







## Bias and unclear outcomes in clinical trials of diabetic retinopathy: a cross-sectional analysis of literature

## Viés e desfechos confusos em ensaios clínicos de retinopatia diabética: uma análise transversal da literatura

Vania Mozetic<sup>1</sup>   
Valéria Mozetic de Barros<sup>2</sup>   
Lucas Denadai<sup>3</sup> 

Matheus Ferreira Santos da Cruz<sup>4</sup>   
Natasha Ferreira Santos da Cruz<sup>5</sup>   
Nilva Simeren Bueno de Moraes<sup>6</sup> 

<sup>1</sup>Instituto Dante Pazzanese de Cardiologia (São Paulo). São Paulo, Brazil. vaniamozetic@gmail.com

<sup>2</sup>Instituto de Cardiologia Dante Pazzanese (São Paulo), Faculdade de Medicina do ABC (São André), Universidade Federal de São Paulo (São Paulo). São Paulo, Brazil. mozeticvaleria@gmail.com

<sup>3</sup>Corresponding author. Universidade Federal de São Paulo (São Paulo). São Paulo, Brazil. lucasdenadai@hotmail.com

<sup>4</sup>Universidade do Estado do Pará (Belém). Pará, Brazil. matheusfcruz@gmail.com

<sup>5-6</sup>Universidade Federal de São Paulo (São Paulo). São Paulo, Brazil. natasha\_cruz27@hotmail.com, nilva.bueno@unifesp.br

**ABSTRACT | BACKGROUND:** Clinical trials are well-designed papers that aim to answer questions in the real world. However, sometimes they present missing, dubious, and unclear outcomes challenging to apply in practice. **OBJECTIVE:** This work aims to evaluate the way and frequency with which the outcomes in randomized clinical trials of intervention in diabetic retinopathy can be presented in an unclear way to readers. Analyze how these dubious presentations can lead to misinterpretations, why this happens and how they can be remedied. **METHODS AND MATERIALS:** We searched for RCT about DR intervention in PubMed published over the past five years. **RESULTS:** Seventy RCTs were included, 27 in peripheral diabetic retinopathy (PDR) and 43 in diabetic macular edema (DME). In the DME group, we found 25.6% reporting and publication bias, 34.9% subjective outcomes, 44.1% presented a lack of presentation of the baseline, and 51.1% underreporting adverse events. In the PDR group, we found 29.6% reporting and publication bias, 44.4% subjective outcomes, 14.8% presented a lack of baseline, and 62.9% underreporting adverse events. **CONCLUSION:** In addition to the resulting bias, we found other forms of publication of unclear RCT outcomes on DR. Most of them occurred due to disrespect for CONSORT parameters. The reader must be attentive to recognize them and know how they can influence the data's interpretation.

**KEYWORDS:** Outcome Bias. Diabetic Retinopathy. Diabetic Macular Edema. Evidence-based medicine. Literature review.

**RESUMO | INTRODUÇÃO:** Os ensaios clínicos são artigos com desenhos bem específicos que visam responder a perguntas do mundo real. Porém, às vezes apresentam desfechos ausentes, duvidosos e confusos que dificultam a aplicação prática. **OBJETIVO:** O objetivo deste trabalho é avaliar o modo e a frequência com que os desfechos em ensaios clínicos randomizados de intervenção em retinopatia diabética podem ser apresentados de forma pouco clara aos leitores. Fazer uma análise de como estas apresentações duvidosas podem levar a erros de interpretação, por que isso acontece e como podem ser sanadas. **MÉTODOS:** Realizamos uma pesquisa por ECR sobre RD no PubMed publicado nos últimos cinco anos. **RESULTADOS:** Setenta ECR foram incluídos, 27 na retinopatia diabética periférica (PDR) e 43 no edema macular diabético (DME). No grupo DME encontramos 25,6% de viés de notificação e viés de publicação; 34,9% desfechos subjetivos, 44,1% apresentaram ausência de apresentação da linha de base e 51,1% subnotificação de eventos adversos. No grupo PDR, encontramos 29,6% de viés de notificação e publicação; 44,4% dos resultados subjetivos, 14,8% apresentaram uma falta de apresentação da linha de base e 62,9% subnotificação de eventos adversos. **CONCLUSÃO:** Além do viés de resultado, encontramos outras formas de publicações de desfechos duvidosos em ECR sobre RD. A maioria deles ocorreu devido ao desrespeito aos parâmetros do CONSORT. O leitor deve estar atento para reconhecê-los e saber como eles podem influenciar na interpretação dos dados.

