



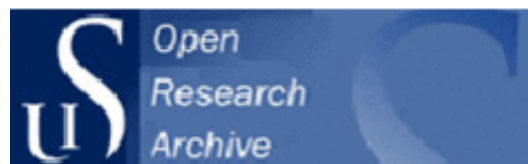
University of
Stavanger

Simon, P., Dettweiler, U. (2019) Current Anti-Doping Crisis: The Limits of Medical Evidence Employing Inductive Statistical Inference. *Sports Medicine*, 100(3), pp. 889-905.

Link to published article:

<https://doi.org/10.1007%2Fs40279-019-01074-0>

(Access to content may be restricted)



UiS Brage

<http://brage.bibsys.no/uis/>

This version is made available in accordance with publisher policies. It is the author's last version of the article after peer-review, usually referred to as post-print. Please cite only the published version using the reference above.



Current Anti-Doping Crisis: The Limits of Medical Evidence Employing Inductive Statistical Inference

Perikles Simon · Ulrich Dettweiler

1 Introduction

The anti-doping system is supposed to level the playing field and protect clean athletes. Doping scandals of the past two decades have seriously questioned the effectiveness of the worldwide anti-doping program, and criminal investigations associated with those scandals have created evidence for its partial ineffectiveness. However, legal action often succeeded because of activities from *within* the anti-doping community, such that the looming ineffectiveness could still be interpreted as a sign of an isolated shortcoming of the drug-testing program, while the overall “system of anti-doping” was still effective and working. In other words, from within the anti-doping system, we may not be able to assess its own effectiveness.

In this situation, recent scientific investigations revealing a high prevalence of doping in elite sports implicate ineffectiveness of the fight against doping in general and stress the importance of independent anti-doping research activities [1]. However, these and other similar research activities investigate praxeological aspects of anti-doping. They will depend on the level of independence and the level of cooperation extended by organized sports to the scientists involved, which will in turn limit the capabilities of such research efforts.

In this issue of *Sports Medicine*, Heuberger and Cohen publish a systematic review that could provide important alternative directions for future anti-doping research [2]. The authors argue from an epistemological point of view and question what constitutes a doping procedure in the first place, and they challenge the methodological robustness of medical evidence that has been generated to elucidate performance-enhancing effects of doping substance classes.

We will use the opportunity of the unexpected outcome generated by Heuberger and Cohen [2] to explicate the possibilities such purely non-praxeological analyses provide for the anti-doping science community. The work of Heuberger and Cohen [2] reminds us that it should in fact be scientists who draw attention to the “more or less” scientific foundation of anti-doping measures. Science is defined by the epistemological and methodological framework in which any praxeology should be situated. We will thus argue that the analysis of Heuberger and Cohen [2], especially those parts that are controversial in nature, helps us to understand why inductive reasoning, as employed by the authors, may not represent the ideal basis for a Prohibited List.

2 Epistemological Allegation: The Very Notion of (Medical) Evidence for Fighting Doping is Challenged

In 2012, the World Anti-Doping Agency (WADA) that oversees the regulatory issues associated with anti-doping set up an expert group that reported on the ineffectiveness of the drug testing program [3]. Interestingly, this report mentioned shortcomings in anti-doping only from a praxeological perspective: “The primary reason for the apparent lack of success of the testing programs does not lie with the science involved.” is one of the central positions presented in the executive summary and throughout all chapters of this report. It is noteworthy that the practical aspect of doping testing is a

level WADA does not take direct responsibility for. The report therefore leaves some room for questions to be addressed by independent scientists. What is the very nature of the science involved in anti-doping, from an epistemological and methodological perspective?

Heuberger and Cohen's contribution [2] falls into this category of research, by assessing the clinical pharmacological evidence for the performance-enhancing effect of doping substance classes included in the Prohibited List. To put the first key point of their analysis on the table, there is no sufficient scientific evidence for 18 out of the 23 substance classes. According to the WADA Code for a substance or method to be added to the Prohibited List, these 18 substance classes could also be placed on the list if they fulfilled both of the two alternative criteria, which are that they represent an actual or potential health risk to the athletes and that they violate the spirit of sports [4]. What about "scientific evidence" for these other two criteria? This has not been investigated by Heuberger and Cohen [2] and it is a provocative question because there is and cannot be a means to provide evidence for the extent of a violation of the spirit of sports, and the WADA Code also does not imply this. For ethical reasons, scientists' abilities to generate evidence for the harm aspect by virtue of randomized controlled trials are even more strictly limited than testing for performance enhancement. In other words, the list is thus for most parts based on practical and ethical considerations, and not based on scientific evidence in the classical sense. Again, it should be noted that the WADA Code is only mentioning the word "scientific evidence" with regard to the performance and the harm aspect, and it does not explicitly state that such evidence is required. The implications of this aspect will be discussed in the following section.

