PLOS ONE

A statistical method for removing unbalanced trials with multiple covariates in meta-analysis

Massimo Attanasio[1], **Fabio Aiello**[2]*, **Fabio Tinè**[3]

1 Dipartimento di Scienze Economiche, Aziendali e Statistiche, Università di Palermo, Palermo, Italy,
2 Facoltà di Scienze Economiche e Giuridiche, Università "Kore" di Enna, Enna, Italy, **3** Azienda Sanitaria Universitaria Giuliano Isontina (ASI GI), Trieste, Italy

* E-mail: fabio.aiello@unikore.it

Abstract

In meta-analysis literature, there are several checklists describing the procedures necessary to evaluate studies from a qualitative point of view, whereas preliminary quantitative and statistical investigations on the "combinability" of trials have been neglected. Covariate balance is an important prerequisite to conduct meta-analysis. We propose a method to identify unbalanced trials with respect to a set of covariates, in presence of covariate imbalance, namely when the randomized controlled trials generate a meta-sample that cannot satisfy the requisite of randomization/combinability in meta-analysis. The method is able to identify the unbalanced trials, through four stages aimed at achieving combinability. The studies responsible for the imbalance are identified, and then they can be eliminated. The proposed procedure is simple and relies on the combined Anderson-Darling test applied to the Empirical Cumulative Distribution Functions of both experimental and control meta-arms. To illustrate the method in practice, two datasets from well-known meta-analyses in the literature are used.

Introduction

Meta-analysis is an analytical technique designed to "combine" findings from multiple studies. It is commonly used to evaluate studies about medical interventions with the aim to provide researchers, policymakers, and clinicians with useful information. Combinability is a technique that integrates data obtained from dissimilar studies. In meta-analysis' literature, combinability is defined as "*the extent to which separate studies are similar enough*" [1], or "*the extent to which separate studies measure the same thing*" [2]. Of significant interest is the scientific process that enables the integration of studies with similar outcomes. Several approaches have been developed to offer rationale and provide procedures on how studies are chosen, how the data are assembled and how the results are reported.

The literature is vast, and several guidelines have been proposed. The most popular guidelines are QUORUM [3, 4], *Quality of Reporting of Meta-Analyses* and PRISMA [5], *Preferred Reporting Items of Systematic reviews and Meta-Analyses*, a technique that has evolved from Quorum. These guidelines provide a checklist that facilitates a "good" meta-analysis or systematic review.

Competing interests: The authors have declared that no competing interests exist.

The PRISMA checklist consists in qualitative issues and does not cover quantitative issues. Its statistical recommendations focus exclusively with *effect size* measures, confidence intervals and with measures that assess heterogeneity, subgroup analysis or other sources of biases, for instance publication bias. Yusuf and Pogue [6] have already stressed how small sample trial meta-analyses are more susceptible to bias and have advised to choose large sample meta-analyses, to obtain more reliable answers and explore interactions among subgroups. additionally, Cochrane Collaboration [7] defines "systematic review as a review of a clearly formulated question that uses systematic and explicit methods to identify, select, and critically appraise relevant research hypotheses to collect and analyze data from the studies that are included in the review".

To summarize, while qualitative issues of combinability are always examined extensively [8], quantitative issues are essentially limited to sample sizes and effect sizes. In reality, the quantitative assessment of clinical combinability studies is unsatisfactory [9, 10], because it lacks specific quantitative criteria to establish when trials can be considered similar enough. That is why here we propose a method to detect the trials responsible for the lack of combinability, i.e., imbalance between the treatment groups, with respect to some potential risk factors.

In a single randomized controlled trial (RCT), covariate imbalance is a very important statistical problem that has been investigated in many scientific papers. Overall, in a single RCT covariate balance occurs when the patients in each group of treatment are similar as close as possible, particularly with regard to prognostic factors [11–13]. When this condition does not hold, then it is referred as covariate imbalance. The issue arises when dissimilarity between the experimental (*exp*) and control (*ctrl*) arms due to covariate imbalances violates the assumption of "combinability", which is a fundamental premise of meta-analysis. Meta-analysis operates on the assumption that, during the random allocation process to the *exp* and *ctrl* arms, the expected level of covariate distribution imbalance should ideally be zero [14]. Covariate balance is not always assessed before conducting a meta-analysis, automatically assuming that individual studies are well-balanced. However, it can happen that some studies do not exhibit covariate imbalance for some or all covariates, or, as Trowman et al., [15] have pointed out that a meta-analysis imbalance may not result just from a baseline imbalance of one trial, but rather from a cumulative effect of smaller imbalances. In both scenarios, the meta-analysis present covariate imbalance. Other scholars have dealt with covariate imbalance. Riley et al. [16] and Ciolino et al. [14] present methods to assess continuous baseline covariate imbalance across treatment groups in clinical trials with a continuous outcome; Clark et al. [17, 18] claim the relevance of bias due to covariate imbalance in meta-analysis studies with respect to allocation concealment; Hicks et al. [19] and Wewege et al. [20] consider baseline imbalance through statistics calculated for each covariate and they remove those studies where differences are not acceptable.

Here, we support the proposal by Trowman et al. [15] that single slightly unbalanced RCTs could generate a meta-sample that cannot satisfy the randomization/combinability requisite of meta-analysis. The upshot is a "rule of thumb" procedure aimed at eliminating unbalanced trials, which cannot be applied when the number of trials involved is large. Alternatively, individual patient data (IPD) may be used instead of meta-analysis, with the caveat that it requires collection of data of all patients involved in all relevant studies [21]. In addition, if a significant portion of the trials included in a systematic review have baseline imbalance, then combining them in a meta-analysis will produce a misleading result [15]. Thus, to mitigate such bias, it is crucial to conduct meta-regressions with balanced trials.

In this context, we propose a method to identify the studies responsible for the imbalance, with respect to a set of covariates. The Proposed statistical method section describes the main

stages of the method to assess the covariate balance in meta-analysis. The Two datasets section illustrates the two meta-analysis datasets used in our application, coming from the Cholesterol Treatment Trialist' (CTT) Collaboration and the Cochrane library. The Notation section defines the objects and the abbreviations used in the paper. The section Application to the two datasets illustrates how the method is applied, presenting the results and employing a logit model for investigating the relation between balanced and unbalanced trials. In the section Conclusion, we discuss concluding points.

The proposed statistical method

This paper starts from the results of a previous work [22], where a tool was developed to assess the covariate imbalance with respect to a single covariate. Recognizing that clinical practice always involves multiple covariates, we now propose a method for detecting unbalanced trials. The method proposed in this work has:

1. extended the combinability procedure in the presence of three covariates, considering that clinical studies often involve more than one covariate. This also led to a generalization of the test statistics used (for a better understanding of this aspect, changes have been made in the introduction),

2. introduced a new section in this work, a kind of ex-post verification, dedicated to estimating the effect size with unbalanced and balanced trials,

3. included a simulation in the S1 Appendix, providing additional strength to the procedure.

   We propose a stepwise procedure for assessing the imbalance between the treatment groups, with respect to potential factors, comparing their distributions in the treatment groups, without any assumption on their shapes. We classify the potential factors of imbalance as:

- study-level variables (SLVs), which usually include design variables, or population structure variables,

- patient-level variables (PLVs), which are all the baseline variables related to the patients.

   To illustrate this new method, we will refer to the objects defined in S2 Appendix, which are:

- the *exp* meta-arm, i.e., representing a collection of similar experimental arms,

- the *ctrl* meta-arm, i.e., representing a collection of similar control arms,

- the Empirical Cumulative Distribution Function (ECDF) of a PLV built for each meta-arm (see Table 1 in S2 Appendix), consisting in a distribution function in which the frequencies are replaced by the sum of the sample sizes of the arms for each PLV value.

   The covariate balance holds if the ECDFs are not statistically different.
   The rationale of the method is to assess the combinability, to identify the studies responsible for the imbalance. Overall, the method here proposed adheres to four sequential stages.

Assessing the marginal combinability

Marginal combinability holds when the randomization process holds with respect to some basic prognostic factors, over all the levels of a given SLV, that is, the PLVs' ECDFs are not statistically different, controlling for the SLV levels [22]. This is investigated both graphically and analytically, through the Anderson-Darling test (see the Notation section).

Assessing the basic combinability

Basic combinability holds when the PLVs' ECDFs for each meta-arm are not statistically different. If basic combinability does not hold, the meta-analysis cannot be conducted without intervention and/or correction [22]. Also in this case, the basic combinability is investigated both graphically and analytically, through the Anderson-Darling test.

Identifying the unbalanced trials

Among the original studies included in a meta-analysis, an iterative procedure is employed to identify the studies responsible for the highest observed imbalance. This process continues until a statistically reasonable balance is achieved between the *exp* and *ctrl* meta-arms. To do this, we establish an iterative procedure based on a pooled quantity over the PLVs, capable of detecting the trials responsible for the imbalances. Once identified the unbalanced trials, it is necessary to remove these trials.

Removing the unbalanced trials

In this case, it is important to consider both qualitative and quantitative criteria (which are not strictly statistical evaluations). Regarding qualitative issues, the eliminated studies should not compromise the meta-analysis because the studies that have an important scientific value cannot be omitted unless one re-defines the meta-analysis parameters. This can happen if one eliminates the studies that represent a specific subgroup (for example the geographic areas, specific dosages, or important subcases such as diabetics, etc.). Otherwise, the objectives of the meta-analysis should be redefined. Instead, quantitatively, one must balance the total number of eliminated studies and the number of patients corresponding to those studies. In fact, from a practical standpoint it would be best not to surpass a convenient limit for the number of studies, or the number of patients eliminated.

The two datasets

The two examples used pertain to studies with higher incidence rates. The first dataset (S3 File), hereafter named *Chol* (Table 1), is drawn from a well-known meta-analysis [23]. It comprises 26 multicentric randomized trials, conducted by the Cholesterol Treatment Trialist' (CTT) Collaboration, involving two types of trials: more intensive statin regimens versus less intensive statin regimens (5 trials) and statin versus control comparisons (21 trials). We selected the 21 trials of the second type (the PRISMA flowchart is depicted in Fig 1).

These trials involve 129,144 participants with treatment durations of at least 2 years for statin (*exp*) versus control (*ctrl*), assessing the efficacy and safety of cholesterol-lowering therapy on the risk of occlusive vascular events in a wide range of individuals [24–44]. The median follow-up is 4.8 years, during which 7136 participants out of 64540 participants (2.8% per annum) allocated to statin therapy experienced their first major vascular events, compared to 8934 out of 64604 participants (3.6% per annum) participants in the control group. In the *Chol* dataset, the trials were first classified as European, mostly European, North American, Australian, Japanese, and multi-continental. Subsequently, we combined the first two into the European category and grouped the others as non-European.