**PALAVRAS-CHAVE:** Viés de desfecho. Retinopatia diabética. Edema macular diabético. Medicina baseada em evidências. Revisão da literatura.

Submitted 11/19/2020, Accepted 03/25/2021, Published 06/04/2021

J. Evid-Based Healthc., Salvador, 2021;3:e3415

<http://dx.doi.org/10.17267/2675-021Xevidence.2021.e3415>

ISSN: 2675-021X

Assigned editor: Luis Claudio Correia

*How to cite this article:* Mozetic V, Barros VM, Denadai L, Cruz MFS, Cruz NFS, Moraes NSB. Bias and unclear outcomes in clinical trials of diabetic retinopathy: a cross-sectional analysis of literature. J Evid-Based Healthc. 2021;3:e3415. <http://dx.doi.org/10.17267/2675-021Xevidence.2021.e3415>



## Introduction

The number of people worldwide impacted by diabetic retinopathy (DR) was predicted to increase from 2.6 million in 2015 to 3.2 million in 2020.<sup>1</sup> With the increase in the number of cases, there will be an increase in the demand for more effective treatments by the decision-maker who finds answers in randomized controlled trials (RCTs). Randomized clinical trials are the appropriate study design to assess the impact of an intervention.<sup>2</sup>

The global number of registered clinical trials increased fivefold between 2004 and 2013.<sup>3</sup> The RCT occupies one of the highest stages in the evidence pyramid and is immediately below the systematic review.<sup>4</sup> It is the most important study project to answer questions in the real world.

However, clinical trials take a long time, require large teams and high economic costs<sup>5</sup>, leading the author to simplify a search by performing biases as generating inadequate randomization, allocation errors, not masking patients and teams evaluators, reporting incomplete or selective outcomes, among other systematized errors<sup>6</sup> that can lead the reader to errors in the evaluation of outcomes.<sup>7</sup> However, there are other dubious forms of publication outcomes beyond bias that can also lead the decision-maker to make mistakes.<sup>8</sup>

This article aims to present some biased and unclear ways of presenting the outcomes in RCT on DR to alert readers about how and why they appeared, how they can influence the interpretation of outcomes, and how they can be avoided in the future.

## Methods

It is a cross-sectional analysis of literature developed at the Department of Ophthalmology, Universidade Federal de São Paulo (UNIFESP), São Paulo, Brazil. We conducted a search strategy according to table 1 for PubMed clinical trials with a filter over the past five years without language restrictions or impact factor of the journals.

**Table 1.** Search strategy

N	Search Strategy
#1	"Diabetic Retinopathy"[Mesh]
#2	(Diabetic Retinopathies) or (Retinopathies, Diabetic) or (Retinopathy, Diabetic)
#3	#1 or #2

We searched PubMed on 08 April 2020 and found 288 articles. Two authors independently selected the only RCT about intervention in DR. Duplicates papers were removed.

We only consider RCT in parallel design in humans of any sex, age and follow-up that has received any treatment for diabetic retinopathy, surgery, laser or intravitreal injection in the intervention group and any intervention or no intervention in the control group.

Two independent authors selected articles by title and abstract using web application Rayyan<sup>9</sup> according to the inclusion criteria, and the discrepancies were solved by consensus. The selected papers were read in total, and primary and secondary outcomes and adverse events were captured and classified according to the criteria proposed by Heneghan et al.<sup>8</sup>, as we will see below.

Poorly chosen outcomes or study design flaws are surrogates, composite and subjective outcomes, complex scales, and lack of relevance for patients and decision-makers.

Surrogates outcomes: Are substitute markers for assessing an outcome with a strong association with the outcome of genuine interest. They are used to infer a direct result related to the outcome sought, usually cheaper and quicker to observe than the clinical result of real interest.<sup>10</sup>

Composite outcomes: The use of combined outcome measures as an association of combined signs or symptoms or multiple causes to culminate in a result.<sup>11</sup> Its use decreases the sample size as it increases the chance of achieving some outcome, consequently decreasing the cost of the study.

