancer screening found more assignments of breast cancer–specific death than expected (compared with the national average) in the registry-based studies [30].

Mortality and the other incidental findings were primarily from different registries (within and between SRs). Mortality data were mainly from death registries, and data for other incidental findings were from patient registries. The data quality between these registries probably differs; however, we did not recognize a difference in results depending on the outcome type. This finding might also be explained by the assumption that missing values and misclassifications are (completely) at random because confounders that might cause differential missing values or

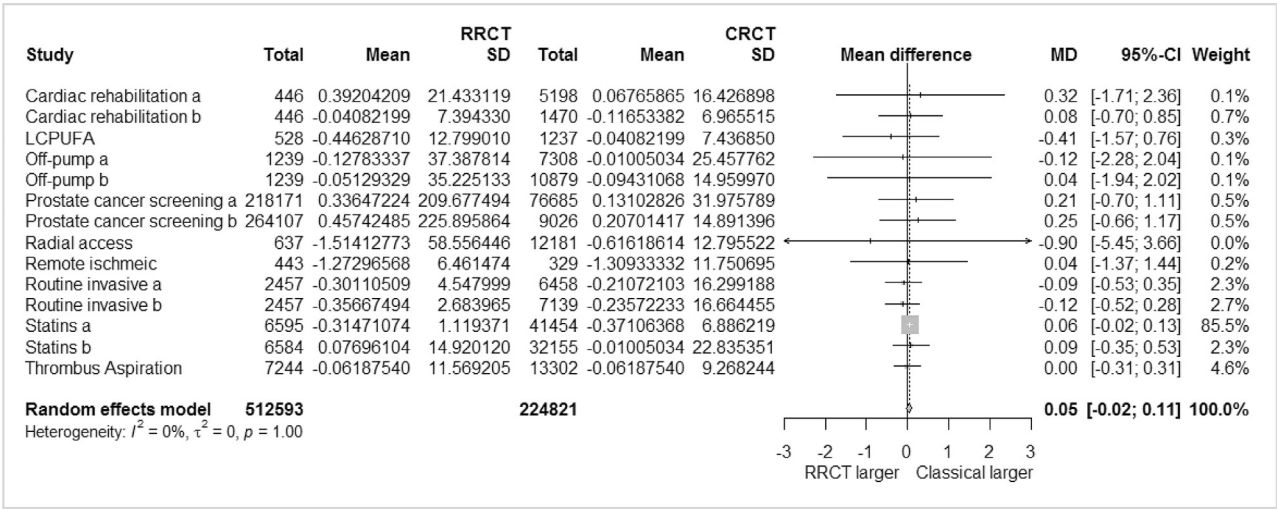

Fig. 4. Meta-analysis of the difference between RRCTs and CRCTs (difference in log odds ratios for incidence measures). For off-pump, prostate cancer and routine invasive, ''a'' indicates a Cochrane SR, and ''b'' indicates a non-Cochrane SR. For cardiac rehabilitation and statins, ''a'' and ''b'' indicate different incidence measures.

misclassifications are balanced at baseline in large RCTs (see Section 1).

Although we found no general difference between effect estimates, we observed conflicting effect directions (RRCT positive effects and CRCT negative effects or vice versa) and point estimators of RRCTs that were not included in the CIs of CRCTs for some comparisons. However, all conflicting effect directions were found at ORs near one (e.g., OR = 1.08 for RRCT vs. OR = 0.99 for CRCT). The point estimators of RRCTs not included in the 95% CIs of CRCT were always near the ends of the 95% CIs, and all 95% CIs of RRCTs overlapped the 95% CIs of CRCTs. In addition, most of these differences observed for comparisons were when the sample size of the meta-analyses was relatively small. Therefore, we propose that the observed differences are rather caused by chance, differences in study characteristics (e.g., patient characteristics), differences in degree of pragmatism (e.g., flexibility of follow-up interventions), or difference in the methodological quality (e.g., risk of bias) than by differences in the data quality. Nevertheless, we could not exclude the possibility that the observed differences are caused by lower data quality. For scientific practice, this indicates that the data quality should be concisely assessed in RCTs assessing outcomes from registries because lower data quality (e.g., larger number of missing values) and consequently higher risk of bias cannot be precluded [30].