3 Methodological Allegation: The Current Level of Scientific Practice in Anti-Doping Science Needs Some Critical Revision

With their second key point, Heuberger and Cohen [2] address a methodological aspect of the studies that presumably demonstrated a performance-enhancing effect. Indeed, the sample sizes of the sparse 11 studies that indicated the performance-enhancing effect of five substance classes were small and the statistical power was hardly comparable to that of typical studies in clinical pharmacology. However, this is only one part of the story because we could also ask whether it is necessary or scientifically sound to apply clinical pharmacological standards to the testing of performance-enhancing effects of substance classes, as we argue below.

The review by Heuberger and Cohen [2] is partially reliant on the outcome of a study by the same group that investigated the performance-enhancing effect of erythropoietin (EPO) on cycling in a time trial and a typical mountain-stage race [5]. Interestingly, their study conducted in the typical fashion of a randomized controlled trial against placebo failed to demonstrate an effect, while suggesting effects on surrogate parameters of performance in the same manner as other previous studies. Here, it should be noted that mountain-stage races and cycling time trials are just two of the many potential endurance-related outcome measures that are of relevance to elite sports relying on various aspects of the construct of endurance. Moreover, if these outcome measures are not tested in an elite cohort subjected to factors that typically accompany EPO use, such as losing weight, as well as increasing the duration and intensity, and modifying the structure of training sessions, how could they be regarded as more valid outcome measures than, for instance, maximal oxygen consumption or submaximal performance traits related to lactate steady state measured in a laboratory test? In other words, the EPO study by Heuberger et al. [5] may have had an excellent internal validity but its external validity might have been so low that a presumably negative outcome should not be accorded great significance. Doping experts in the field certainly are aware of *some* practical advantages of administering EPO to cyclists.

The problem here is that Heuberger et al. [5] apply inductive reasoning for verifying a performance-enhancing effect of a very concrete doping procedure. This assumes that dopers behave like patients who would subsequently receive EPO prescribed under medical supervision. However, dopers do not use EPO as tested by the authors. They use a more creative concept of evidence and come to the conclusion to use EPO at dosages never tested in clinical trials in combination with other blood-boosting agents such as insulin-like growth factor 1, testosterone, and blood doping, and they modify their living and training circumstances profoundly. Thus, what is the point of conducting randomized controlled trials to presumably test the performance-enhancing effects of doping substance classes?

The rationale behind the concept of evidence is not restricted to classical inductive statistical inference, as Charles Sanders Peirce has pointed out [6]. He introduced the concept of “abduction” into modern logic. Hereby, one seeks to find the simplest and most likely explanation for observations based on studying facts and devising a theory to explain them. This discursive process, unlike deductive reasoning or inductive (statistical) inference, yields a plausible conclusion but does not positively verify it. Interestingly, this is very prominent in criminology where conclusions need be justified on often “purely” circumstantial “evidence”.

4 Some Praxeological Perspectives on Anti-Doping Science: Present and Future

When considered at the praxeological level, the review by Heuberger and Cohen [2] indicates that the scientific community dealing with anti-doping topics will now need to critically re-think the “evidence” involved in anti-doping practice and research from epistemological and methodological perspectives. We argue here that it is not reasonable to justify the Prohibited List on the basis of classical inductive statistical inference. We should in fact reconsider more critically the concept of conducting classical pharmacological studies because their external validity will in any case be low, or for most parts insignificant. Criminology is guided by concepts and frameworks such as law systems, which provide the basics for utilizing abductive reasoning, producing circumstantial evidence in a manner accepted by society. Such a framework appears to be missing in the anti-doping science field and we need to ask the anti-doping community whether the Prohibited List should be based on more than the ideas and interests of those who have the sports political power. We envision that there is some work to be done to level the playing field for compiling a Prohibited List in the future.