The second dataset (S4 File), hereafter named *Hep* (S1 Table), is derived from a Cochrane review on Hepatitis C [45]. The review commenced with 72 studies, of which 32 were excluded, based on various criteria (the PRISMA flowchart is provided in Fig 2). We identified 40 studies [46–85], involving 2999 patients in the *ctrl* arms and 4108 in the *exp* arms, conducted in Europe and North America.

Table 1. Cholesterol Treatment Trialists' (CTT) Collaborators 21 studies (*Chol* dataset): Selected SLV and PLVs.

| Trial name[a]	SLV	Arm								
		Ctrl					Exp			
		No. of patients	No. of any major vascular event	PLVs			No. of patients	No. of any major vascular event	PLVs		
	Continent			*mean (age)*	*p(diab)*	*p(male)*			*mean (age)*	*p(diab)*	*p(male)*
SSSS	European	2223	796	58.60	0.04	0.81	2221	555	58.60	0.05	0.82
WOSCOPS	European	3293	318	55.10	0.01	1.00	3302	232	56.30	0.01	1.00
CARE	North AM	2078	553	59.00	0.15	0.86	2081	433	59.00	0.14	0.86
Post-CABG	North AM	677	100	61.60	0.09	0.91	674	79	61.40	0.09	0.93
AFCAPS	North AM	3301	201	58.00	0.05	0.85	3304	143	58.00	0.07	0.85
LIPID	Australia	4502	1153	62.00	0.09	0.83	4512	936	62.00	0.09	0.83
GISSI-P	European	2133	231	60.00	0.14	0.86	2138	208	59.70	0.13	0.86
LIPS	Mostly EU	833	195	60.00	0.10	0.83	844	164	60.00	0.14	0.84
HPS	European	10267	2043	65.20	0.29	0.75	10269	1511	65.20	0.29	0.75
PROSPER	European	2913	495	75.30	0.11	0.48	2891	431	75.40	0.11	0.48
ALLHAT-LLT	North AM	5185	812	66.30	0.34	0.51	5170	758	66.40	0.36	0.51
ASCOT-LLA	European	5137	307	63.20	0.25	0.81	5168	217	63.10	0.24	0.81
ALERT	Mostly EU	1052	140	50.00	0.19	0.65	1050	135	49.50	0.19	0.67
CARDS	European	1410	123	61.80	1.00	0.68	1428	81	61.50	1.00	0.68
ALLIANCE	North AM	1225	293	61.30	0.21	0.82	1217	254	61.10	0.23	0.82
4D	European	636	162	65.70	1.00	0.54	619	144	65.70	1.00	0.54
ASPEN	Multi-Continent	1199	136	61.00	1.00	0.67	1211	114	61.10	1.00	0.66
MEGA	Japan	3966	140	58.40	0.21	0.31	3866	102	58.20	0.21	0.32
JUPITER	Multi-Continent	8901	194	66.00	0.00	0.62	8901	105	66.00	0.00	0.61
GISSI-HF	European	2289	174	68.00	0.25	0.79	2285	172	68.00	0.27	0.76
AURORA	Multi-Continent	1384	368	64.30	0.25	0.65	1389	362	64.10	0.28	0.61
	Totals	64604	8934				64540	7136			

[a]Trial names are consistent with the work of CCT.

As already said, we aim to evaluate the combinability of the studies in a metanalysis, with respect to the PLVs, considering two different types of balance [22]. The first type regards the combinability of the trials concerning the levels of the SLVs (marginal combinability). The second type regards the combinability of the trials with respect to the treatment, *exp* or *ctrl* (basic combinability).

Notation

To avoid cumbersome notation, we have introduced the following terminology:

1. Let $S = \{S_1, S_2, \ldots, S_I\}$ be the original set of I trials, with cardinality $|S| = I$, collected for the meta-analysis. Each of the I trials, S_i (for $i = 1, 2, \ldots, I$), has at least two (k) arms, the control ($k = 1$) arm and the experimental ($k = 2$) arm (see S1 Appendix).

2. Let $S_{(-i)} = \{S \backslash \{S_i\}\}$, for $i = 1, 2, \ldots, I$, be a set of $I–1$ trials. In this way, we get I sets of this kind, each with cardinality $|I–1|$.
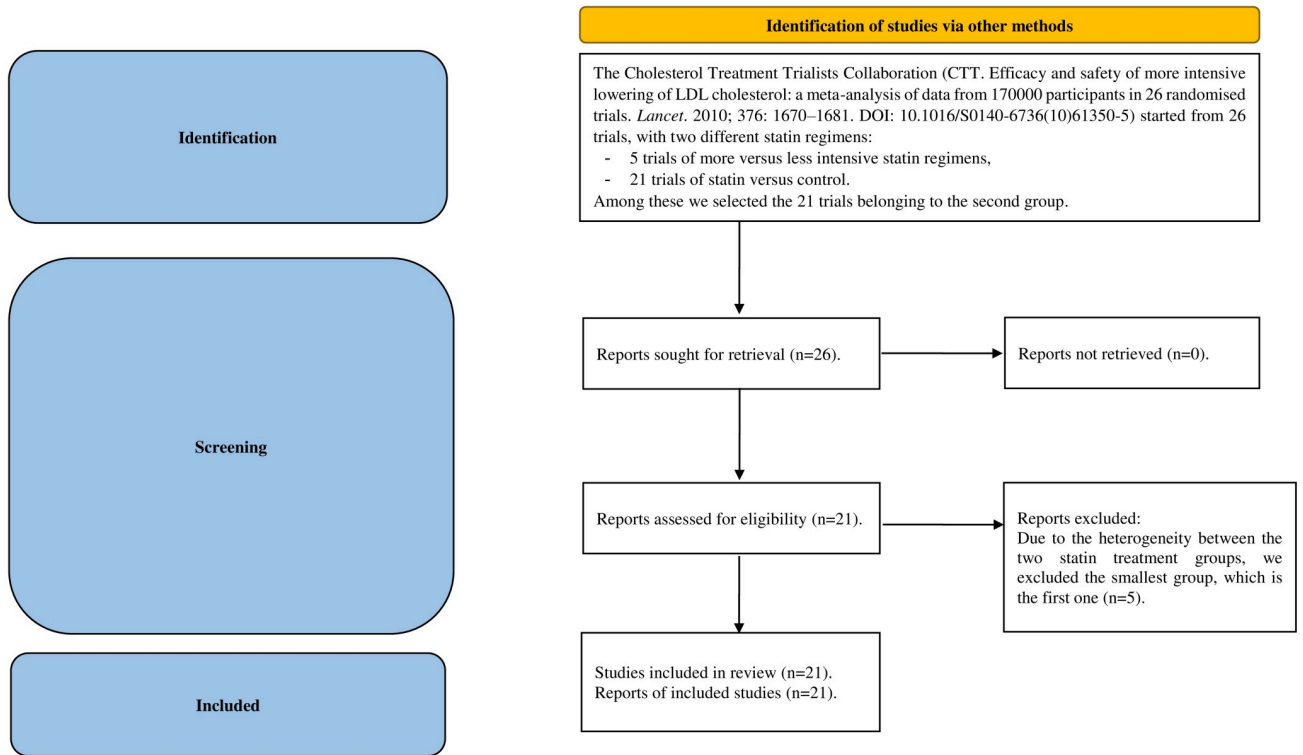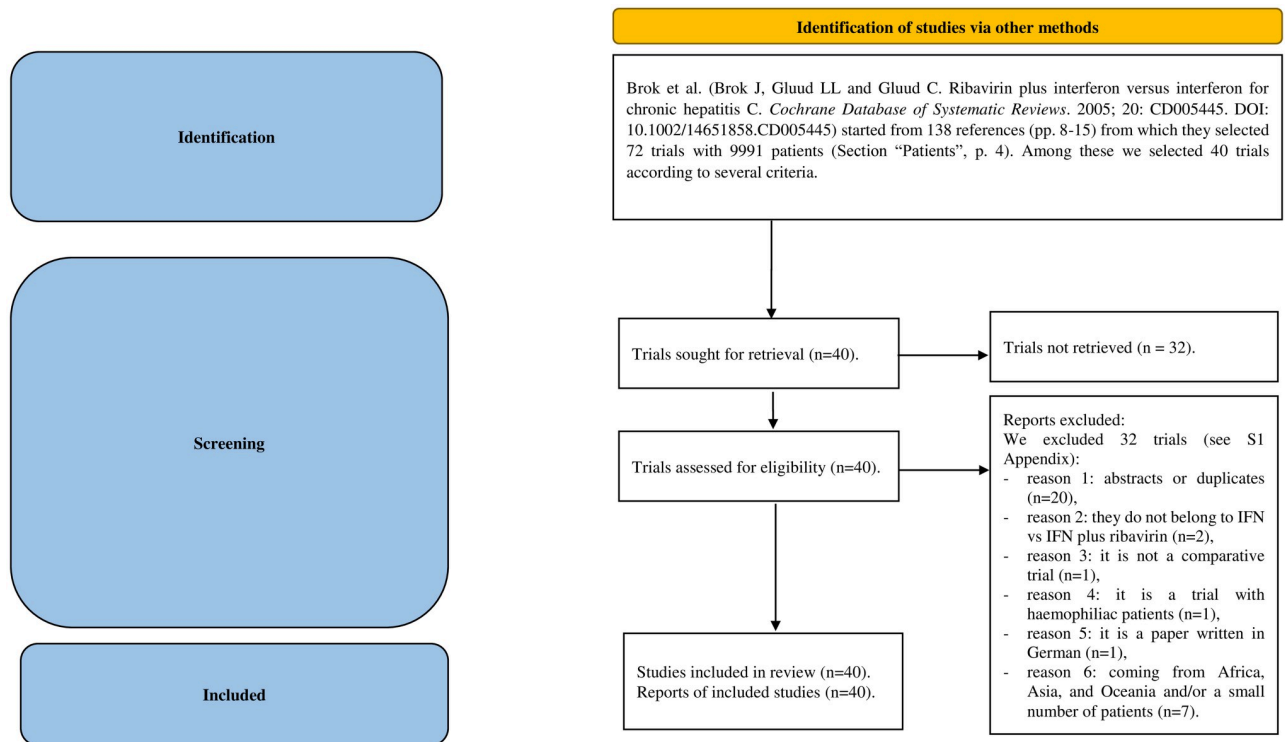
Fig 1. The PRISMA flowchart of *Chol* dataset. *Consider, if feasible to do so, reporting the number of records identified from each database or register searched (rather than the total number across all databases/registers). **If automation tools were used, indicate how many records were excluded by a human and how many were excluded by automation tools. *From*: Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ 2021; 372: n71. doi: 10.1136/bmj.n71. For more information, visit: http://www.prisma-statement.org/.

https://doi.org/10.1371/journal.pone.0295332.g001

3. Let $\mathbf{S}_{(-i)(-i')} = \{\mathbf{S}\backslash\{S_i, S_{i'}\}\}$, for $i' = 1, 2, \ldots, I-1$, be a set of $I-2$ trials. In this way, we get $I-1$ sets of this kind, each with cardinality $|I-2|$.

