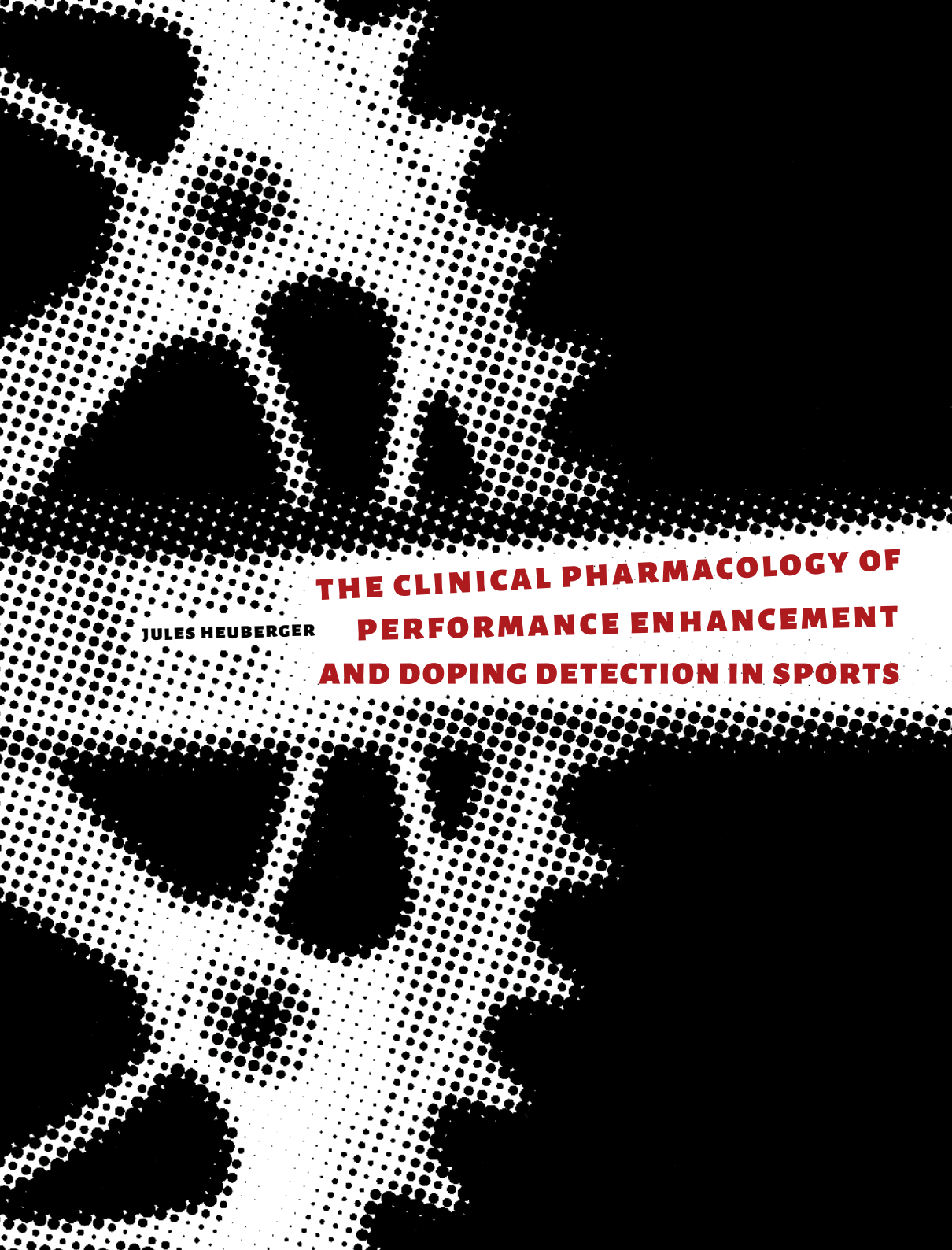

THE CLINICAL PHARMACOLOGY OF PERFORMANCE ENHANCEMENT AND DOPING DETECTION IN SPORTS
JULES HEUBERGER

Design: Caroline de Lint, Voorburg (caro@delint.nl)

Publication of this thesis was financially supported by the foundation Centre for Human Drug Research, Leiden, the Netherlands.

THE CLINICAL PHARMACOLOGY OF PERFORMANCE ENHANCEMENT AND DOPING DETECTION IN SPORTS

PROEFSCHRIFT

ter verkrijging van de graad van Doctor
aan de universiteit Leiden, op gezag van
Rector Magnificus prof. mr. C.J.J.M. Stolker,
volgens besluit van het College voor Promoties
te verdedigen op donderdag 16 mei 2019
klokke 11:15 uur

DOOR

Jules Heuberger
geboren te Maastricht in 1989

PROMOTOR
Prof. Dr. A.F. Cohen

CO-PROMOTORES
Dr. G.J. Groeneveld
Dr. J.I. Rotmans

LEDEN PROMOTIECOMMISSIE
Prof. Dr. J.M.A. van Gerven
Prof. Dr. C.M. Cobbaert
Dr. J.M.A. Daniels (*Department of Pulmonary Diseases, VU University Medical Centre, Amsterdam*)
Prof. Dr. E.C.M. de Lange (*Research Division of Systems Biomedicine and Pharmacology, LACDR, Leiden*)
Prof. Dr. P.C.W. Hogendoorn

