

The Art of Reading a Scientific Article

A arte de ler um artigo científico

Luís Cláudio Lemos Correia¹, Gabriela Oliveira Bagano²

¹Corresponding author. BAHIANA – School of Medicine and Public Health, São Rafael Hospital, D'Or Institute for Research and Education. Salvador, Bahia, Brazil. ORCID: 0000-0002-6910-1366. luis.correia@bahiana.edu.br

²BAHIANA – School of Medicine and Public Health. Salvador, Bahia, Brazil. ORCID: 0000-0002-6541-0372. gabrielabagano14.2@bahiana.edu.br

I do not appreciate much the bureaucratic and simplistic connotation of the terms "guide" or "checklist" for analysis of scientific evidence. Reading an article requires further than "checking" for details. Checklist is for an airplane that will take off or surgery that will start. Scientific interpretation should not be based on checklists. It requires a vision of the whole, involvement with the subject, intuitive perception brought by experience, capacity of reflection, a balance between skepticism and pragmatism. It is like reading a book or watching a movie, our interpretation is not just a checklist.

Imagine a movie critic who analyzes movies by the structured form of a checklist. It would not reach its goal. On the other hand, in the analysis of scientific evidence we cannot have artistic freedom to interpret as we wish, as it would be an invitation for confirmation bias of our beliefs. We need to control our personal biases during reading; we need to know how to look for illusory components of the studies, to notice when noises are mixed with signs; we must follow scientific principles. These principles should guide us as a compass.

Reading an article is not a passive process. More than reading, it is a data-based thinking process. We think first, then look for the answer in the work. We should build up a mental framework of how much that evidence will influence our thinking.

With these statements I do not want to convey the idea that the interpretation of evidence is difficult or laborious. On the contrary, reading an article should be a light, natural, and even fun activity.

As I always propose, the article reading begins before the article. It begins within us, with the search for personal biases that avert a rational analysis. We need to prepare ourselves mentally for the analysis of external evidence, because it may come against our internal perspectives. This preparation goes through the questioning of what are our internal evidences (beliefs) about the subject that is exposed in the title of the work of interest. This process is a proposal of mental preparation.

Continuing with the pre-article reflection, we need to know what the article is about. For this we will have an initial contact with the content of the work, by reading only two sentences of the abstract: purpose and conclusion. These are the pillars of our analysis. The purpose shows the hypothesis tested or the reality described. The conclusion reflects what the author wants us to believe as the final message of the article. At this point we will exercise skepticism, because the reading of the article must be an attempt to refute the conclusion. A rigorous analysis of noise (bias and random) versus signal (truth). A low signal-to-noise ratio will make us doubt the veracity of the work conclusion.

Reading the conclusion also gives us an important insight into how much the author is biased. I refer to the "spin" attitude, when the author recognizes a negative primary result, but then creates a positive trend by presenting a secondary result¹. For example, "in the overall analysis, there was no difference between groups in left ventricular ejection fraction [primary outcome], but there was improvement in troponin [secondary outcome]." Spin is a strong marker of bias. Detection from the beginning will make us more attentive.

Also in the pre-article reading, it is worth noticing if (1) the study is an initiative from the manufacturer of the product tested, (2) it is carried out by an independent group, but receives manufacturer's cost aid or (3) has no financing relationship with the manufacturer. This numbering represents a scale of the degree of involvement between the manufacturer and the scientific work. Theoretically, the smaller the involvement, the more exempt would be the study. Recent work published in the *Annals of Internal Medicine* has shown that the greater the link with the pharmaceutical industry, the more use of inadequate statistical methods aiming to achieve the desired result².

We must also investigate if the authors have a conflict of interest, whether related to industry, to intellectual conflicts or to professional activity³. Independent of industry ties, there is the conflict of interest led by the very nature of the professional activity of the authors. It may be even a greater conflict than that of those who receive industry sponsorship. For example, works that test accuracy of diagnostic methods or treatment

efficacy may be performed by professionals whose primary activity is tied to the method or treatment being tested. Intellectual conflict of interest, on the other hand, occurs when an author is part of a legion of scholars and believers of a subject.

I do not propose here to invalidate the study if it has an interested funding source or a conflict of interest. But we must increase our awareness to the risk of bias. A Cochrane's systematic review shows that work with conflict of interest is more likely to deliver positive results than works without manufacturer financing⁴, a bias that should not exist. It is interesting to notice that, from Cochrane's traditional bias risk assessment, this study did not detect a greater risk of bias in studies with a conflict of interest. Actually, industry-funded studies are "well done" according to a superficial methodological assessment: they follow the standard methodology (sample size, randomization, blind). The noises arising from those works are usually detected by the type of reading we propose, something that goes beyond a checklist present in systematic reviews.