4. So forth for the triples, until $\mathbf{S}(all)_{(-i)(-i')\ldots(-(I-1))}$.

5. Let $\boldsymbol{ctrl}$, $\boldsymbol{ctrl}_{(-i)}$, $\boldsymbol{ctrl}_{(-i)(-i')}$ be the control meta-arms built over the sets $\mathbf{S}$, $\mathbf{S}_{(-i)}$, $\mathbf{S}_{(-i)(-i')}$.

6. Let $\boldsymbol{exp}$, $\boldsymbol{exp}_{(-i)}$, $\boldsymbol{exp}_{(-i)(-i')}$ be the experimental meta-arms built over the sets $\mathbf{S}$, $\mathbf{S}_{(-i)}$, $\mathbf{S}_{(-i)(-i')}$.

7. $\mathbf{PLV} = \{PLV_1, PLV_2, \ldots, PLV_H\}$ (for $h = 1, 2, \ldots, H$) be a given vector of H PLVs.

8. Let $\mathrm{T}\left(\mathrm{A}_{hkN}^2\right)$ be the k-sample Anderson-Darling test defined as:

$$\mathrm{T}\left(\mathrm{A}_{hkN}^2\right) = \frac{\mathrm{A}_{hkN}^2 - \mu_{hN}}{\sigma_{hN}} \tag{1}$$

where $N = m + n$, that are the sample sizes of the two meta-arms, the A_{hkN}^2 is the k-sample Anderson-Darling criterion, computed for the k meta-arms and for an individual hth PLV, and where $\mu_{hN} = k-1$ and σ_{hN} are the mean and the standard deviation of A_{hkN}^2. Details on the statistical distributions are included in [86].

Fig 2. The PRISMA flowchart of *Hep* dataset. *Consider, if feasible to do so, reporting the number of records identified from each database or register searched (rather than the total number across all databases/registers). **If automation tools were used, indicate how many records were excluded by a human and how many were excluded by automation tools. *From*: Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ 2021; 372: n71. doi: 10.1136/bmj.n71. For more information, visit: http://www.prisma-statement.org/.

https://doi.org/10.1371/journal.pone.0295332.g002

9. Let $A_c^2 = \sum_{h=1}^{H} A_{hkN}^2$ be the combined Anderson-Darling criterion, always computed for the k meta-arms, summing the A_{hkN}^2 criteria over all PLVs. Hence, the overall test is given by:

$$T\left(A_c^2\right) = \frac{A_c^2 - \mu_c}{\sigma_c} \qquad (2)$$

where $\mu_c = \sum_{h=1}^{H} \mu_{hN}$ and $\sigma_c = \sqrt{\sum_{h=1}^{H} \sigma_{hN}^2}$ are the mean and the standard deviation of A_c^2. The $T\left(A_c^2\right)$ statistic is the combined Anderson-Darling k-sample test under the hypothesis that the independent arms within each trial come from a common unspecified continuous distribution.

In meta-analysis, all arms of all trials are assumed to be independent and from identical continuous distributions. Both the individual criterion, A_{hkN}^2, and the combined Anderson-Darling criterion, A_c^2, (and the corresponding standardized statistics, $T\left(A_{hkN}^2\right)$ and T_c, respectively) are used to test simultaneously whether the arms of each trial come from the same continuous distribution function, i.e., whether they are balanced. These standardized statistics are expected to be zero under the null hypothesis. Thus, the larger the statistics, the greater the overall imbalance.

In the case of dependent samples, one can refer to the suggestions made by Lin and Sullivan [87] and Han et al. [88].

The application to the two datasets

This section demonstrates the application of the proposed method. The first two stages are applied to both datasets, but for brevity, we will illustrate the iterative procedure only for the *Chol* dataset, while the results will be reported for both datasets.

The *Chol* and *Hep* datasets comprise of 21 and 40 trials, respectively. In both datasets, we refer to the SLV *Continent* (European, EU, and Non-European, Non-EU), because it reflects different epidemiological populations (S2 Table). The PLVs consist of well-known risk factors associated to the disease under study. In the *Chol* dataset, these include the patients' mean age, *mean*(*age*), the proportion of diabetics, *p*(*diab*), the proportion of males, *p*(*male*). In the *Hep* dataset, the PLVs include the patients' mean age, *mean*(*age*), the proportion of cirrhotic patients, *p*(*cirr*), and the proportion of males, *p*(*male*). We applied the method in three stages as described above.

Assessing the marginal combinability

We construct the ECDFs of each PLV, controlling for the two levels of the SLV *Continent*, i.e., European, and Non-European trials. We then compared each pair of ECDFs using the Anderson-Darling test for both datasets. Fig 3 illustrates that the ECDFs are noticeably different from each other, with all the *p*-values being significant. Therefore, the distributions of the PLVs are structurally and statistically different in European and Non-European trials. For brevity, we applied our method only to the subsets of European studies for both datasets: **S1** (with $|\mathbf{S1}| = 11$) for the *Chol* dataset, and **S2** (with $|\mathbf{S2}| = 34$) for the *Hep* dataset.

Assessing the basic combinability

We constructed the ECDFs of each PLV separately for both the *exp* and *ctrl* meta-arms within the **S1** and **S2** subsets of the *Chol* and *Hep* datasets, respectively. We then investigated the basic combinability of data by comparing all pairs of ECDFs, both graphically (Fig 4) and analytically (Table 2), using the Anderson-Darling statistics based on the quantities A^2_{hkN}, μ_{hN}, σ_{hN}.

The ECDFs for the first dataset (see Fig 4a–4c) exhibit closeness, while in the second dataset (see Fig 4d–4f) the ECDFs are less similar. This difference is likely due the varied distribution among the studies in the datasets.

Table 2 reports the quantities, namely, A^2_{hkN}, μ_{hN}, σ_{hN}, to compute the individual and the combined Anderson-Darling statistics, $T(A^2_{hkN})$ and T_h (for $h = 1, 2, 3$), which measure the basic imbalance concerning the PLVs. The overall T_c is obtained by summing the three $T(A^2_{hkN})$. The *p*-values are all significant, denoting that the ECDFs are all statistically different and hence the trials are not balanced concerning the PLVs under consideration.

Identifying the unbalanced trials

We developed a backward reduction procedure to select the balanced trials, implemented using the "kSample" package of R software [89]. It compares the ECDFs of the meta-arms using the A-D test. To avoid ties in the ECDFs, the values were perturbed by a random component. The procedure identifying the unbalanced studies is based on comparing the quantities $T_{c(-i)}$, calculated over the subsets:

$$\mathbf{S1}_{(-i)} = \{\mathbf{S1}\backslash\{S1_i\}\} \Rightarrow T_{c(-i)} \text{ where } |\mathbf{S1}_{(-i)}| = 10, \forall\, i = 1, \ldots, 11.$$
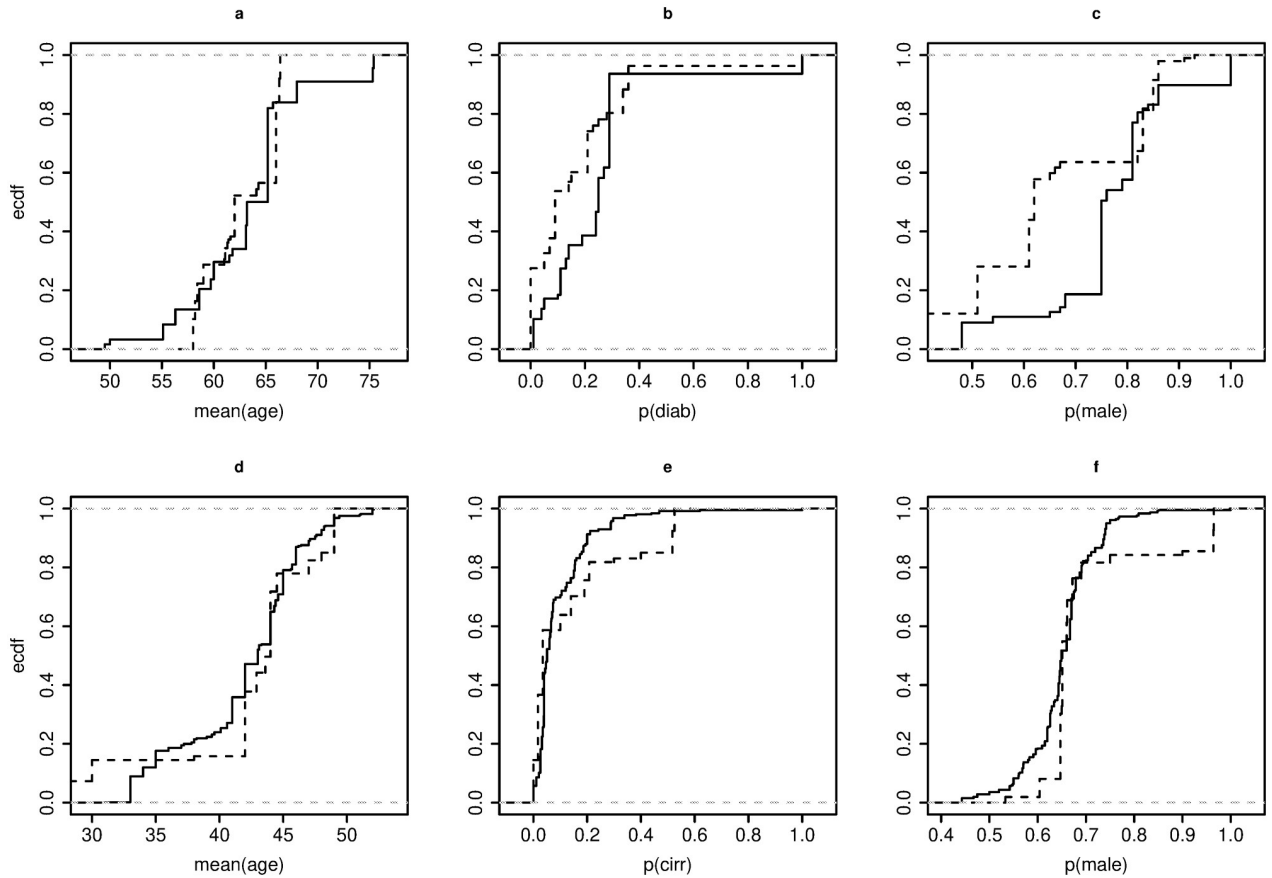
Fig 3. ECDFs of the European (——) and Non-European (— —) meta-arms, with respect to the PLVs. *Chol* dataset (a, b, c); *Hep* dataset (d, e, f).

