European Journal of Physical and Rehabilitation Medicine 2022 December;58(0):888-91
DOI: 10.23736/S1973-9087.22.07767-X

SPECIAL ARTICLE

Trust/untrust is not the same as true/false. Lessons learned and ethical questions on the application of untrustworthiness scales to judge individuals

Stefano NEGRINI [1, 2] *, Maria G. CERAVOLO [3], Giorgio FERRIERO [4, 5], the EJPRM International Review Commission ‡

[1]Department of Biomedical, Surgical and Dental Sciences, University "La Statale", Milan, Italy; [2]IRCCS Istituto Ortopedico Galeazzi, Milan, Italy; [3]Department of Experimental and Clinical Medicine, "Politecnica delle Marche" University, Ancona, Italy; [4]Physical and Rehabilitation Medicine Unit, Scientific Institute of Tradate, IRCCS, Istituti Clinici Scientifici Maugeri, Tradate, Varese, Italy; [5]Department of Biotechnology and Life Sciences, University of Insubria, Varese, Italy

‡Members are listed at the end of the paper

*Corresponding author: Stefano Negrini, Department of Biomedical, Surgical and Dental Sciences, University "La Statale", Milan, Italy.
E-mail: stefano.negrini@unimi.it

ABSTRACT

This special paper reflects on trustworthiness and its implications for scientific medical journals and all the communities they serve: health professionals, policymakers, the public, and a specific discipline, in our case, Physical and Rehabilitation Medicine. We start from a recent episode: a paper claimed the untrustworthiness of two randomised controlled trials (RCTs) published in the *European Journal of Physical and Rehabilitation Medicine* based on a newly developed trustworthiness scale, used until now only in systematic reviews. This likely represents the first case of applying such a scale focusing on a single leading author. Developing a proper answer to this case led us to present some insights from the perspective of a Journal editor. We discuss the impact of false research results, why trust is needed in science and medicine, the difference between untrust and false results, the problems in judging trustworthiness, the unfortunately weak capacity of the peer review system in preventing these issues, the problems of "post-hoc" judgements and the emerging ethical issues. We conclude with some suggestions for the future based on prevention at the system level.

(*Cite this article as*: Negrini S, Ceravolo MG, Ferriero G; the EJPRM International Review Commission. Trust/untrust is not the same as true/false. Lessons learned and ethical questions on the application of untrustworthiness scales to judge individuals. Eur J Phys Rehabil Med 2022;58:888-91. DOI: 10.23736/S1973-9087.22.07767-X)

KEY WORDS: Ethics; Serial publications; Trust.

The facts

In the last few months, our journal has suffered a case of potential untrustworthiness, raised by a paper published in the journal *Pain*[1] (PJP). The PJP is worth reading to understand better this special paper. In synthesis, while producing a Cochrane systematic review (SR), the authors found three Randomised Controlled Trials (RCTs) reported by a specific group of authors (SGA) with results considerably different from any other RCT published on the same topic. For this reason, they excluded these RCTs from the metanalysis. Due to the high impact of these RCTs

on guidelines and SRs production, the PJP authors decided to check other papers by the SGA on the same topic using a newly developed trustworthiness scale. PJP found ten RCTs by the SGA, applied a trustworthiness scale they developed to be used within SRs production, tested this scale probably for the first time to judge a series of papers produced by the SGA, and suggested the untrustworthiness of all the RCTs. The *European Journal of Physical and Rehabilitation Medicine* (EJPRM) published two RCTs included in this list.[2, 3] As a second step, the PJP authors decided to extend the check beyond authors "testing" the journals' reactions. The theory is that the interest of all involved parties can cause untrustworthiness. Authors are driven by the "publish or perish" need and can emphasize, when not even fake their results, partly or completely. Editors pursue the Impact Factor and don't care much about what they publish, provided it is citable and well-written.