Subjective results: Occur when there is a need for the observer's judgment on a result. There is no objective measure or when the result is self-reported by the subject.<sup>12</sup>

Complex scales: Combinations of symptoms and signs forming a scale determined by the author and not standardized.<sup>13</sup>

The lack of relevance for patients and decision-makers: They are presenting results that do not have a practical purpose either for the patient or for the decision-maker. These are outcomes with little or no applicability to the real world.<sup>14</sup>

According to poorly collected outcomes, it fails in methods that are missing data and poorly specified outcomes.

Missing data: It is the attrition bias of the incomplete outcomes data. It is considered a loss of data in relation to the baseline. The 'rule 5 and 20' (ie, if > 20% of missing data, then the study is highly biased; if < 5%, then low risk of bias) exists to help the reader understand the missing size.<sup>8,16</sup>

Poorly specified results: It is essential to clarify in the publication protocol and methodology how the problem will be defined, how the evaluation will be technically made, and how the outcomes will be considered, as poorly specified results can confuse. Self-report measures are particularly prone to bias due to their subjectivity.<sup>12,15</sup>

Selectively reported results or problems related to publication are publication bias, report bias, and underreporting of an adverse event.

Publication bias: It is an actual bias. It is directly related to the registration of the protocol before starting the registration of the participants<sup>16</sup> determined the study design in detail.

Reporting bias: It is an actual bias, systematized error, and, as such, interferes with mixed results of clinical trials. Selective bias occurs when a study is published, but some of the measured and analyzed results have not been reported. It is directly related to the publication of the protocol that will serve as a parameter of the research's real interest regardless of the results obtained.<sup>16</sup>

Underreporting of adverse events: It is characterized by the absence of mention of any form of adverse events or side effects, mild, moderate, and severe, that are beneficial or not.<sup>8</sup>

We can observe relative measures, spin, multiplicity, and core outcomes sets according to inappropriately interpreted outcomes.

Relative measures. The outcome should always be presented in relative and absolute numbers as determined by CONSORT.<sup>18</sup> This way, they can be recalculated, giving credibility to the study, dates can be used to be meta-analyzed according to others authors' needs and so that they are not overvalued given the strictly relative presentation.

Spin results are characterized by the exaltation of discourse in the abstract that does not match the data obtained or even try to divert the reader by presenting secondary results or subgroup analyses and not the direct result, or to focus the reader on another study objective far from the result statistically not significant and also a multiplicity of results particularly when there are several times of evaluation of the outcome.<sup>8</sup>

Core outcomes set when observing multiple outcomes that can take the central focus of the research. It considers the set of outcomes of real interest to the patient that must be evaluated. Much like the lack of relevance, we will be considering the set of outcomes<sup>8</sup> and the need for interpretation.

The results were compiled in Table 2 and presented in a descriptive analysis evaluating the results.

**Table 2.** Summary of findings

	Design					Methods		Publication			Interpretation			
	Surrogate	Composite	Subjective	Complex scale	Lack of relevance	Missing data	Poorly specified	Publication bias	Reporting bias	Underreporting Adverse events	Relative measures	Spin	Multiplicity	Core outcomes sets
<b>DME (43)</b>	0	3	15	2	3	2	16	11	11	22	22	4	6	0
<b>Peripheral DR (27)</b>	0	1	12	6	3	5	14	8	8	17	14	4	0	0
<b>DR (70)</b>	0	4	27	8	6	25	30	19	19	39	36	8	6	0

DME: diabetic macular edema; DR: diabetic retinopathy

## Results

Seventy randomized clinical trials were found, 27 (38.6%) addressing peripheral diabetic retinopathy (PDR) and 43 (61.4%) diabetic macular edema (DME); 35 (50%) multi-centre, 17 (24.2%) single centre and 18 (25.7%) could not be identified in this criteria. 23 (32.9%) clinical trials did not present a protocol registration; three registered in the oriental language, one registered in a university database, but with public domain, and two registered after data collection. Four trials assessed the quality of life (QoL) (four, 5.7%) remains low, and 13 trials do not measure visual acuity (VA) (18.6%). Twenty-three trials were published in magazines A1 (impact factor  $\geq 3.8$ ), 13 in A2 ( $3.799 \geq \text{impact factor} \geq 2.5$ ), 10 in B1 ( $2.499 \geq \text{impact factor} \geq 1.3$ ), 22 in B2 ( $1.299 \geq \text{impact factor} \geq 0.001$ ) and two without impact factor.