A challenge in the interpretation of the results of registry-based studies is that assessing the validity of the registry data for an individual study is often not possible. The reason for this problem (within one study) is that it could not usually be followed if the outcome is not in the registry because it did not occur (e.g., hospital admission) or it is not in the registry because it is missing (e.g., admission to a hospital that does not contribute data to the registry). This problem might be less important for studies with published reports that described the validity of data because this can increase knowledge of the data quality [35]. Therefore, it is beneficial to use all available external information on registry quality (e.g., annual validation reports) to assess the risk of bias. However, if validation reports for the data quality exist, the validity of the data for an individual patient remains unclear because such a report could only provide ''average'' population metrics for data quality description (e.g., proportion of missing values). For example, for an individual participant or groups of participants, it could remain unclear whether the missing value is indeed a missing entry or a nonevent. Moreover, the applicability of the validation report results to the entire study population is questionable because the study population is only a subpopulation of the total population in the registry [36]. Moreover, validation reports refer only to a certain time point or time frame that might not correspond with the time frame of data collection for the RRCT. Because the registry data quality might vary with time, this can be an additional challenge for the judgment of risk of bias. A solution for these problems might be study-specific data validation, for example, by comparing the registry data with a random sample of actively collected data (e.g., by an independent review board) [37].

Detection bias (blinding) is important when interpreting the results from RRCTs. On the one hand, the person who collects the data for the registry is usually not involved in the planning and performance of the study. Thus, data are collected by persons without an interest in the study or, similar to blinding, even without knowledge of the study intervention [37]. On the other hand, for registry data,

generally, no attempts are made to blind the outcome assessment (e.g., outcome adjunction by an independent review board) although detection bias might be an issue. Often, the patient history may not be excluded (e.g., patient records), including received study interventions (e.g., chemotherapy for cancer treatment), and are accessible to the persons (e.g., clinicians) who make the registry entries. Therefore, it might be unclear whether the assessor of the outcome (e.g., remission) is aware of the study intervention. Usually, there is not enough information on the collection of registry data to determine whether the outcome assessment can be considered blinded or unblinded.

Another explanation for our results is that the baseline balance of confounders for missing values and misclassifications between the randomized groups must be judged in the view of the data quality of the possible alternative active data collection method(s). In particular, for very long follow-ups, there is usually also a high risk of a large number of missing values for other data collection methods (e.g., postal or mailed questionnaires, follow-up visits). Therefore, there might be situations in which using registries might be the best alternative although the quality of the registry data is not perfect. In particular, when there is a high risk of low data quality for registry data as well as actively collected data, cross-linking different types of data collection or sources can improve data quality and is usually the best alternative [38]. Using different data sources is especially important in the case of rare events (e.g., side effects) because small changes might have large impacts on the effect estimates and statistical significance [39].

Our analyses of the impact of RRCTs in a meta-analysis show that in some research areas, RRCTs are already the main evidence regarding the number of studies and even more so regarding the number of patients (see, e.g., screening breast cancer, screening prostate cancer). Notably, we excluded one SR on abdominal aortic aneurysm screening because all studies used registry data [40]. The large impact of RRCTs suggests that for research areas such as screening and other preventive interventions in which long-term follow-ups are necessary and for trials on rare events and/or low financial interests of the industry (e.g., public health, health services research), using registries for data collection might be a feasible method to conduct RCTs. Our results indicate no general bias in the effect estimates of RRCTs. Moreover, an RRCT is always at a lower risk of bias than a nonrandomized registry-based study. Therefore, the main question is not whether we should use registries or other preexisting routinely collected health care data for trials but what is the best way to use these data in trials and how can we improve registries for utilization in future research (e.g., public available validation reports). Despite the larger sample sizes in RRCTs, these trials had larger standard deviations than CRCTs. The larger standard deviations suggest a more pragmatic approach (larger variation in patients, intervention) of RRCTs compared with CRCTs, consistent with previous reports [3].