The PJP authors made all the journals that published the SGA paper aware of their publication and asked to be informed of what we decided to do. We started a formal process of verification of the RCTs published in EJPRM. This journal is published in Italy, it is managed by the Italian Society of PRM, and the two Chief-Editors are Italian, like the involved SGA. To avoid any conflict of interest, we appointed an International Review Commission (IRC) including highly-reputed Editorial Board members with a non-Italian majority. Under the leadership of the senior Deputy Editor, Maria Gabriella Ceravolo (Ita), the IRC members were Rolf Frischknecht (Ch), Jaro Karppinen (Fin), Ayse A. Küçükdeveci (Tur), and Mauro Zampolini (Ita).

The IRC analysed the main characteristics of papers[2, 3] in face of the application of the Cochrane Pregnancy and Childbirth review group's Trustworthiness Screening Tool (CPC-TST), described by O'Connell et al,[1] and the criteria for retracting the papers according to the Committee on Publication Ethics (COPE) Guidelines.[4]

After the IRC inquiry for further clarifications, the SGA decided to retract one of the two RCTs[3] they published in EJPRM because they found mistakes in the data collection.[5] The IRC supported this decision. Concerning the other paper,[2] the IRC concluded that it did not fulfil the COPE Guidelines indications for paper retractions[4] because "there is no clear evidence that the findings are unreliable, or that they are a result of plagiarism, or have been published elsewhere. The only concern is about the unavailability of prior trial registration and formal approval by the Local Ethics Committee. However, this is not enough, at present, to ask for the retraction of the paper." The IRC added: "Looking at the risks that allegations of untrustworthiness may be raised in the future, the EJPRM Commission advises the Journal Editors to adopt a stricter publication policy, requesting the authors to provide evidence of the Ethics Committee approval as a mandatory step before starting the peer review process."

We support the SGA decision and agree with the IRC judgment. The fate of the other SGA RCTs published in the other journals is not homogeneous. Two journals retracted their papers,[6, 7] and one reached the opposite decision on one paper[8] because the peer-review process was adequate, the results were exciting, and "discrepancy is a master stone in science" (personal communication). A fourth journal is still in the decisional phase.

Discussion

The science philosopher Karl Popper proposed the Falsification Principle[9] to distinguish science from non-science. A theory is scientific when it can be tested and eventually proven false. From another perspective, the principle could also suggest the possibility of producing false results that further tests could reject. The fake production of scientific results has already been shown by some famous examples, from cold fusion[10] to the efficacy of hydroxychloroquine for COVID-19.[11, 12] The consequences of such falsifications are heavy for everybody: authors, journals, the scientific community, and the public.

Trust in science is everything. Trust and untrust do not correspond to true and false: if you do not trust a discovery, you are not saying that it is wrong – you only doubt that it is true. Trust and untrust are also not simple questions of disagreement. We can disagree on concepts and interpretations (introduction and discussion) but trust facts (methods and results). Untrust means there are reasons to suspect that facts were not properly collected and managed, whether for inaccuracy or fraud. Trust is even more important in medicine because it relates to the health of human beings. At the EJPRM, we are fully aware of the importance of trust in what we publish.

Trust is fundamental because publishing false results in scientific journals impacts people. Specifically, medical journals serve a few communities, with the most important being the farthest ones: patients and the public. We reach them through health professionals and eventually policymakers – they shape their decisions on the results we publish, and what they do impacts patients and the public. We also serve the communities of authors and scientists; specifically, the EJPRM serves the field of Physical and Rehabilitation Medicine. Publication of false results offends all

these communities, and our journal is fully committed to preventing this from happening. Untrust is one of the ways to identify possible false results.

What journals publish is entirely based on trust in authors and not fact-checking, allowing us to "certify" true or false. The only possible test for trust is the review process. Discovering methodological mistakes or strange results/presentations/methods can raise suspicions and distrust. Unfortunately, some excellent writers perfectly mask papers' flaws behind a methodological accuracy curtain. In the end, reviewers must judge the methodology, and almost only methodological faults raise their attention unless they have very high knowledge of the topic and what is specifically studied. On one side, it is enough to know the system to write a methodologically sound paper; reviewers will probably keep their trust, and the article will be published. On the other, reviewers are already the weakest chain of the publication process. They do voluntary, unpaid, and unrewarded work to improve science. Unfortunately, most reviewers are overloaded and have to refuse, others at any stage of their careers refuse systematically, and those at the start of their careers accept to learn. In these conditions, trust check becomes even more challenging.