$\mathbf{S2}_{(-j)} = \{\mathbf{S2}\backslash\{S2_j\}\} \Rightarrow T_{c(-j)}$ where $|\mathbf{S2}_{(-j)}| = 33$, $\forall j = 1, \ldots, 34$.

For simplicity, let us assume that all the risk factors have equal weight (although different weights can also be applied) and we will proceed with the subset **S1** of the *Chol* dataset, even though the results will be reported for both datasets.

At each step, the decision rule eliminates the study associated with the greatest overall imbalance from the initial set, **S1**. The steps of the iterative procedure are:

1. Build I sets $\{\mathbf{S1}\backslash\{S1_i\}\}$, for $i = 1, 2, \ldots, I$, whose cardinality is $|I{-}1|$.

2. Compute the quantities $T_{h(-i)}$, for $h = 1, 2, 3$, and $\{T_{c(-i)}\}$, $\forall i = 1, \ldots, I$.

3. Identify the $min_i\{T_{c(-i)}\}$, and then the corresponding ith study.

4. Remove the ith study, $S1_i$, and consider the set $\mathbf{S1}_{(-i)} = \{\mathbf{S1}\backslash\{S1_i\}\}$.

5. Rename $\mathbf{S1} = \{\mathbf{S1}\backslash\{S1_i\}\}$.

6. If $min_i\{T_{c(-i)}\}$ is not significant at the level of $\alpha = 0.05$, then stop,

7. Otherwise, go to step 1.

The procedure terminates when the value $min_i\{T_{c(-i)}\}$ is not significant, and the "latest" **S1**, consisting of the non-removed trials, is *balanced* with respect to the chosen covariates.
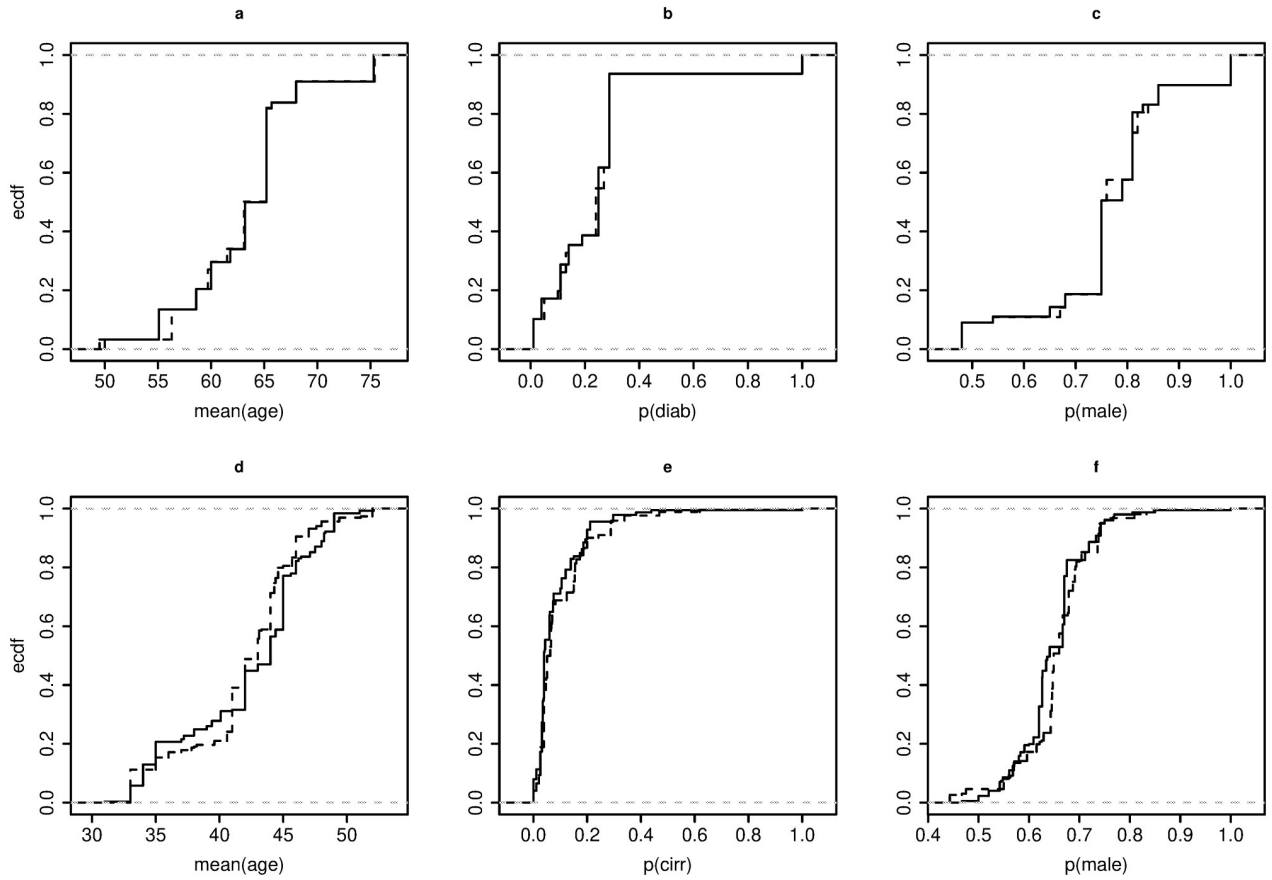
Fig 4. ECDFs of the *exp* (——) and *ctrl* (——) meta-arms in the European studies, for the PLVs. S1 (a, b, c) of the *Chol* dataset and **S2** (d, e, f) of the *Hep* dataset.

https://doi.org/10.1371/journal.pone.0295332.g004

1st iteration of the procedure.

1. Build 11 sets $\mathbf{S1}_{(-i)} = \{\mathbf{S1}\backslash\{S1_i\}\} \; \forall \; i = 1, \ldots, 11$, whose cardinality is 10.

2. Compute $T_{1(-i)}$, $T_{2(-i)}$, $T_{3(-i)}$, and $T_{c(-i)}$ (S3 Table).

3. As the $min_i\{T_{c(-i)}\}$ is $T_{1(-12)} = 58.9$, then identify $S1_{12}$.

4. Remove $S1_{12}$; consider $\mathbf{S1}_{(-12)} = \{\mathbf{S1}\backslash\{S1_{12}\}\}$.

Table 2. Quantities of the Anderson-Darling statistics and *p*-values, for S1 (*Chol* dataset) and S2 (*Hep* datasets).

	S1				S2			
	mean(age)	*p(diab)*	*p(male)*	*Combined*	*mean(age)*	*p(cirr)*	*p(male)*	*Combined*
A^2_{hkN}	11.09	51.42	25.29	87.81	158.20	2450.00	1086.00	3694.20
μ_{hN}	1	1	1	3	1	1	1	3
σ_{hN}	0.761	0.761	0.761	1.319	0.761	0.761	0.761	1.319
$T(A^2_{hkN})$	13.25	66.22	31.91	–	206.50	3218.00	1425.00	–
T_c	–	–	–	64.30	–	–	–	2798.48
p	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000

https://doi.org/10.1371/journal.pone.0295332.t002

5. Rename $\mathbf{S1} = \{\mathbf{S1}\backslash\{S1_{12}\}\}$.

6. Since $T_{c(-12)}$ is significant, return to step 1.

2nd iteration of the procedure.

1. Build 10 sets $\mathbf{S1}_{(-i)} = \{\mathbf{S1}\backslash\{S1_i\}\} \; \forall \; i = 1, \ldots, 10$, whose cardinality is 9.

2. Compute $T_{1(-i)}$, $T_{2(-i)}$, $T_{3(-i)}$, and $T_{c(-i)}$ (S4 Table).

3. As the $min_i\{T_{c(-i)}\}$ is $T_{c(-1)} = 43.1$, then identify $S1_1$.

4. Remove $S1_1$; consider $\mathbf{S1}_{(-1)} = \{\mathbf{S1}\backslash\{S1_1\}\}$.

5. Rename $\mathbf{S1} = \{\mathbf{S1}\backslash\{S1_1\}\}$.

6. Since $T_{c(-1)}$ is significant, return to step 1.

Now, let's skip to the last one, keeping in mind that we removed 5 studies ($S1_{12}$, $S1_1$, $S1_2$, $S1_8$, $S1_{20}$).

6th iteration.

1. Build 6 sets $\mathbf{S1}_{(-i)} = \{\mathbf{S1}\backslash\{S1_i\}\} \; \forall \; i = 1, \ldots, 6$, whose cardinality is 5.

2. Compute $T_{1(-i)}$, $T_{2(-i)}$, $T_{3(-i)}$, and $T_{c(-i)}$ (S5 Table).

3. As the $min_i\{T_{c(-i)}\}$ is $T_{c(-7)} = 1.1$, then identify $S1_7$.

4. Remove $S1_7$; consider $S1_{(-7)} = \{\mathbf{S1}\backslash\{S1_7\}\}$.

5. Rename $\mathbf{S1} = \{\mathbf{S1}\backslash\{S1_7\}\}$.

6. Since $T_c(-7)$ is not significant ($p = 0.138$), stop the procedure.

Table 4 summarizes the results of the backward reduction procedure at each step, for both datasets, *Chol* and *Hep*. The iterations are 6 for **S1**, and 4 for **S2**, leading to the final subset of balanced trials:

- **BAL1** = ($S1_9$, $S1_{10}$, $S1_{13}$, $S1_{14}$, $S1_{16}$), for *Chol* dataset.

- **BAL2** = ($S2_1$, $S2_2$, $S2_4$, $S2_5$, $S2_6$, $S2_7$, $S2_9$, $S2_{10}$, $S2_{11}$, $S2_{12}$, $S2_{13}$), for *Hep* dataset.

    the unbalanced trials are:

- **UNB1** = **S1** −**BAL1** for *Chol* dataset.