Table 2 shows the classification of outcomes according to the bias.

### DME outcomes

In the 43 studies addressing diabetic macular edema, no surrogate outcomes were found; composite outcomes three (6.9%); 15 (34.9%) subjective outcomes (nine due to individual analysis unit and six uncertain analysis unit). Two (4.7%) complex scale; three (6.9%) outcomes with lack of relevance to patients and decision-makers; one (2.3%) trial missing data and 19 (44.1%) it is impossible to affirm due to the lack of presentation of the baseline characterizing poorly specified data; 11 (25.6%) publication bias, that is, without protocol registration and two of which were registered late, after the collection of results began as well as reporting bias because they depend on the protocol registration in an appropriate period, the others did not present selective reporting of the data; 22 (51.2%) underreporting of adverse events and relative measures; four (9.3%) spin; six (13.9%) multiplicity and zero core outcome sets inconsistencies.

### PDR outcomes

In the 27 clinical trials of peripheral diabetic retinopathy, no one presented surrogate outcomes,

one (3.7%) composite outcome; 12 (44.4%) subjective outcomes; six (22.2%) complex scales and three (11.1%) lack of relevance. One (3.7%) trial presents missing data, and four (14.8%) it is impossible to affirm due to the lack of presentation of the baseline. 14 (51.9%) poorly specified results, eight (29.6%) publication bias (one of them with a late protocol record) as well as reporting bias which is directly dependent on the protocol record, 17 (62.9%) underreporting of adverse events, 14 (51.9%) with only relative measures, four (14.8%) spin; four (14.8%) multiplicity and no core outcome sets inconsistencies.

## Discussion

The primary purpose of all medical science is to guarantee the quality of life for humans. For the RCT to fulfill this function, the results must be presented in a relevant, appropriate, and essential way for patients in the real world. In a clinical trial, bias is a dangerous practice with or without intent in collecting, analyzing, interpreting, publishing, or reviewing data that can lead to conclusions that are different from the truth.

According to Heneghan et al.<sup>8</sup>, there are dubious ways of reporting outcomes in RCT that, as well as bias, can lead to misinterpretation.

In the present study, we can observe several of these unclear ways of presenting the outcomes. We separated the analysis into central diabetic retinopathy (diabetic macular disease, DME) and peripheral diabetic retinopathy (outside the great vascular arches) because some treatments and outcomes are observed differently.

In this study's sample, we did not find surrogate outcomes in either macular studies or peripheral retinal studies, probably because there are well-determined methods for evaluating well-known outcomes in diabetic retinopathy.<sup>18</sup> When surrogate outcomes occur, caution is required when interpreting, as it provides less direct relevant evidence than the relevant results.<sup>19</sup>

Three (6.9%) and one (3.7%) composite outcomes were found in DME and peripheral DR, respectively. An advantage of using it is to increase statistical efficiency making the study shorter and cheaper<sup>11</sup>; however, it can lead to exaggerated treatment estimates.<sup>13</sup> It is a practice with high prevalence in cardiological studies. Suspicion for the occurrence of composite outcomes would be in pathology with multiple clinical signs and symptoms. Its occurrence imposes the need for data imputation, which still does not present a safe rule and ends up incurring errors.

In this sample, we found subjective outcomes in 15 (34.9%) DME trials and 12 (44.4%) of peripheral DR. Most of them occur due to the unit of analysis be considered the individual. The presentation of subjective outcomes is usually overvalued and tends to give positive results, and papers with positive results tend to be more published.<sup>12,15,21,22</sup>

A complex scale is another form to present unclear outcomes characterized by creating a scale composed of multiple signs and symptoms not standardized. In this sample, we observed two (4.7%) trials in DME and six (22.2%) in peripheral DR. Studies without validated scales usually give more positive results, and these are more likely to be published because it induces more favorable results that are always more publishable.<sup>15</sup> The creation of scales for the classification of diabetic retinopathy is not justified since it is well established and fulfills proposing treatment and determining prognosis.