We want to point out that inductive statistical inference is nevertheless very essential for other aspects of anti-doping, such as for instance the quality-control issues surrounding classical doping test procedures. A recent review has drawn into question the general methodological robustness of the drug testing program [7]. The scientific community still refrains from presenting figures on the sensitivity and specificity of drug testing. By applying proper statistics, it is already evident that for certain testing procedures, such as the present EPO testing, the test outcomes must comprise a significant proportion of false-positive doping tests [7]. Here, more studies using inductive statistical inference and displaying greater transparency with respect to study outcomes are warranted.

In the future, science may need to address more rigorously the serious epistemological and methodological shortcomings in the anti-doping program. However, science should also highlight persisting misinterpretation and false communication of anti-doping measures that may endanger the integrity of sports almost as much as doping itself. For example, stating or implying that science would have provided evidence that the present anti-doping system is protecting the integrity of sports or the integrity of elite athletes as individuals is an unfounded assumption. This has never been shown.

Finally, we want to consider how the outcome of the review by Heuberger and Cohen [2] could be disseminated to the public. If we do not have classical pharmacological evidence for a performance-enhancing effect for most of the substance classes on the Prohibited List, would it not be the first and most important obligation of the anti-doping system to transport this message immediately to all athletes? As it stands now, the Prohibited List may even serve as a blueprint for persons who want to

learn what might be performance enhancing. We have to ensure we communicate that such thinking is wrong. In so doing, we contribute to a preventive measure the anti-doping community may not be exploiting enough—harm reduction.

5 Conclusions

There is a lack of evidence for a performance-enhancing effect of most of the substance classes in the Prohibited list of WADA [2]. Here, we critically discuss why classical pharmacological studies investigating monocausal drug–performance relationships will most likely fail to deliver a valid outcome. While patients are requested to follow treatment regimens that have been shown to provide beneficial effects, doping in elite sports will for most parts be conducted by treatment regimens based on abductive inference. Doping is variable in nature and often appears to involve the use of multiple substances. It is adapted to the many accompanying environmental factors, such as training, doping tests, and body weight modifications, which can be highly specific to a respective sport. Therefore, it appears to be reasonable to consider under which circumstances it is more valid to apply abductive instead of inductive statistical inference, to justify why certain substance classes are placed on the Prohibited List. Principally, evidence based on interdisciplinary and transparent abductive reasoning within a still to be defined framework could lead to a Prohibited List that serves both analytical and preventive goals.

Compliance with Ethical Standards

Funding

No sources of funding were used to assist in the preparation of this editorial.

Conflict of interest

Perikles Simon and Ulrich Dettweiler have no conflicts of interest that are directly relevant to the content of this editorial.

References

Ulrich R, Pope HG Jr, Cléret L, Petróczi A, Nepusz T, Schaffer J, et al. Doping in two elite athletics competitions assessed by randomized-response surveys. *Sports Med.* 2018;48(1):211–9.

Heuberger JAAC, Cohen AF. Review of WADA prohibited substances: limited evidence for performance-enhancing effects. *Sports Med.* 2018. <https://doi.org/10.1007/s40279-018-1014-1> (epub ahead of print).

Report to the WADA Executive Committee by the working group established following foundation board meeting of 18 May 2012. Lack of effectiveness of testing programs. Available from: <https://www.wadaama.org/sites/default/files/resources/files/2013-05-12-Lack-of-effectiveness-of-testing-WG-Report-Final.pdf>. Accessed 18 Dec 2018.

The World Anti-Doping Agency. World Anti-Doping Code 2015 with 2018 amendments. Available from: https://www.wada-ama.org/sites/default/files/resources/files/wada_antidoping_code_2018_english_final.pdf. Accessed 18 Dec 2018.

Heuberger JAAC, Cohen Tervaert JM, Schepers FM, Vliegthart AD, Rotmans JI, Daniels JM, et al. Erythropoietin doping in cycling: lack of evidence for efficacy and a negative risk-benefit. *Br J Clin Pharmacol.* 2013;75(6):1406–21.

Peirce CS. Abduction and induction. In: Buchler J, editor. *Philosophical writings of Peirce*. New York: Dover Publications; 1955. p. 5–22.

Simon P, Neuberger EW, Wang G, Pitsiladis YP. Antidoping science: important lessons from the medical sciences. *Curr Sports Med Rep*. 2018;17(10):326–31.