- **UNB2** = **S2** −**BAL2** for *Hep* dataset.

    The last columns of Table 3 report, for both datasets, the percentage of lost patients at each iteration, which reaches 49.5% and 24.3%, respectively.

    As expected, the ECDFs built over **BAL1** and **BAL2** show reasonable overlapping between the *exp* and *ctrl* meta-arms (Fig 5).

## The effect of the *imbalance* on the outcome variable

The previous procedure effectively identifies trials that may be responsible for imbalances between *exp* and *ctrl* meta-arms, without taking into account the outcome variables. The proposed solution aims to remove the unbalanced trials to conduct a "proper" meta-analysis. However, in this section we apply a meta-regression to evaluate the effects of the treatment (i.e., the arm type) and the "imbalance" on the outcomes of the *Chol* and *Hep*

Table 3. Results of the backward reduction procedure by iteration: Statistics, *p*-values, and reduction of both studies and patients. S1 (*Chol* dataset) and S2 (*Hep* dataset).

Iteration r	S1							S2						
	T_c statistics	p	Deleted Study	Deleted pts	No. of studies	No. of pts	Pts' Reduction (%)	T_c statistics	p	Deleted Study	Deleted pts	No. of studies	No. of pts	Pts' Reduction (%)
0	64.3	< 0.001	0	0	11	64401	0	2798.5	< 0.001	0	0	34	6473	0
1	59.8	0.008	$S1_{12}$	10305	10	54096	-16.0	14.9	< 0.001	$S2_3$	303	33	6170	-4.7
2	43.1	< 0.001	$S1_1$	4444	9	49652	-22.9	6.2	< 0.001	$S2_{25}$	832	32	5338	-17.5
3	18.7	< 0.001	$S1_{20}$	4574	8	45078	-30.0	3.3	< 0.05	$S2_{18}$	376	31	4962	-23.3
4	13.8	< 0.001	$S1_8$	1677	7	43401	-32.6	1.8	0.061	$S2_8$	60	30	4902	-24.3
5	10.5	< 0.001	$S1_2$	6595	6	36806	-42.8	-	-	-	-	-	-	-
6	1.1	0.138	$S1_7$	4271	5	32535	-49.5	-	-	-	-	-	-	-

https://doi.org/10.1371/journal.pone.0295332.t003

dataset. The goal of this application is to illustrate the effect that including unbalanced trials would have had on the outcome. The outcomes are the probability *p* of "occlusive vascular events" for the *Chol* dataset, and the "sustained response" for the *Hep* dataset. To achieve this, we will use a meta-regression logit model with two dummy variables, defined
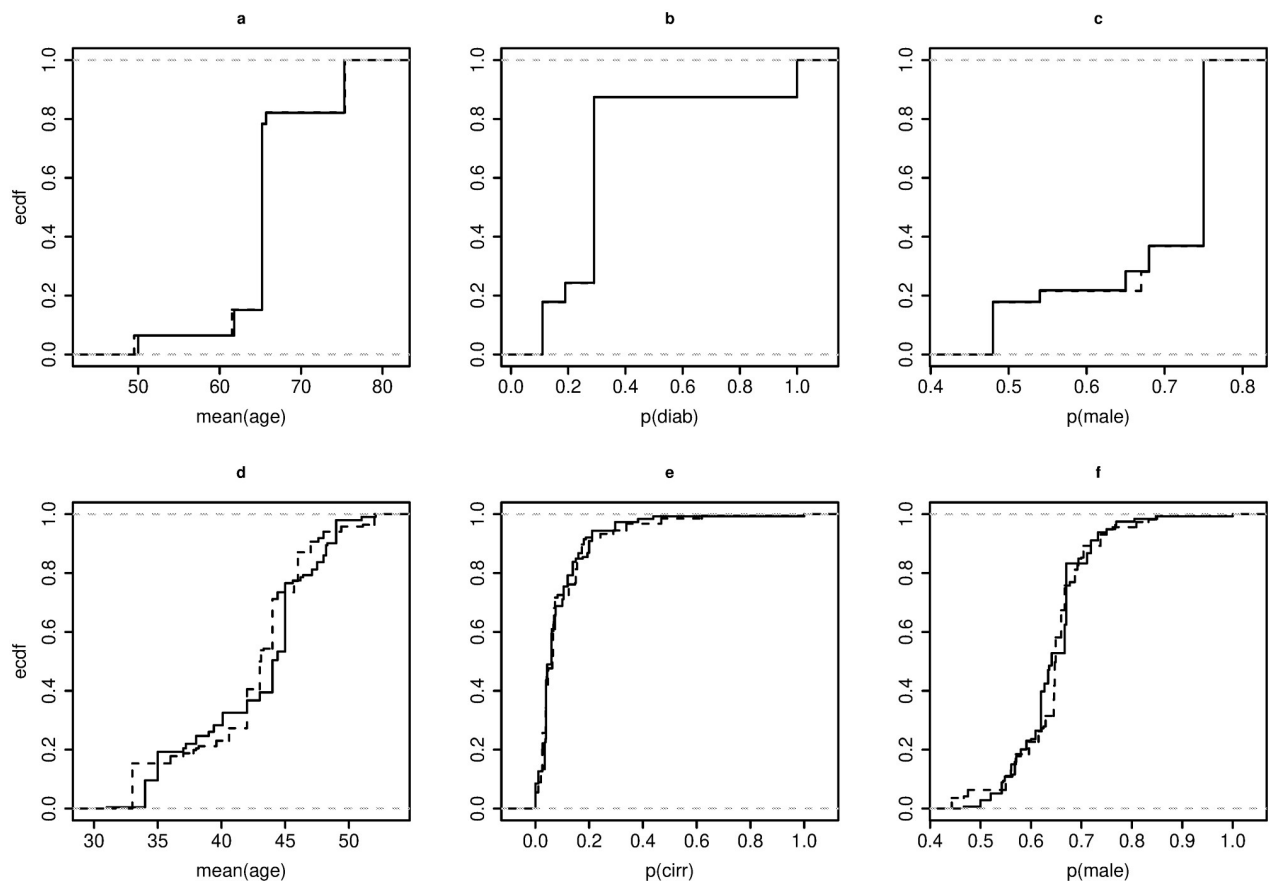
Fig 5. ECDFs of the *exp* (——) and *ctrl* (——) meta-arms in the European studies, for the PLVs. BAL1 (a, b, c) of the *Chol* dataset and BAL2 (d, e, f) of the *Hep* dataset.

https://doi.org/10.1371/journal.pone.0295332.g005

as follows:

$$arm_k = \begin{cases} k = 0 \text{ if } ctrl \\ k = 1 \text{ if } exp \end{cases}$$

$$imb_m = \begin{cases} m = 0 \begin{cases} \text{if } \mathbf{BAL1} \in Chol \\ \text{if } \mathbf{BAL2} \in Hep \end{cases} \\ m = 1 \begin{cases} \text{if } \mathbf{UNB1} \in Chol \\ \text{if } \mathbf{UNB2} \in Hep \end{cases} \end{cases}$$

The model is defined as $logit(p_{km}) = \beta_0 + \beta_1 arm_k + \beta_2 imb_m$, where $k = 0, 1$ and $m = 0, 1$. The parameter estimates and their standard errors are reported in Table 4.

It is evident that imb has a significant effect on the outcomes, while the interactions are not significant. This suggests that the presence of unbalanced trials should always be investigated, to avoid biased estimates of treatment effects. It is important to emphasize that logistic regression with the inclusion of a dummy variable indicating the presence of unbalanced studies does not resolve the problem when the sample size of balanced studies obtained through the procedure is limited. Instead, it serves as a warning about the effect size of both balanced and unbalanced studies on the outcome.

Conclusions

As highlighted in the introduction, there is a lack of statistical methods for assessing systematic differences in patients' characteristics in meta-analysis studies, even though numerous methods and procedures exist for correcting covariate imbalances in individual RCTs [11, 19]. It is important to note that incorporating unbalanced trials can have a significant effect on the assessment of the response [90]. In this context, we conducted a meta-regression aiming at illustrating the effect that including balanced trials would have had on the outcome. In fact, the meta-regression equation tells us just that the presence of unbalanced trials may change (if the parameter is significant) the effect size.

In clinical practice, researchers always encounter trials with multiple covariates, and it becomes essential to evaluate whether a trial can be considered balanced as a whole, regardless of the balance of individual covariates. In this regard, to address this issue, we presented a method for removing trials that simultaneously considers three covariates, building a prior study [22] that tackled the issue in the presence of a single covariate. The method involves constructing meta-arms, which are collections of similar randomized experimental or control arms. These meta-arms are then compared through their ECDFs to determine whether the randomization concerning a set of risk factors holds. If randomization is not upheld, the trials responsible for the imbalance are identified iteratively using a statistical test based on the

Table 4. Estimated parameters of the meta-regression logistic models. S1 (*Chol* dataset) and S2 (*Hep* dataset).

Coefficients	S1			S2		
	Estimate	Std.Err.	*p*	Estimate	Std.Err.	*p*
intercept	-1.502	0.018	< 0.001	-1.942	0.059	< 0.001
arm	-0.301	0.023	< 0.001	1.344	0.068	< 0.001
imb	-0.427	0.023	< 0.001	-0.139	0.069	< 0.05

distance between the ECDFs. We have also conducted a simulation study with various scenarios to strengthen to the method's validity.

One limitation of this method is that it may lead to a reduction of the number of trials involved in the meta-analysis. Therefore, investigators must decide whether the meta-analysis is still meaningful after the removal of many unbalanced trials.

Finally, our work proposes a method of backward elimination of studies. Nowadays, meta-analyses have the potential to include many studies, so the removal of trials should not compromise the conduct of the meta-analysis itself. However, there are alternative statistical methods, such as propensity score methods, which address imbalance through re-weighting procedures and could offer a solution. Nonetheless, these methods are generally more computationally intensive, and employ a distinct approach.