An observation for ophthalmological studies is appropriate here. Different standardized scales can measure visual acuity (VA); Snellen and LogMar are two more commons in scientific publications.<sup>23</sup> It cannot be classified as presenting a perverted presentation of outcome since the scales are standardized, known, and quickly transcribed between them, but they need to choose the logarithmic scale as the ideal form for publications persists (24). In the DME sample, two (4.7%) studies assessed VA by Snellen, 38 (88.3%) studies assessed by Log Mar, and three (7%) studies did not assess VA. In the peripheral DR test (3.7%) measured vision by Snellen, 16 (59.3%) tests in Log Mar.

Lack of relevance to patients and decision-makers occurs when the laboratory markers do not present

clinical correlation in outcome important to patients or decide in real-world, although every study has its importance. In our sample, we considered the lack of relevance in 6 trials, three in each subgroup, as the measurement of serum markers which, although essential for research, are of little practical interest to the patient.

VA, QoL (quality of life), comfort with no pain matters are absolute values for the patient in ophthalmological studies. Only four trials assessed QoL (5.7%), which means keeping very low, but the most surprising is the presence of studies that do not measure VA is 13 (18.6%). The end of all science is the human being and what matters to the individual is well-being, so in ophthalmological studies, one of the most important outcomes is VA and QoL. We cannot forget it.

Missing data was found in one (2.3%) trial on DME sample and 19 (44.1%) it is impossible to affirm due to the lack of presentation of the baseline or inadequate baseline, featuring 20 (46.5%) poorly specified data. One (3.7%) trial in peripheral DR sample with missing data and four (14.8%) is impossible to affirm due to the baseline's lack of presentation. The missing data is considered a bias.<sup>16</sup> a systematic error that affects all outcomes. Data loss throughout the study is frequent, particularly in extended studies, so if > 20% missing data, then the study is highly biased; if <5%, then a low risk of bias. Evaluating the outcome compared to the baseline is a determining factor to prove the study's result; therefore, its non-presentation raises suspicion.<sup>24</sup>

Poorly specified outcomes are another distortion of the methodology and occur due to the lack of an adequate description of the outcome, the intervention, the evaluation technique, and how the study methodology and protocols will be treated. Our sample found 16 (37.2%) in DME and 14 (51.6%) in the peripheral DR sample.

It is worth remembering that the unit of analysis defined as an individual in ophthalmological studies is restricted to economic analysis, systemic effects, adverse events, and QoL. It must always be clearly defined in the methodology. The analysis unit in ophthalmological studies must always be defined and shown as the eye.<sup>25</sup>

In the analysis of this study's DME, two trials presented both units of analysis in the same study, seven trials used the individual, and two remained undetermined. In the peripheral DR sample, six trials presented the outcomes with the unit of analysis individual, and the other six do not make clear requiring interpretation by the reader.

Publication bias is a real bias related to the lack of publication of the study protocol. It was found in 11 (25.6%) in DME, and 8 (29.6%) in our sample's peripheral DR. The purpose of the protocol record is to show that the author's original idea was not disturbed after the results were collected.<sup>26</sup> The lack of a protocol record is not justified as it is an easy practice and without economic cost.

Reporting bias is a real bias too. It is the publication of only selected outcomes and not all predicted outcomes. It is only possible to know when the protocol is published and the lack of protocol registration. It occurred in 11 (25.6%) and eight (29.6%) in DME and peripheral DR, respectively. Evaluating the outcome compared to the initial research interest is a determining factor in proving the study's result if it is being presented without hiding bad results; therefore, the presentation of the protocol gives the publication veracity and confidence.<sup>26</sup>

In our sample, we could observe that two records of PDR and four of DME were registered in a different platform than [clinicaltrials.gov](http://clinicaltrials.gov), but all of them are public. Two protocols registers in DME trials and one in peripheral DR trial were done after results. Several years have passed from complaints of lack of registration, and the problem persists.<sup>27</sup>