4.1. Limitations

4.1.1. Our study has some limitations

First, our analysis was only based on 15 and 14 comparisons of 12 and 10 unique PIC for mortality and incidence measures, respectively. All RCTs included patients either with cardiovascular diseases or cancer, and most considered a preventive intervention. Therefore, the generalizability of our results to other research areas is limited.

Second, the SRs may have not reported that a study is an RRCT, and we classified this as CRCT. Although we verified each RCT included in the meta-analysis systematically, if registries were used for outcome assessment, we cannot guarantee that studies were not misclassified as CRCTs. Therefore, data from the CRCTs may be contaminated by registry data, which may cloud the difference between RRCTs and CRCTs.

Third, because of the small sample size, we did not match or adjust for differences in study characteristics (e.g., country, mean age), pragmatism, or risk of bias. We believe that it is unlikely that these possible confounders would have covered an existing difference of the ORs of ORs as found in our analysis. However, we could not exclude the possibility that differences in study characteristics, degree of pragmatism, or risk of bias might have caused some of the differences observed in the comparison of effect directions and CI overlap.

Fourth, we only considered mortality, which is an objective outcome, and most of the analyzed incidence measures were also rather objective (e.g., cardiovascular events and hospital admission). Our results might not be applicable to other outcomes, in particular, to more subjective outcomes (e.g., quality of life, cause-specific death), which are at a higher risk of ascertainment bias and misclassification.

5. Conclusion

Our analysis indicates that for objective outcomes, there seems to be no systematic difference between effect estimates from RRCTs and CRCTs. Our sample size was small; we identified only RCTs in cardiovascular conditions and cancer, and we assessed primarily objective outcomes. Therefore, further research is needed to replicate these findings, especially for more subjective outcomes, such as patient-reported outcomes (e.g., quality of life) and other conditions. A meta-matching for study characteristics, degree of pragmatism, and risk of bias should be performed in these studies, if applicable.

Acknowledgments

Funding: This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Conflicts of interest: none.

Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jclinepi.2018.09.011.

References

[1] Rothwell PM. Commentary: external validity of results of randomized trials: disentangling a complex concept. Int J Epidemiol 2010;39:94−6.

[2] Gliklich R, Dreyer NA, Leavy MB, Velentgas P, Khurana L. Standards in the conduct of registry studies for patient-centered outcomes research. Washington: Patient-Centered Outcomes Research Institute (PCORI); 2012.

[3] Mathes T, Buehn S, Prengel P, Pieper D. Registry-based randomized controlled trials merged the strength of randomized controlled trails and observational studies and give rise to more pragmatic trials. J Clin Epidemiol 2018;93:120−7.

[4] Travers K, Sallum RH, Burns MD, Barr CE, Beattie MS, Pashos CL, et al. Characteristics and temporal trends in patient registries: focus on the life sciences industry, 1981-2012. Pharmacoepidemiol Drug Saf 2015;24(4):389−98.

[5] Li G, Sajobi TT, Menon BK, Korngut L, Lowerison M, James M, et al. Registry-based randomized controlled trials- what are the advantages, challenges, and areas for future research? J Clin Epidemiol 2016;80:16−24.

[6] Lauer MS, D'Agostino RBS. The Randomized registry trial — The next disruptive technology in clinical research? N Engl J Med 2013;369(17):1579−81.

[7] Liu JB, D'Angelica MI, Ko CY. The randomized registry trial: two birds, one stone. Ann Surg 2017;265(6):1064−5.