Judging trustworthiness is not easy, and unjust accusations offend all the communities we serve as much as the discovery of fraud. If a result is accurate but not published because of suspected untrustworthiness, we impede or delay growth and knowledge. Nevertheless, the most impacted are the authors, whose reputation is at stake. For this reason, we must be incredibly cautious without becoming corporatists – another serious mistake. We must judge authors with fairness and look at the facts. Otherwise, we would harm our most important community: patients and the public.

Nowadays, trustworthiness has become even more important for some specific papers. RCTs are the most important ones because they provide the best evidence and are synthesized in SRs, with Cochrane SRs being the reference standard. RCTs and SRs are used to produce guidelines, finally spread in all countries. The World Health Organization, health authorities, policymakers, and health professionals produce/use guidelines, SRs, and RCTs to improve evidence-based decisions impacting the health of populations and individuals. Consequently, the trustworthiness of RCTs is crucial because it will mean trustable SRs and guidelines. Cochrane is always searching for the best methods to produce the best SRs. The need to check for RCTs' trustworthiness is part of this process, and some Cochrane groups are developing untrustworthiness scales like the one used in this episode. Nevertheless, these scales provide checks that are good for SRs because they are "a posteriori" (after publication). The same scales are inappropriate for journals that need judgments "a priori" (before publication).

Lessons learned and possible solutions

We must reflect on the cure and prevention of these issues. PJP was a "post-hoc" suggestion of trustworthiness issues. We wonder if the approach was the correct one because a trustworthiness scale provides assumptions and not objective facts. Such a scale judges trust and trust is not necessarily the truth. Much progress is still needed, but we have also to recognise that we do not see real alternatives at the moment. In the end, we need a treatment for this disease caused by the "publish or perish" dogma.

Nevertheless, the ethical issues remain and should be thoroughly explored by COPE. We would suggest reducing personalization as much as possible. We support the implementation of trustworthiness scales in SRs to exclude possibly untrustworthy papers, but pointing fingers at individuals is something different. Still, individual behaviours can be the problem and must be censored. Authors should fear that they could be identified and censored. We fully recognise the ethical tension between these extremes: trust and truth, the judgment of scientific data and individual behaviours, bad data management, and fraud. Are papers like the PJP the right approach, or should we look for something different?

Pending this dilemma solution, we believe that prevention should remain the mainstream. This is particularly important for journal editors who are responsible for choosing what to publish and what not. Is prevention feasible, and by what means?

We see a couple of solutions:

- mandatory publication of data in public repositories to allow "post-hoc" double checks without passing through burdensome and sometimes unanswered requests to the authors. Such a decision would reduce the phenomenon but not avoid it. Still, the possibility of being checked by others could be a good deterrent;
- publishing the ethical committee approvals information (and not only a statement of approval) in the papers. This would also serve as a deterrent for false declarations.

The EJPRM decided to request from now on the authors the mandatory publication of data in open-access repositories and to publish in the text the reference of the ethical committee approval.

Finally, there is the usual, old way of prevention: the review process. In case of untrust, we strongly suggest reviewers reject the papers because the first communities that medical journals serve are the patients and public, not authors and the "publish or perish" dogma.

References

1. O'Connell NE, Moore RA, Stewart G, Fisher E, Hearn L, Eccleston C. de C Williams AC. Investigating the veracity of a sample of divergent published trial data in spinal pain. Pain 2022. [Epub ahead of print].

2. Monticone M, Ambrosini E, Vernon H, Rocca B, Finco G, Foti C, *et al.* Efficacy of two brief cognitive-behavioral rehabilitation programs for chronic neck pain: results of a randomized controlled pilot study. Eur J Phys Rehabil Med 2018;54:890–9.