It is surprising the underreporting of adverse events observed in this work, 55.7%. Twenty-two trials (51.1%) in DME, 17 did not mention having evaluated any adverse events, 11 of these trials evaluated the use of VEGF, five assessed strictly ocular events. 17 (62.9%) in PDR, 15 did not mention having evaluated any adverse events, four of these trials evaluated VEGF, two assessed strictly ocular events. Studies have shown that journal publications underestimate side effects and exaggerate treatment benefits when assessing the risk-benefit ratio.<sup>27,28</sup>

In the interpretation block, we find 22 (50%) of the results in DME trials and 14 (51.8%) trials in PDR of results presented only in a relative way, which can exaggerate interpretation. As indicated by CONSORT (17), the results must be presented in an absolute and relative way, proving the results' clarity. Also, 25 (35.7%) trials presented the author's report as a trial with insufficient sample size, a problem that could be solved by systematic review. For that, the trials need to present the outcomes in the absolute and relative way to facilitate the reviewer work.

In the case of the spin evaluation, there were no misleading titles in our sample, but four (9.3%) spin in the macula trials and four (14.8%) in the peripheral retina, all related to the loss of focus of the primary outcome presenting subgroup analysis and secondary outcomes. Spin shifts the reader's focus on another objective of the study other than the primary outcome, keeping it away from the statistically insignificant result.<sup>29</sup>

All works that did not present a protocol registration may have incurred a multiplicity error since the original proposal's lack of access frees the author to enrich the study with other analyses. In addition to the lack of registration, we found six (13.9%) and four (14.8%) trials of the macula and non-central retina respectively, with illusory graphics without scale and unnecessary photographs taking the author's focus away. The risk of multiplicity lies in the greater the number of outcomes, the greater the chance of false-positive results and unfounded claims of effectiveness.<sup>30</sup>

Already proposed by COMET<sup>31</sup>, core outcome sets bring together relevant resources to facilitate the development of a proposal with a limited number of outcomes that the reader does not get lost. In our sample, we do not observe bad core set outcomes.

24 trials show limited follow-up, and 25 studies with a deficiency in the sample size reported by the author. Systematic reviews can help in the sample size, compiling tests with the same PICO (Population, Intervention, Comparative, Outcome), but it is necessary to have uniformity in the technical analysis, to be added in a single meta-analysis.

This study aimed to show how the resulting bias and other dubious presentations of outcomes can appear in RCT on DR and alert readers to read the articles with more criteria. Many of these biases have already been reported in the literature<sup>32</sup> and could be corrected with strict compliance with CONSORT rules.<sup>17</sup> We observed the persistence of a lack of standardization in assessing outcomes, which makes it challenging to carry out systematic reviews.

One of the limitations of this paper was not to have limited RCT analysis published by the impact factor of the journals. On the other hand, it was not our objective to evaluate the quality of publications but rather evaluate the quality of the RCTs found in the literature that has guided decision-makers conduct. The impact factor is another metric that the reader should view with great care because it is multifactorial and influences inattentive readers; this would be a proposal for future research. The message that remains is to be very careful when reading scientific articles.

## Conclusion

This study showed that the outcome biases and other unclear forms of presentation outcomes in RCT about DR are frequent and that most would not exist if there were strict compliance with the parameters dictated by CONSORT. Given the risk in understanding that these errors can cause, the reader must be alert not to be deceived when making a decision.

## Author contributions

Mozetic V and De Barros V conceived and designed the study. All authors participated in data acquisition and analysis, drafting the report, revising and approving the content in its final version.

## Competing interests

No financial, legal, or political competing interests with third parties (government, commercial, private foundation, etc.) were disclosed for any aspect of the submitted work (including but not limited to grants, data monitoring board, study design, manuscript preparation, statistical analysis, etc.).