So far we have first analyzed our bias as a reader and the bias of the author. The next step is the analysis of the pre-test probability of the study hypothesis. The scientific experiment is the test, while the pre-test is the knowledge that exists prior to the work. There are two components of the pre-test probability: the plausibility of the hypothesis (logical mechanism, knowledge of basic science that supports this mechanism) and previous works that tested the hypothesis.

This stage is important because the predictive value of a study depends not only on the study, but also on the pre-test probability of the idea suggested by the study. By the end of an article, we want to know how likely that conclusion is to be true. This also depends on the pre-test quality of the idea.

Studies whose result proposes meaningless ideas will have low positive predictive value, independent of the strength of evidence. Just as negative works will have lower negative predictive value when the idea is very promising. This is the Bayesian perspective that I advocate for the process of critical analysis of an evidence. Evidence should not be analyzed in the vacuum, but rather in the context of the idea.

The next step will be the search of the previous records of the protocol, with the intention of figuring out whether: the protocol was published a priori (databases of protocols or protocol articles)? The a priori protocol coincides with the methods published in the article or there have been changes a posteriori? Answers "yes" to the first question and "no" to the second question will bring credibility to the results. A recent search in the clinicaltrials.org records showed that 30% of the studies change the primary outcome defined a priori. These changes substantially reduce the predictive value of the study⁵.

We finally came to reading the article. Now we will look for systematic errors (biases) and random errors (random) that produce noise and illusory results. This reading of the article comes down to "methods" and "results".

A parenthesis: "introduction" and "discussion" do not need to be read. I hardly ever read. Not because it is bad to read, but my laziness does not allow me to waste time with the unnecessary. But if you read, do not simply believe (faith) neither in the author's arguments for the idea behind the work (introduction), nor in his interpretation of results (discussion). Do not read believing, read doubting. Even if in the end you agree with the conclusion.

We have to know what to look for in methods and results. What to look for depends on the purpose of the study: treatment efficacy, diagnostic accuracy, prognostic accuracy. The analysis also depends on whether the study is positive or negative. The direction of the article reading is different depending on the direction of the result. Different things produce false-negatives or false-positives.

Another interesting question is the different level of attention we should have in methods and results. In methods we predominantly evaluate systematic errors. In the results we evaluate random errors (P-value).

We detect potential biases when we notice flaws in the study design described in the methods.

However, it is in the results that we will see the degree of statistical significance or the imprecision of the confidence intervals, aspects that represent

the risk of random error. In the results we will see if the sample complied with the assumptions the sample size calculation described in the methods. For example, the calculation (methods) is perfect and supposes a certain incidence of the outcome, but we will see in results that this premise was not obeyed. Therefore, this would be a work with greater risk of random error. It will also be in the results that we will see deviations of scientific integrity that predispose to random error: the study was interrupted early because the result appeared positive (truncated), the author begins to value description of secondary outcomes or subgroup analysis (reporting bias).

After all this analysis, in the absence of protocol changes and when we observe low risk of bias or random error, we will conclude that we have a study with a high level of evidence.

Finally, the final answer will be Bayesian: we must calculate the post-test probability. When we have a moderately probable hypothesis a priori, a low risk of bias and random error study is sufficient to generate a high positive predictive value. In case of unlikely hypothesis, a good study will have moderate positive predictive value and we require a second study to confirm the idea (reproducibility).

In the case of a negative study, when to give up the idea? Normally, if the hypothesis is not promising, we tend to think that it is no longer worth insisting on. But if the idea is very likely, we will need more than one negative study to make us give up the concept.