Supporting information

S1 Appendix. Simulations.
(PDF)

S2 Appendix. Meta-arms and ECDF.
(PDF)

S1 File. PRISMA checklist.
(PDF)

S2 File. The studies' selection procedure for the *Hep* dataset.
(PDF)

S3 File. *Chol* dataset. File of the *Chol* dataset.
(TXT)

S4 File. *Hep* dataset. File of the *Hep* dataset.
(TXT)

S1 Table. Ribavirin plus interferon versus interferon for chronic hepatitis C's 40 studies (*Hep* dataset) selected SLV and PLVs.
(XLSX)

S2 Table. European and Non-European studies in *Chol* and *Hep* datasets.
(XLSX)

S3 Table. 1st iteration: Anderson-Darling test statistics $T_{1(-i)}$, $T_{2(-i)}$, $T_{3(-i)}$, and $T_{c(-i)}$. *Chol* dataset.
(XLSX)

S4 Table. 2nd iteration: Anderson-Darling test statistics $T_{1(-i)}$, $T_{2(-i)}$, $T_{3(-i)}$, and $T_{c(-i)}$. *Chol* dataset.
(XLSX)

S5 Table. 6th iteration: Anderson-Darling test statistics $T_{1(-i)}$, $T_{2(-i)}$, $T_{3(-i)}$, and $T_{c(-i)}$. *Chol* dataset.
(XLSX)

Acknowledgments

Author Contributions

Conceptualization: Massimo Attanasio.

Data curation: Fabio Aiello, Fabio Tinè.

Formal analysis: Massimo Attanasio, Fabio Aiello.

Funding acquisition: Massimo Attanasio.

Methodology: Massimo Attanasio, Fabio Aiello.

Software: Fabio Aiello.

Supervision: Massimo Attanasio, Fabio Aiello, Fabio Tinè.

Writing – original draft: Massimo Attanasio, Fabio Aiello, Fabio Tinè.

Writing – review & editing: Massimo Attanasio, Fabio Aiello.

References

1. Sacks HS, Berrier J, Reitman D, Ancona-Berk VA, Chalmers T. Meta-analyses of randomized controlled trials. N Engl J Med. 1987; 316(8):450–455. https://doi.org/10.1056/NEJM198702193160806 PMID: 3807986.

2. Macarthur C, Foran PJ, Bailar JC III. Qualitative assessment of studies included in a meta-analysis: DES and the risk of pregnancy loss. J Clin Epidemiol. 1995; 48(6):739–747. https://doi.org/10.1016/0895-4356(94)00190-2 PMID: 7769404.

3. Clarke M. The QUORUM Statement. Lancet. 2000; 355(9205):756–757. https://doi.org/10.1016/S0140-6736(05)72172-3 PMID: 10703836.

4. Moher D, Cook DJ, Eastwood S, Olkin I, Rennie D, Stroup DF, et al. Improving the quality of reports of meta-analyses of randomized controlled trials: the QUOROM statement. Quality of Reporting of Meta-analyses. Lancet. 1999; 354(9193):1896–1900. https://doi.org/10.1016/s0140-6736(99)04149-5 PMID: 10584742.

5. Liberati A, Altman DG, Tetzlaff J, Mulrow C, Gøtzsche PC, Ioannidis JP, et al. The PRISMA Statement for Reporting Systematic Reviews and Meta-Analyses of Studies That Evaluate Health Care Interventions: Explanation and Elaboration. PLOS Med. 2009; 6(7):1–28. https://doi.org/10.1371/journal.pmed.1000100 PMID: 19621070.

6. Pogue J, Yusuf S. Overcoming the limitations of current meta-analysis of randomized controlled trials. Lancet. 1998; 351(9095):47–52. https://doi.org/10.1016/S0140-6736(97)08461-4 PMID: 9433436.

7. COCHRANE METHODS. Editors: Hopewell S, Clarke M, Higgins J. Wiley-Blackwell, Oxford; 2011.

8. Haidich AB. Meta-analysis in medical research. Hippokratia. 2010; 14(1):29–37. PMID: 21487488.

9. Feinstein AR. Meta-analysis: Statistical alchemy for the 21st century. J Clin Epidemiol. 1995; 48(1):71–79. https://doi.org/10.1016/0895-4356(94)00110-c PMID: 7853050.

10. Borenstein M, Hedges LV, Higgins JPT, Rothstein HR. When does it make sense to perform a meta-analysis? Introduction to meta-analysis. New York: John Wiley & Sons. 2009. p. 19–40.

11. Altman DG. Comparability of randomised groups. J Royal Stat Soc. Ser D (The Statistician). 1985; 34(1):125–136. https://doi.org/10.2307/2987510

12. Altman DG. Covariate imbalance, adjustment for. In: Armitage P, Colton T, editors. Encyclopedia of Biostatistics. Chichester: Wiley; 1998. p. 1000–1005.

13. Berger VW, Weinstein S. Ensuring the comparability groups: is randomization enough? Control Clin Trials. 2004; 25(5):515–524. https://doi.org/10.1016/j.cct.2004.04.001 PMID: 15465620

14. Ciolino JD, Martin RH, Zhao W, Hill MD, Jauch EC, Palesch YY. Measuring continuous baseline covariate imbalances in clinical trial data. Stat Methods Med Res. 2015; 24(2):255–272. https://doi.org/10.1177/0962280211416038 PMID: 21865270.

15. Trowman R, Dumville JC, Torgerson DJ, Cranny G. The impact of trial baseline imbalances should be considered in systematic reviews: a methodological case study. J Clin Epidemiol. 2007; 60(12):1229–1233. https://doi.org/10.1016/j.jclinepi.2007.03.014 PMID: 17998076.

16. Riley RD, Kauser I, Bland M, Thijs L, Staessen JA, Wang J, et al. Meta-analysis of randomized trials with a continuous outcome according to baseline imbalance and availability of individual participant data. Stat. Med. 2013; 32(16):2747–2766. https://doi.org/10.1002/sim.5726 PMID: 23303608

17. Clark L, Fairhurst C, Hewitt CE, Birks Y, Brabyn S, Cockayne S, et al. A methodological review of recent meta-analyses has found significant heterogeneity in age between randomized groups. J Clin Epidemiol. 2014; 67(9):1016–1024. https://doi.org/10.1016/j.jclinepi.2014.04.007 PMID: 24909873.

18. Clark L, Fairhust C, Torgerson DJ. Allocation concealment in randomised controlled trials: are we getting better? BMJ. 2016; 355:i5663. https://doi.org/10.1136/bmj.i5663 PMID: 27856428

19. Hicks A, Fairhust C, Torgerson DJ. A simple technique investigating baseline heterogeneity helped to eliminate potential bias in meta-analyses. J Clin Epidemiol. 2017; 95:55–62. https://doi.org/10.1016/j.jclinepi.2017.10.001 PMID: 29032245

20. Wewege MA, Hansford HJ, Shah B, Gilanyi YL, Douglas SRG, Parmenter BJ, et al. Hypertens Res. 2022; 45:1643–1652. https://doi.org/10.1038/s41440-022-00984-3 PMID: 35882996

21. Simmonds MC, Higgins JP. Covariate heterogeneity in meta-analysis: criteria for deciding between meta-regression and individual patient data. Stat Med. 2007; 26(15):2982–2999. https://doi.org/10.1002/sim.2768 PMID: 17195960

22. Aiello F, Attanasio M, Tinè F. Assessing covariate imbalance in meta-analysis studies. Stat Med. 2011; 30(22):2671–2682. https://doi.org/10.1002/sim.4311 PMID: 21786283

23. Cholesterol Treatment Trialists (CTT) Collaboration. Efficacy and safety of more intensive lowering of LDL cholesterol: a meta-analysis of data from 170000 participants in 26 randomised trials. Lancet. 2010; 376(9753):1670–1681. https://doi.org/10.1016/S0140-6736(10)61350-5 PMID: 21067804.

24. Scandinavian Simvastatin Survival Study Group. Randomised trial of cholesterol lowering in 4444 patients with coronary heart disease: the Scandinavian Simvastatin Survival Study (4S). Lancet. 1994; 344(8934):1383–1389. https://doi.org/10.1016/S0140-6736(94)90566-5 PMID: 7968073

25. Shepherd J, Cobbe SM, Ford I, Isles CG, Lorimer AR, Macfarlane PW, et al. Prevention of coronary heart disease with pravastatin in men with hypercholesterolemia. N Engl J Med. 1995; 333(20):1301–1308. https://doi.org/10.1056/NEJM199511163332001 PMID: 7566020.

26. Sacks FM, Pfeffer MA, Moyé LA. The effect of pravastatin on coronary events after myocardial infarction in patients with average cholesterol levels. Cholesterol and Recurrent Events Trial investigators. N Engl J Med. 1996; 335(14):1001–1009. https://doi.org/10.1056/NEJM199610033351401 PMID: 8801446.

27. The Post Coronary Artery Bypass Graft Trial Investigators (Post-CABG). The effect of aggressive lowering of low-density lipoprotein cholesterol levels and low-dose anticoagulation on obstructive changes in saphenous-vein coronary-artery bypass grafts. N Engl J Med. 1997; 336(3):153–162. https://doi.org/10.1056/NEJM199701163360301 PMID: 8992351.

28. Downs JR, Clearfield M, Weis S. Primary prevention of acute coronary events with lovastatin in men and women with average cholesterol levels: results of AFCAPS/TexCAPS. JAMA. 1998; 279(20):1615–1622. https://doi.org/10.1001/jama.279.20.1615 PMID: 9613910

29. The Long-Term Intervention with Pravastatin in Ischaemic Disease (LIPID) Study Group. Prevention of cardiovascular events and death with pravastatin in patients with coronary heart disease and a broad range of initial cholesterol levels. N Engl J Med. 1998; 339(19):1349–1357. https://doi.org/10.1056/NEJM199811053391902 PMID: 9841303.

30. Prevenzione Investigators (Gruppo Italiano per lo Studio della Sopravvivenza nell'Infarto Miocardico). Results of the low dose (20 mg) pravastatin GISSI Prevenzione trial in 4271 patients with recent myocardial infarction: do stopped trials contribute to overall knowledge? Ital Heart J. 2000; 1(12):810–820. PMID: 11302109.

31. Serruys PW, Patrick WJ, de Feyter P, Macaya C, Kokott N, Puel J. Lescol Intervention Study Investigators (LIPS). Fluvastatin for prevention of cardiac events following successful first percutaneous coronary intervention: a randomized controlled trial. JAMA. 2002; 287(24):3215–3222. https://doi.org/10.1001/jama.287.24.3215 PMID: 12076217

32. Heart Protection Study Collaborative Group. MRC/BHF Heart Protection Study of cholesterol lowering with simvastatin in 20536 high-risk individuals: a randomised placebo-controlled trial. Lancet. 2002; 360(9326):7–22. https://doi.org/10.1016/S0140-6736(02)09327-3 PMID: 12114036.