## References

1. Flaxman SR, Bourne RRA, Resnikoff S, Ackland P, Braithwaite T, Cicinelli MV, et al. Global causes of blindness and distance vision impairment 1990-2020: a systematic review and meta-analysis. *Lancet Glob Health*. 2017;5(12):e1221-34. [https://doi.org/10.1016/S2214-109X\(17\)30393-5](https://doi.org/10.1016/S2214-109X(17)30393-5)
2. Evans SR. Fundamentals of clinical trial design. *J Exp Stroke Transl Med*. 2010;3(1):19-27. <https://doi.org/10.6030/1939-067X-3.1.19>
3. Viergever RF, Li K. Trends in global clinical trial registration: an analysis of numbers of registered clinical trials in different parts of the world from 2004 to 2013. *BMJ Open* 2015;5(9):e008932. <https://doi.org/10.1136/bmjopen-2015-008932>
4. Murad MH, Asi N, Alsawas M, Alahdab F. New evidence pyramid. *Evid Based Med*. 2016;21(4):125-7. <http://dx.doi.org/10.1136/ebmed-2016-110401>
5. Moore TJ, Zhang H, Anderson G, Alexander GC. Estimated Costs of Pivotal Trials for Novel Therapeutic Agents Approved by the US Food and Drug Administration, 2015-2016. *JAMA Intern Med*. 2018;178(11):1451-7. <https://doi.org/10.1001/jamainternmed.2018.3931>
6. Higgins JPT, Thomas J, Chandler J, Cumpston M, Li T, Page MJ, et al, editors. *Cochrane Handbook for Systematic Reviews of Interventions*. Chichester UK: Wiley-Blackwell; 2020. Available from: <https://training.cochrane.org/handbook/archive/v6.1>
7. Smith CT, Hickey H, Clarke M, Blazeby J, Williamson P. The trials methodological research agenda: results from a priority setting exercise. *Trials*. 2014;15(32). <https://doi.org/10.1186/1745-6215-15-32>
8. Heneghan C, Goldacre B, Mahtani, KR. Why clinical trial outcomes fail to translate into benefits for patients. *Trials*. 2017;18(122). <https://doi.org/10.1186/s13063-017-1870-2>
9. Ouzzani M, Hammady H, Fedorowicz Z, Elmagarmid A. Rayyan — a web and mobile app for systematic reviews. *Syst Rev*. 2016;5:210. <https://doi.org/10.1186/s13643-016-0384-4>
10. Twaddell S. Surrogate outcome markers in research and clinical practice. *Aust Prescr [Internet]*. 2009;32(2):47-50. Available from: <https://www.nps.org.au/assets/4de02c581d38e5f6-8ed90bde43e6-0a7fd02d9378b4a1811c30b83618ecb16c7fca111484ef2055ed4940f4e9.pdf>
11. Freemantle N, Calvert M, Wood J, Eastaugh J, Griffin C. Composite outcomes in randomized trials: greater precision but with greater uncertainty? *JAMA*. 2003;289(19):2554-9. <https://doi.org/10.1001/jama.289.19.2554>