[8] Kjoller E, Hilden J, Winkel P, Galatius S, Frandsen NJ, Jensen GB, et al. Agreement between public register and adjudication committee outcome in a cardiovascular randomized clinical trial. Am Heart J 2014;168:197−204.e1-4.

[9] Computerised record linkage: compared with traditional patient follow-up methods in clinical trials and illustrated in a prospective epidemiological study. The West of Scotland Coronary Prevention Study Group. J Clin Epidemiol 1995;48:1441−52.

[10] Dignam JJ, Huang L, Ries L, Reichman M, Mariotto A, Feuer E. Estimating breast cancer-specific and other-cause mortality in clinical trial and population-based cancer registry cohorts. Cancer 2009; 115(22):5272−83.

[11] Hemkens LG, Contopoulos-Ioannidis DG, Ioannidis JPA. Agreement of treatment effects for mortality from routinely collected data and subsequent randomized trials: meta-epidemiological survey. BMJ 2016;352:i493.

[12] McGowan J, Sampson M, Salzwedel DM, Cogo E, Foerster V, Lefebvre C. PRESS peer review of electronic search strategies: 2015 guideline statement. J Clin Epidemiol 2016;75:40−6.

[13] Jørgensen AW, Hilden J, Gøtzsche PC. Cochrane reviews compared with industry supported meta-analyses and other meta-analyses of the same drugs: systematic review. BMJ 2006;333:782.

[14] Moher D, Tetzlaff J, Tricco AC, Sampson M, Altman DG. Epidemiology and reporting characteristics of systematic reviews. PLoS Med 2007;4:e78.

[15] Franklin JM, Dejene S, Huybrechts KF, Wang SV, Kulldorff M, Rothman KJ. A bias in the evaluation of bias comparing randomized trials with nonexperimental studies. Epidemiol Methods 2017;6(1): 1−14.

[16] Anderson L, Thompson DR, Oldridge N, Zwisler AD, Rees K, Martin N, et al. Exercise-based cardiac rehabilitation for coronary heart disease. Cochrane Database Syst Rev 2016;Cd001800.

[17] Taylor F, Huffman MD, Macedo AF, Moore TH, Burke M, Davey Smith G, et al. Statins for the primary prevention of cardiovascular disease. Cochrane Database Syst Rev 2013;Cd004816.

[18] Chou R, Dana T, Blazina I, Daeges M, Jeanne TL. Statins for prevention of cardiovascular disease in adults: evidence report and systematic review for the US preventive services task force. JAMA 2016; 316:2008−24.

[19] Moller CH, Penninga L, Wetterslev J, Steinbruchel DA, Gluud C. Off-pump versus on-pump coronary artery bypass grafting for ischaemic heart disease. Cochrane Database Syst Rev 2012;Cd007224.

[20] Dieberg G, Smart NA, King N. On- vs. off-pump coronary artery bypass grafting: a systematic review and meta-analysis. Int J Cardiol 2016;223:201−11.

[21] Fanning JP, Nyong J, Scott IA, Aroney CN, Walters DL. Routine invasive strategies versus selective invasive strategies for unstable angina and non-ST elevation myocardial infarction in the stent era. Cochrane Database Syst Rev 2016;Cd004815.

[22] Elgendy IY, Kumbhani DJ, Mahmoud AN, Wen X, Bhatt DL, Bavry AA. Routine invasive versus selective invasive strategies for non-ST-elevation acute coronary syndromes: an updated meta-analysis of randomized trials. Catheter Cardiovasc Interv 2016; 88(5):765−74.

[23] Huynh T, Perron S, O'Loughlin J, Joseph L, Labrecque M, Tu JV, et al. Comparison of primary percutaneous coronary intervention and fibrinolytic therapy in ST-segment-elevation myocardial infarction: bayesian hierarchical meta-analyses of randomized controlled trials and observational studies. Circulation 2009;119:3101−9.