33. Shepherd J, Blauw GJ, Murphy MB, Bollen EL, Buckley BM, Cobbe SM, et al. Pravastatin in elderly individuals at risk of vascular disease (PROSPER): a randomised controlled trial. Lancet. 2002; 360(9346):1623–1630. https://doi.org/10.1016/s0140-6736(02)11600-x PMID: 12457784.

34. The ALLHAT Officers and Coordinators for the ALLHAT Collaborative Research Group. Major outcomes in moderately hypercholesterolemic, hypertensive patients randomized to pravastatin vs usual care: The Antihypertensive and Lipid-Lowering Treatment to Prevent Heart Attack Trial (ALLHAT-LLT). JAMA. 2002; 288(23):2998–3007. https://doi.org/10.1001/jama.288.23.2998 PMID: 12479764

35. Sever PS, Dahlof B, Poulter NR, Wedel H, Beevers G, Caulfield M, et al. Prevention of coronary and stroke events with atorvastatin in hypertensive patients who have average or lower-than-average cholesterol concentrations, in the Anglo-Scandinavian Cardiac Outcomes Trial-Lipid Lowering Arm

(ASCOT-LLA): a multicentre randomised controlled trial. Lancet. 2003; 361(9364):1149–1158. https://doi.org/10.1016/S0140-6736(03)12948-0 PMID: 12686036.

36. Holdaas H, Fellstrom B, Jardine AG, Holme I, Nyberg G, Fauchald P, et al. Effect of fluvastatin on cardiac outcomes in renal transplant recipients: a multicentre, randomised, placebo-controlled trial. Lancet. 2003; 361(9374):2024–2031. https://doi.org/10.1016/S0140-6736(03)13638-0 PMID: 12814712.

37. Colhoun HM, Betteridge DJ, Durrington PN, Hitman GA, Neil AW, Livingstone SJ, et al. Primary prevention of cardiovascular disease with atorvastatin in type 2 diabetes in the Collaborative Atorvastatin Diabetes Study (CARDS): multicentre randomised placebo-controlled trial. Lancet. 2004; 364(9435):685–696. https://doi.org/10.1016/S0140-6736(04)16895-5 PMID: 15325833.

38. Koren MJ, Hunninghake DB. Clinical outcomes in managed-care patients with coronary heart disease treated aggressively in lipid-lowering disease management clinics: the alliance study. J Am Coll Cardiol. 2004; 44(9):1772–1779. https://doi.org/10.1016/j.jacc.2004.07.053 PMID: 15519006.

39. Wanner C, Krane V, Marz W, Olschewski M, Mann FE, Ruf G, et al. Atorvastatin in patients with type 2 diabetes mellitus undergoing hemodialysis. N Engl J Med. 2005; 353(3):238–248. https://doi.org/10.1056/NEJMoa043545 PMID: 16034009.

40. Knopp RH, d'Emden M, Smilde JG, Pocock SJ, Colwell J, Schork A, et al. Efficacy and safety of atorvastatin in the prevention of cardiovascular end points in subjects with type 2 diabetes. The Atorvastatin Study for Prevention of Coronary Heart Disease Endpoints in non-insulin-dependent diabetes mellitus (ASPEN). Diabetes Care. 2006; 29(7):1478–1485. https://doi.org/10.2337/dc05-2415 PMID: 16801565

41. Nakamura H, Arakawa K, Itakura H, Kitabatake A, Goto Y, Toyota T, et al. Primary prevention of cardiovascular disease with pravastatin in Japan (MEGA Study): a prospective randomised controlled trial. Lancet. 2006; 368(9542):1155–1163. https://doi.org/10.1016/S0140-6736(06)69472-5 PMID: 17011942.

42. Ridker PM, Genest J, Boekholdt SM, Libby P, Gotto AM, Nordestgaard BG, et al. HDL cholesterol and residual risk of first cardiovascular events after treatment with potent statin therapy: an analysis from the JUPITER trial. Lancet. 2010; 376(9738):333–339. https://doi.org/10.1016/S0140-6736(10)60713-1 PMID: 20655105.

43. GISSI-HF investigators. Effect of rosuvastatin in patients with chronic heart failure (the GISSI-HF trial): a randomised, double-blind, placebo-controlled trial. Lancet. 2008; 372(9645):1231–1239. https://doi.org/10.1016/S0140-6736(08)61240-4 PMID: 18757089.

44. Fellstrom BC, Jardine AG, Schmieder RE, Holdaas H, Bannister K, Beutler J, et al. Rosuvastatin and cardiovascular events in patients undergoing hemodialysis. N Engl J Med. 2009; 360(14):1395–1407. https://doi.org/10.1056/NEJMoa0810177 PMID: 19332456.

45. Brok J, Gluud LL, Gluud C. Ribavirin plus interferon versus interferon for chronic hepatitis C. Cochrane Database Syst Rev. 2005; 2:CD005445. https://doi.org/10.1002/14651858.CD005445 PMID: 16034976

46. Andreone P, Cursaro C, Gramenzi A, Fiorino S, Di Giammarino L, Miniero R, et al. Interferon alpha plus ketoprofen or interferon alpha plus ribavirin in chronic hepatitis C non-responder to interferon alpha alone: results of a pilot study. Ital J Gastroenterol Hepatol. 1999; 31(8):688–94. PMID: 10730561.

47. Andreone P, Gramenzi A, Cursaro C, Sbolli G, Fiorino S, Di Giammarino L, et al. Interferon-alpha plus ribavirin in chronic hepatitis C resistant to previous interferon-alpha course: results of a randomized multicenter trial. J Hepatol. 1999; 30(5):788–93. https://doi.org/10.1016/s0168-8278(99)80130-5 PMID: 10365803.

48. Barbaro G, Di Lorenzo G, Soldini M, Giancaspro G, Bellomo G, Belloni G, et al. Interferon-alpha-2B and ribavirin in combination for chronic hepatitis C patients not responding to interferon-alpha alone: an Italian multicenter, randomized, controlled, clinical study. Am J Gastroenterol. 1998; 93(12):2445–51. https://doi.org/10.1111/j.1572-0241.1998.00702.x PMID: 9860407.

49. Barbaro G, Di Lorenzo G, Belloni G, Ferrari L, Paiano A, Del Poggio P, et al. Interferon alpha-2B and ribavirin in combination for patients with chronic hepatitis C who failed to respond to, or relapsed after, interferon alpha therapy: a randomized trial. Am J Med. 1999; 107(2):112–8. https://doi.org/10.1016/s0002-9343(99)00160-6 PMID: 10460040.

50. Barbaro G, Di Lorenzo G, Soldini M, Giancaspro G, Pellicelli A, Grisorio B, et al. Evaluation of long-term efficacy of interferon alpha-2b 2b and ribavirin in combination in naive patients with chronic hepatitis C: an Italian multicenter experience. Ribavirin-Interferon in Chronic Hepatitis Italian Group Investigators. J Hepatol. 2000; 33(3):448–55. https://doi.org/10.1016/s0168-8278(00)80281-0 PMID: 11020001.

51. Bell H, Hellum K, Harthug S, Myrvang B, Ritland S, Maeland A, et al. Treatment with interferon-alpha2a alone or interferon-alpha2a plus ribavirin in patients with chronic hepatitis C previously treated with interferon-alpha2a. Scand J Gastroenterol. 1999; 34(2):194–8. https://doi.org/10.1080/00365529950173087

52. Bellobuono A, Mondazzi L, Tempini S, Silini E, Vicari F, Idéo G. Ribavirin and interferon-alpha combination therapy vs interferon-alpha alone in the retreatment of chronic hepatitis C: a randomized clinical trial. J Viral Hepat. 1997; 4(3):185–91. https://doi.org/10.1046/j.1365-2893.1997.00142.x PMID: 9181527.

53. Bellobbuono A, Mondazzi L, Tempini S, Chiodo F, Magliano E, Furione L, et al. Early addition of ribavirin to interferon in chronic hepatitis C not responsive to interferon monotherapy. J Hep. 2000; 33:463–8. https://doi.org/10.1016/s0168-8278(00)80283-4 PMID: 11020003.

54. Berg T, Hoffmann RM, Teuber G, Leifeld L, Lafrenz M, Baumgarten R, et al. Efficacy of a short-term ribavirin plus interferon alfa combination therapy followed by interferon alfa alone in previously untreated patients with chronic hepatitis C: a randomized multicenter trial. Liver. 2000; 20(6):427–36. https://doi.org/10.1034/j.1600-0676.2000.020006427.x PMID: 11169056

55. Boucher EJ, Jacquelinet S, Canva V, Turlin B, Jacquelinet C, Colimon R, et al. High rate of long-term virological response after a 1-year course of interferon +/- ribavirin in chronic hepatitis C relapsers. Results of a 191 patients randomized trial. Liver Int. 2003; 23(4):255–61. https://doi.org/10.1034/j.1600-0676.2003.00836.x PMID: 12895265

56. Bresci G, Parisi G, Bertoni M, Capria A. High-dose interferon plus ribavirin in chronic hepatitis C not responding to recombinant alpha-interferon. Dig Liver Dis. 2000; 32(8):703–7. https://doi.org/10.1016/s1590-8658(00)80334-5 PMID: 11142581

57. Brillanti S, Garson J, Foli M, Whitby K, Deaville R, Masci C, et al. A pilot study of combination therapy with ribavirin plus interferon alfa for interferon alfa-resistant chronic hepatitis C. Gastroenterology. 1994; 107(3):812–7. https://doi.org/10.1016/0016-5085(94)90131-7 PMID: 7521308.

58. Brouwer JT, Nevens F, Bekkering FC, Bourgeois N, Van Vlierberghe H, Weegink CJ, et al. Reduction of relapse rates by 18-month treatment in chronic hepatitis C. A Benelux randomized trial in 300 patients. J Hepatol. 2004; 40(4):689–95. https://doi.org/10.1016/j.jhep.2003.12.017 PMID: 15030987.

59. Chemello L, Cavalletto L, Bernardinello E, Guido M, Pontisso P, Alberti A. The effect of interferon alfa and ribavirin combination therapy in naive patients with chronic hepatitis C. J Hepatol. 1995; 23 Suppl 2:8–12. PMID: 8720287.

60. Davis GL, Esteban-Mur R, Rustgi V, Hoefs J, Gordon SC, Trepo C, et al. Interferon alfa-2b alone or in combination with ribavirin for the treatment of relapse of chronic hepatitis C. N Engl J Med. 1998; 339 (21):1493–9. https://doi.org/10.1056/NEJM199811193392102 PMID: 9819447.