3. Monticone M, Ambrosini E, Portoghese I, Rocca B. Multidisciplinary program based on early management of psychological factors reduces disability of patients with subacute low back pain: one-year results of a randomized controlled study. Eur J Phys Rehabil Med 2021;57:959–67.

4. COPE: Committee on Publication Ethics. Retraction guidelines; 2022 [Internet]. Available from: https://doi.org/10.24318/cope.2019.1.4 [cited 2022, Nov 3].

5. Monticone M, Ambrosini E, Portoghese I, Rocca B. Retraction of: Multidisciplinary program based on early management of psychological factors reduces disability of patients with subacute low back pain: one-year results of a randomized controlled study. Eur J Phys Rehabil Med 2022. [Epub ahead of print].

6. Monticone M, Ambrosini E, Rocca B, Cazzaniga D, Liquori V, Lovi A, *et al.* Multimodal exercises integrated with cognitive-behavioural therapy improve disability of patients with failed back surgery syndrome: a randomized controlled trial with one-year follow-up. Disabil Rehabil 2022;44:3422–9.

7. Monticone M, Ambrosini E, Rocca B, Cazzaniga D, Liquori V, Pedrocchi A, *et al.* Group-based multimodal exercises integrated with cognitive-behavioural therapy improve disability, pain and quality of life of subjects with chronic neck pain: a randomized controlled trial with one-year follow-up. Clin Rehabil 2017;31:742–52.

8. Monticone M, Ambrosini E, Rocca B, Cazzaniga D, Liquori V, Foti C. Group-based task-oriented exercises aimed at managing kinesiophobia improved disability in chronic low back pain. Eur J Pain 2016;20:541–51.

9. Popper K. The Logic of Scientific Discovery (2002 pbk; 2005 ebook ed.). London: Routledge; 1959.

10. Ondir Freire L, Alves de Andrade D. Preliminary survey on cold fusion: it's not pathological science and may require revision of nuclear theory. J Electroanal Chem (Lausanne) 2021;903.

11. Mehra MR, Desai SS, Ruschitzka F, Patel AN. RETRACTED: Hydroxychloroquine or chloroquine with or without a macrolide for treatment of COVID-19: a multinational registry analysis. Lancet 2020;S0140-6736(20)31180-6. [Epub ahead of print] [Retraction in: Lancet. 2020 Jun 5;null. Erratum in: Lancet. 2020 May 30; Erratum in: Lancet. 2020 Jul 18;396] [10245] [:e2-e3]

12. Mehra MR, Desai SS, Kuy S, Henry TD, Patel AN. Cardiovascular Disease, Drug Therapy, and Mortality in Covid-19. N Engl J Med 2020;382:e102. [Retraction in: N Engl J Med. 2020 Jun 4]

Conflicts of interest.—The authors certify that there is no conflict of interest with any financial organization regarding the material discussed in the manuscript.

Authors' contributions.—All authors read and approved the final version of the manuscript.

Group name.—Members of the group include the following (in alphabetical order): Maria G. CERAVOLO (Department of Experimental and Clinical Medicine, "Politecnica delle Marche" University, Ancona, Italy); Rolf FRISCHKNECHT (Swiss Society of Physical Medicine and Rehabilitation, Switzerland, European Board of Physical and Rehabilitation Medicine, the Nederlands); Jaro KARPPINEN (Medical Research Center Oulu, University of Oulu and Oulu University Hospital, Oulu, Finland); Ayse A. KÜÇÜKDEVECI (Faculty of Medicine, Department of Physical Medicine and Rehabilitation, Ankara University, Ankara, Turkey); Mauro ZAMPOLINI (Department of Rehabilitation, San Giovanni Battista Hospital, Foligno, Perugia, Italy).

History.—Article first published online: November 3, 2022. - Manuscript accepted: November 3, 2022. - Manuscript received: October 27, 2022.