12. Moustgaard H, Bello S, Miller FG, Hróbjartsson A. Subjective and objective outcomes in randomized clinical trials: definitions differed in methods publications and were often absent from trial reports. *J Clin Epidemiol*. 2014;67(12):1327–34. <https://doi.org/10.1016/j.jclinepi.2014.06.020>
13. Hobart JC, Cano SJ, Zajicek JP, Thompson AJ. Rating scales as outcome measures for clinical trials in neurology: problems, solutions, and recommendations. *Lancet Neurol*. 2007;6(12):1094–105. [https://doi.org/10.1016/S1474-4422\(07\)70290-9](https://doi.org/10.1016/S1474-4422(07)70290-9)
14. Fogel DB. Factors associated with clinical trials that fail and opportunities for improving the likelihood of success: A review. *Contemp Clin Trials Commun*. 2018;11:156-64. <https://dx.doi.org/10.1016%2Fj.conctc.2018.08.001>
15. Sackett DL. Bias in analytic research. *J Chronic Dis*. 1979;32(1-2):51–63. [https://doi.org/10.1016/0021-9681\(79\)90012-2](https://doi.org/10.1016/0021-9681(79)90012-2)
16. Sterne JAC, Savović J, Page MJ, Elbers RG, Blencowe NS, Boutron I, et al. RoB 2: a revised tool for assessing risk of bias in randomised trials. *BMJ*. 2019;366:l4898. <https://doi.org/10.1136/bmj.l4898>
17. Schulz KF, Altman DG, Moher D, CONSORT Group. CONSORT 2010 statement: updated guidelines for reporting parallel group randomised trials. *BMJ*. 2010;340:c332. <https://doi.org/10.1136/bmj.c332>
18. Aronson JK. Biomarkers and surrogate endpoints. *Br J Clin Pharmacol*. 2005;59(5):491-4. <https://dx.doi.org/10.1111%2Fj.1365-2125.2005.02435.x>
19. Yudkin JS, Lipska KJ, Montori VM. The idolatry of the surrogate. *BMJ*. 2011;343:d7995. <https://doi.org/10.1136/bmj.d7995>
20. Ferreira-González I, Busse JW, Heels-Ansdell D, Montori VM, Akl EA, Bryant DM, et al. Problems with use of composite end points in cardiovascular trials: systematic review of randomised controlled trials. *BMJ*. 2007;334(7597):786. <https://doi.org/10.1136/bmj.39136.682083.ae>
21. Song F, Parekh S, Hooper L, Loke YK, Ryder J, Sutton AJ, et al. Dissemination and publication of research findings: an updated review of related biases. *Health Technol Assess*. 2010;14(8):iii, ix-xi, 1-193. <https://doi.org/10.3310/hta14080>
22. Hopewell S, Loudon K, Clarke MJ, Oxman AD, Dickersin K. Publication bias in clinical trials due to statistical significance or direction of trial results. *Cochrane Database Syst Rev*. 2009;(1):MR000006. <https://doi.org/10.1002/14651858.mr000006.pub3>
23. Elliott DB. The good (logMAR), the bad (Snellen) and the ugly (BCVA, number of letters read) of visual acuity measurement. *Ophthalmic Physiol Opt*. 2016;36(4):355–8. <https://doi.org/10.1111/opo.12310>
24. Nunes EV, Pavlicova M, Hu MC, Campbell AN, Miele G, Hien D, et al.. Baseline matters: the importance of covariation for baseline severity in the analysis of clinical trials. *Am J Drug Alcohol Abuse*. 2011;37(5):446-52. <https://doi.org/10.3109/00952990.2011.596980>
25. Higgins JPT, Green S, editors. *Cochrane Handbook for Systematic Reviews of Interventions* Version 5.1.0 [Internet]. The Cochrane Collaboration; 2011. Available from: <https://handbook-5-1.cochrane.org/>
26. Aslam A, Imanullah S, Asim M, El-Menyar A. Registration of Clinical Trials: Is it Really Needed?. *N Am J Med Sci*. 2013;5(12):713-5. <https://doi.org/10.4103/1947-2714.123266>
27. Wieseler B, Wolfram N, McGauran N, Kerekes MF, Vervölgyi V, Kohlepp P, et al. Completeness of reporting of patient-relevant clinical trial outcomes: comparison of unpublished clinical study reports with publicly available data. *PLoS Med*. 2013;10(10):e1001526. <https://doi.org/10.1371/journal.pmed.1001526>
28. Hart B, Lundh A, Bero L. Effect of reporting bias on meta-analyses of drug trials: reanalysis of meta-analyses. *BMJ*. 2012;344:d7202. <https://doi.org/10.1136/bmj.d7202>
29. Boutron I, Dutton S, Ravaud P, Altman DG. Reporting and interpretation of randomized controlled trials with statistically nonsignificant results for primary outcomes. *JAMA*. 2010;303(20):2058–64. <https://doi.org/10.1001/jama.2010.651>
30. Sun X, Briel M, Busse JW, You JJ, Akl EA, Mejza F, et al. Credibility of claims of subgroup effects in randomised controlled trials: systematic review. *BMJ*. 2012;344:e1553. <https://doi.org/10.1136/bmj.e1553>
31. COMET initiative. *Core Outcome Measures in Effectiveness Trials* [Internet]. [cited 2020 May 15]. Available from: <http://www.comet-initiative.org/Studies>
32. Mozetic V, Leonel L, Pacheco RL, Latorraca COC, Guimarães T, Logullo P, et al. Reporting quality and adherence of randomized controlled trials about statins and/or fibrates for diabetic retinopathy to the CONSORT checklist. *Trials*. 2019;20(1):729. <https://doi.org/10.1186/s13063-019-3868-4>