61. Dettmer R, Reinus JF, Clain DJ, Aytaman A, Levendoglu H, Bloom AA, et al. Interferon-alpha-2b for retreatment of chronic hepatitis C. Hepatogastroenterology. 2002; 49:758–63. PMID: 12063985.

62. Ferenci P, Stauber R, Steindl-Mundi P, Gschwantler M, Fickert P, Datz C, et al. Treatment of patients with chronic hepatitis C not responding to interferon with high-dose interferon alpha with or without ribavirin: final results of a prospective randomised trial. Eur J Gastroenterol Hepatol. 2001; 13(6):699–705. https://doi.org/10.1097/00042737-200106000-00014 PMID: 11434597

63. Fried MW, Shiffman ML, Reddy KR, Smith C, Marinos G, Goncales F Jr, et al. Peginterferon alfa-2a plus ribavirin for chronic hepatitis C viral infection. N Engl J Med. 2002; 347(13):975–82. https://doi.org/10.1056/NEJMoa020047 PMID: 12324553.

64. Lédinghen V, Trimoulet P, Winnock M, Foucher J, Bourliérec M, Desmorat H, et al. Daily or three times a week interferon alfa-2b in combination with ribavirin or interferon alone for the treatment of patients with chronic hepatitis. J Hepatol. 2002; 36(5):672–80. https://doi.org/10.1016/s0168-8278(02)00026-0 PMID: 11983451.

65. Lédinghen VL, Trimoulet P, Bernard PH, Bourliere M, Portal I, Rémy AJ, et al. Daily or three times per week interferon alpha-2b combination with ribavirin or interferon alone for the treatment of patients with chronic hepatitis C not responding to previous interferon alone. J Hepatol. 2002; 36(6):819–26. https://doi.org/10.1016/s0168-8278(02)00071-5 PMID: 12044534.

66. Malik AH, Kumar KS, Malet PF, Ostapowicz G, Adams G, Wood M, et al. A randomised trial of high-dose interferon alpha-2b, with or without ribavirin, in chronic hepatitis C patients who have not responded to standard dose interferon. Aliment Pharmacol Ther. 2002; 16(3):381–8. https://doi.org/10.1046/j.1365-2036.2002.01201.x PMID: 11876690

67. Mangia A, Villani MR, Minerva N, Leandro G, Bacca D, Cela M, et al. Efficacy of 5MU of interferon in combination with ribavirin for naive patients with chronic hepatitis C virus: a randomised controlled trial. J Hepatol. 2001; 34(3):441–6. https://doi.org/10.1016/s0168-8278(00)00024-6 PMID: 11322207.

68. Di Marco DV, Ferraro D, Almasio P, Vaccaro A, Parisi P, Cappello M, et al. Early viral clearance and sustained response in chronic hepatitis C: a controlled trial of interferon and ribavirin after high-dose interferon induction. J Viral Hepat. 2002; 9(5):354–9. https://doi.org/10.1046/j.1365-2893.2002.00370.x PMID: 12225330

69. McHutchison JG, Gordon SC, Schiff ER, Shiffman ML, Lee WM, Rustgi VK, et al. Interferon alfa-2b alone or in combination with ribavirin as initial treatment for chronic hepatitis C. N Engl J Med. 1998; 339(21):1485–92. https://doi.org/10.1056/NEJM199811193392101 PMID: 9819446.

70. Milella M, Santantonio T, Pietromatera G, Maselli R, Casalino C, Mariano N, et al. Retreatment of non-responder or relapser chronic hepatitis C patients with interferon plus ribavirin vs interferon alone. Ital J Gastroenterol Hepatol. 1999; 31(3):211–5. PMID: 10379482

71. Pockros PJ, Reindollar R, McHutchinson J, Reddy R, Wright T, Boyd DG, et al. The safety and tolerability of daily infergen plus ribavirin in the treatment of naive chronic hepatitis C patients. J Viral Hepat. 2003; 10(1):55–60. https://doi.org/10.1046/j.1365-2893.2003.00402.x PMID: 12558913

72. Pol S, Couzigou P, Bourlière M, Abergel A, Combis JM, Larrey D, et al. A randomized trial of ribavirin and interferon-alpha vs. interferon-alpha alone in patients with chronic hepatitis C who were non-responders to a previous treatment. Multicenter Study Group under the coordination of the Necker Hospital, Paris, France. J Hepatol. 1999; 31(1):1–7. https://doi.org/10.1016/s0168-8278(99)80157-3 PMID: 10424277.

73. Pol S, Nalpas B, Bourlière M, Couzigou P, Tran A, Abergel A, et al. Combination of ribavirin and interferon-alfa surpasses high doses of interferon-alfa alone in patients with genotype-1b-related chronic hepatitis. Hepatology. 2000; 31(6):1338–44. https://doi.org/10.1053/jhep.2000.8089 PMID: 10827161.

74. Portal I, Bourliere M, Halfon P, De Ledinghen V, Couzigou P, Bernard PH, et al. Retreatment with interferon and ribavirin vs interferon alone according to viraemia in interferon responder-relapser hepatitis C patients: a prospective multicentre randomized controlled study. J Viral Hepat. 2003; 10(3):215–23. https://doi.org/10.1046/j.1365-2893.2003.00426.x PMID: 12753341

75. Poynard T, Marcellin P, Lee SS, Niederau C, Minuk GS, Ideo G, et al. Randomised trial of interferon alpha2b plus ribavirin for 48 weeks or for 24 weeks versus interferon alpha2b plus placebo for 48 weeks for treatment of chronic infection with hepatitis C virus. Lancet. 1998; 352(9138):1426–32. https://doi.org/10.1016/s0140-6736(98)07124-4 PMID: 9807989.

76. Reichard O, Norkrans G, Fryden A, Braconier JH, Sonnerborg A, Weiland O. Randomised, double-blind, placebo-controlled trial of interferon alpha-2b with and without ribavirin for chronic hepatitis C. The Swedish Study Group. Lancet. 1998; 351(9096):83–7. https://doi.org/10.1016/s0140-6736(97)06088-1 PMID: 9439491.

77. Salmeron J, Ruiz-Extremera A, Torres C, Rodriguez-Ramos L, Lavin I, Quintero D, et al. Interferon versus ribavirin plus interferon in chronic hepatitis C previously resistant to interferon: a randomized trial. Liver. 1999; 19(4):275–80. https://doi.org/10.1111/j.1478-3231.1999.tb00049.x PMID: 10459624.

78. Scotto G, Fazio V, Tantimonaco G. Pilot study of a short course of ribavirin and alpha interferon in the treatment of chronic active hepatitis C not responding to alpha-interferon alone. Ital J Gastroenterol. 1996; 28(9):505–11. PMID: 9131395.

79. Scotto G, Campanozzi F, D'Adduzio A, Grimaldi M, Fazio V. Interferon-alpha (IFN alpha) Daily Dose Versus IFN Alpha plus Ribavirin for Treatment-Naive Chronic Hepatitis C Patients Infected by Genotype 1b. BioDrugs. 2003; 17(4):281–6. https://doi.org/10.2165/00063030-200317040-00007 PMID: 12899645

80. Senturk H, Ersoz G, Ozaras R, Kaymakoglu S, Bozkaya H, Akdogan M, et al. Interferon-alpha2b induction treatment with or without ribavirin in chronic hepatitis C: a multicenter, randomized, controlled trial. Dig Dis Sci. 2003; 6:1124–9. https://doi.org/10.1023/a:1023725014751 PMID: 12822874

81. Shiffman ML, Hofmann CM, Gabbay J, Luketic VA, Sterling RK, Sanyal AJ, et al. Treatment of chronic hepatitis C in patients who failed interferon monotherapy: effects of higher doses of interferon and ribavirin combination therapy. Am J Gastroenterol. 2000; 95(10):2928–35. https://doi.org/10.1111/j.1572-0241.2000.02321.x PMID: 11051370

82. Sostegni R, Ghisetti V, Pittaluga F, Marchiaro G, Rocca G, Borghesio E, et al. Sequential versus concomitant administration of ribavirin and interferon alfa-n3 in patients with chronic hepatitis C not responding to interferon alone: results of a randomized, controlled trial. Hepatology 1998; 28(2):341–6. https://doi.org/10.1002/hep.510280208 PMID: 9695995.

83. Toccaceli F, Grimaldi M, Rosati S, Palazzini E, Laghi V. Ribavirin plus human leucocyte interferon alpha for the treatment of interferon resistant chronic hepatitis C: a controlled trial. Hepatol Res. 1997; 8(2):106–12. https://doi.org/10.1016/S1386-6346(97)00056-9

84. Trippi S, Di Gaetano G, Soresi M, Cartabellotta F, Vassallo R, Carroccio A, et al. Interferon-alfa alone versus interferon-alfa plus ribavirin in patients with chronic hepatitis C not responding to previous interferon-alfa treatment. BioDrugs. 2000; 13(4):299–304. https://doi.org/10.2165/00063030-200013040-00008 PMID: 18034536

85. Verbaan HP, Widell HEA, Bodeson TL, Lindgren SC. High sustained response rate in patients with histologically mild (low grade and stage) chronic hepatitis C infection. A randomised, double blind, placebo-controlled trial of interferon alpha-2b with or without ribavirin. Eur J Gastroenterol Hepatol. 2002; 14:627–33. https://doi.org/10.1097/00042737-200206000-00007 PMID: 12072596

86. Scholz FW, Stephens MA. K-sample Anderson-Darling Tests. J Am Stat Assoc. 1987; 82(399):918–924. https://doi.org/10.2307/2288805

87. Dan-Yu L, Sullivan PF. Meta-Analysis of Genome-wide Association Studies with Overlapping Subjects. Am J Hum Genet. 2009; 85(6): 862–872. https://doi.org/10.1016/j.ajhg.2009.11.001 PMID: 20004761

88. Han B Duong D, Sul JH, de Bakker PI, Eskin E, Raychaudhuri S. A general framework for meta-analyzing dependent studies with overlapping subjects in association mapping. Hum Mol Gen. 2016; 25 (9):1857–1866. https://doi.org/10.1093/hmg/ddw049 PMID: 26908615

89. R Development Core Team. *R*: *a language and environment for statistical computing*. R Foundation for Statistical Computing: Vienna, Austria, 2009. ISBN 3-900051-07-0, URL http://www.R-project.org.

90. Egger M, Smith GD, Altman D. Systematic Reviews in Health Care: Meta-analysis in context. 2nd ed. London: John Wiley and Sons BMJ Books; 2001.