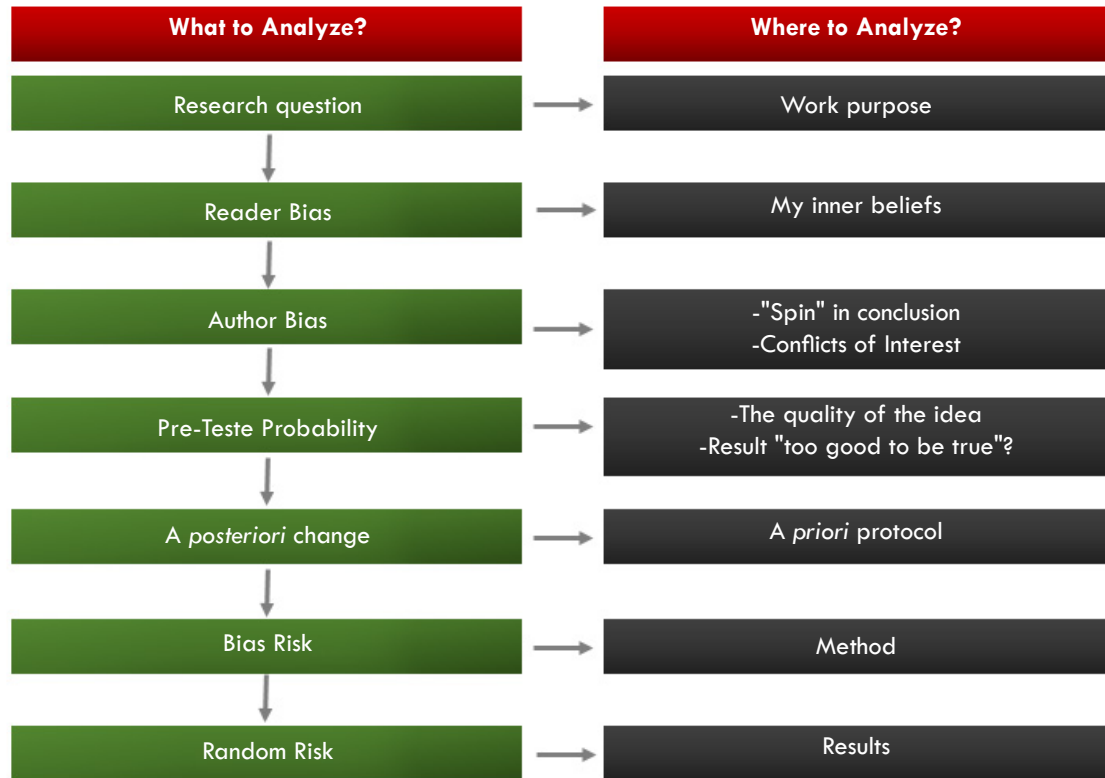
Science does not prove non-existence, science proves existence of concepts. Since it is impossible to prove that something does not exist, what is at stake in the analysis of a negative study is whether it should make us give up trying the idea or whether we should continue trying to prove it. We must evaluate when we are being stubborn or when it would be deemed wise to simply abandon an idea. This depends on the negative predictive value of the study, which depends on the pre-test probability and the signal-to-noise ratio of the study.

In summary, we must evaluate – Figure 1:

- Bias of the reader (my inner beliefs)
- Intrinsic bias of the author (spin)

- Extrinsic bias of the author (conflict of interest)
- Pre-test probability of the idea
- Subsequent protocol changes
- Study signal-to-noise ratio (methods and results)
- And finally, we will mentally calculate the predictive value of the work

Figure 1. Process of article interpretation



I call this exercise a compass to critical reading, because there are no accurate maps or GPS when it comes to true scientific thinking. Our compasses will stimulate a thoughtful, dogma-free, and a thinking based on uncertainty, change and cause.

References

1. Boutron I, Altman DG, Hopewell S, Vera-Badillo F, Tannock I, Ravaud P. Impact of spin in the abstracts of articles reporting results of randomized controlled trials in the field of cancer: the SPIIN randomized controlled trial. *J Clin Oncol.* 2019;32(36):4120-6. doi: [10.1200/JCO.2014.56.7503](https://doi.org/10.1200/JCO.2014.56.7503)
2. Schwartz LM, Woloshin S, Zheng E, Tse T, Zarin DA. ClinicalTrials.gov and Drugs@FDA: A comparison of results

reporting for new drug approval trials. *Ann Intern Med.* 2016;165(6):421-430. doi: [10.7326/M15-2658](https://doi.org/10.7326/M15-2658)

3. Bou-Karroum L, Hakoum MB, Hammoud MZ, Khamis AM, Al-Gibbawi M, Badour S et al. Reporting of Financial and Non-financial Conflicts of Interest in Systematic Reviews on Health Policy and Systems Research : A Cross Sectional Survey. *Int J Health Policy Manag.* 2018;7(8):711-7. doi: [10.15171/ijhpm.2017.146](https://doi.org/10.15171/ijhpm.2017.146)

4. Lundh A, Lexchin J, Mintzes B, Schroll JB, Bero L. Industry sponsorship and research outcome (review). *Cochrane Database Syst Rev.* 2017;(2). doi: [10.1002/14651858.MR000033.pub3](https://doi.org/10.1002/14651858.MR000033.pub3)

5. Delgado AF, Delgado AF. Outcome switching in randomized controlled oncology trials reporting on surrogate endpoints : a cross- sectional analysis. *Sci Rep.* 2017;7(1):9206. doi: [10.1038/s41598-017-09553-y](https://doi.org/10.1038/s41598-017-09553-y)