[24] Gunaratne AW, Makrides M, Collins CT. Maternal prenatal and/or postnatal n-3 long chain polyunsaturated fatty acids (LCPUFA) supplementation for preventing allergies in early childhood. Cochrane Database Syst Rev 2015;Cd010085.

[25] Hewitson P, Glasziou P, Irwig L, Towler B, Watson E. Screening for colorectal cancer using the faecal occult blood test, hemoccult. Cochrane Database Syst Rev 2007;Cd001216.

[26] Ilic D, Neuberger MM, Djulbegovic M, Dahm P. Screening for prostate cancer. Cochrane Database Syst Rev 2013;Cd004720.

[27] Djulbegovic M, Beyth RJ, Neuberger MM, Stoffs TL, Vieweg J, Djulbegovic B, et al. Screening for prostate cancer: systematic review and meta-analysis of randomised controlled trials. BMJ 2010;341:c4543.

[28] Usman Ali M, Miller J, Peirson L, Fitzpatrick-Lewis D, Kenny M, Sherifali D, et al. Screening for lung cancer: a systematic review and meta-analysis. Prev Med 2016;89:301−14.

[29] Elgendy AY, Elgendy IY, Mahmoud AN, Bavry AA. Long-term outcomes with aspiration thrombectomy for patients undergoing primary percutaneous coronary intervention: a meta-analysis of randomized trials. Clin Cardiol 2017;40(8):534−41.

[30] Gotzsche PC, Jorgensen KJ. Screening for breast cancer with mammography. Cochrane Database Syst Rev 2013;Cd001877.

[31] Le Page S, Bejan-Angoulvant T, Angoulvant D, Prunier F. Remote ischemic conditioning and cardioprotection: a systematic review and meta-analysis of randomized clinical trials. Basic Res Cardiol 2015;110:11.

[32] Ferrante G, Rao SV, Juni P, Da Costa BR, Reimers B, Condorelli G, et al. Radial versus femoral access for coronary interventions across the entire spectrum of patients with coronary artery disease: a meta-analysis of randomized trials. JACC Cardiovasc Interv 2016;9(14): 1419−34.

[33] Chu R, Walter SD, Guyatt G, Devereaux PJ, Walsh M, Thorlund K, et al. Assessment and implication of prognostic imbalance in randomized controlled trials with a binary outcome − a simulation study. PLoS One 2012;7:e36677.

[34] Breaux S, Perez CA. Pitfalls in the use of death certificates for assessing cause of death: a study of tonsil carcinoma patients. Am J Clin Oncol 1984;7:375−80.

[35] Pieper D, Mathes T, Lefering R, Neugebauer EA. Register-basierte Studien in Evidenzsynthesen - was wir nicht wissen, aber wissen sollten. 9. Ulm, Germany: Jahrestagung der Deutschen Gesellschaft für Epidemiologie; 2014.

[36] Benchimol EI, Smeeth L, Guttmann A, Harron K, Moher D, Petersen I, et al. The REporting of studies conducted using

observational routinely-collected health data (RECORD) statement. PLoS Med 2015;12(10):e1001885.

[37] Mc Cord KA, Al-Shahi Salman R, Treweek S, Gardner H, Strech D, Whiteley W, et al. Routinely collected data for randomized trials: promises, barriers, and implications. Trials 2018; 19(1):29.

[38] Hugosson J, Carlsson S, Aus G, Bergdahl S, Khatami A, Lodding P, et al. Mortality results from the Göteborg randomised population-based prostate-cancer screening trial. Lancet Oncol 2010;11(8): 725−32.

[39] Bolland MJ, Barber PA, Doughty RN, Mason B, Horne A, Ames R, et al. Vascular events in healthy older women receiving calcium supplementation: randomised controlled trial. BMJ 2008;336:262−6.

[40] LeFevre ML, on behalf of the USPSTF. Screening for abdominal aortic aneurysm: U.s. preventive services task force recommendation statement. Ann Intern Med 2014;161:281−90.