1
INTRODUCTION

Since the discovery of pharmacologically active plants and other products, these have been applied to treat disease.[1] But apart from therapeutic use, man has applied such substances for other purposes as well. Opium and coca for example were used in ancient times by priests in religious ceremonies for their euphoric effects, mushrooms and cacti were taken for spiritual reasons, and substances like alcohol, caffeine and nicotine have long been used recreationally (see for example[2–4]). One other distinct deployment of pharmacologically active agents has been in sports, a practice that also dates back to ancient times. In the old Greek Olympic Games, alcohol, mushrooms and sesame seeds were consumed to enhance performance, and stimulants were used by gladiators in Roman times against fatigue and injury.[5] Since those first applications of pharmacological products, our understanding of the involved mechanisms and the related science has evolved tremendously, and with this increase in pharmacological knowledge the possibility and need to determine whether substances are actually beneficial grew. This led to the conception of evidence-based medicine, of which the fundamental aim is to prevent patients from being treated with ineffective agents or getting exposed to unnecessary harm. Pioneers such as Archie Cochrane and Iain Chalmers postulated the importance of thoroughly evaluating the available evidence before implementing a certain treatment in medicine.[6] Such ideas led to the formation of systems to rate the quality of evidence for a treatment, such as the Grading of Recommendations Assessment, Development and Evaluation (GRADE) system[7] and a similar system from the Oxford Centre for Evidence-Based Medicine (OCEBM), partly depicted in Table 1.[8]

According to this latter method, systematic reviews of randomised trials and randomised trials are considered the two strongest levels of evidence for the benefit of a treatment (level 1 and 2 evidence). These carry more weight than for example non-randomised controlled cohort studies, case-control studies, or mechanism-based reasoning. This shows the study design is very important in classifying the strength of the evidence, but an additional contributing factor is the knowledge about the link between the measured pharmacodynamic marker(s) in a study and the actual outcome that is being studied.[9] More specifically, evaluating the effect of a drug on a surrogate marker, for example levels of Tau protein for effects on Alzheimer's Disease, only gives a high level of evidence of a beneficial effect if there is a clear and proven link between the surrogate marker and the intended endpoint. In turn, to understand this link a thorough knowledge about the link between the biological mechanism of the drug and the clinical effect is needed. These two factors are schematically plotted in Figure 1, depicting that the basis for obtaining evidence on the beneficial effect of a drug is to increase the knowledge about the relation between mechanism of action, marker and clinical effect.[9]

This overall concept termed evidence-based medicine has been accepted as an essential part of medicine, crucial to ensuring good quality healthcare by making therapeutic decisions based on weighed and evaluated science. These principles are therefore indeed strictly applied for drugs intended for therapeutic use, and products will not be approved for medical use unless there is a high level of evidence for benefit in patients. Moreover, prescribing drugs without such a level of evidence is considered bad medical practice. However, when pharmacological agents are used in different contexts, these same standards are not applied as strictly. Both in recreational drug use and drug use in sports, there seems to be a discrepancy between the extent of use and the available evidence on the intended and harmful effects. In this thesis we focus on the use of pharmacological agents in sports (termed doping) and apply standards of evidence-based medicine and clinical pharmacology to doping research. That pharmacological agents are being used in sports is clear from the (many) confessions by athletes and the World Anti-Doping Agency (WADA) testing figures,[10] despite regulation aimed to ban or at least restrict doping use.[11] And so if these substances are being used in sports, it is important to acquire knowledge about the related effects and harms. The biological, methodological and pharmacological fundaments are the same as in therapeutics, so why should these basic principles not apply in doping research?

We believe they should, the problem is that these principles are not yet embedded in doping research, as can be seen from some examples of reactions to our research:

- A referee (vigorously) defending the WADA Prohibited List in response to our review on the evidence for effects of testosterone: '*Various testimonies during [legal] trials indicate that athletes use testosterone to recover faster. This is particularly valid for multiday events. The number of volunteers in many intervention studies is often too limited to demonstrate a marginal but still existing effect on performance.*

Most studies are underpowered when dealing with elite athletes due to the difficulty to access this population for obvious ethical and availability reasons. You cannot administer doping substance to elite and competing athletes, this would seriously contravene the rules in force.'

- Dr. Olivier Rabin, WADA's science director, in The Times in response to our model-based publication describing the fundamental flaws in the salbutamol anti-doping procedures and the applied threshold level for triggering a violation:[12] '*... the organisation would vigorously defend its position and had already read the Leiden paper. "I read the article you refer to, and no, no concern at all. Nothing new as their model is based on three well-known studies," Rabin said. He added:* "WADA *has conducted several studies on salbutamol and continues to conduct studies on beta-2 agonists. We believe the current threshold is solid considering the scientific literature published on salbutamol over the past 20 years. Based upon the published and unpublished information in our possession, we see no reason to change the salbutamol threshold.*'
- In a letter to the editor regarding our publication of the effects of erythropoietin in cyclists:[13] '*In summary, the interpretation of their findings with* VO_{2peak} *testing and real-life competition is not convincing [...] In subjects with a similar training background and tested under carefully controlled laboratory conditions with a double-blind study design, erythropoietin clearly improves both* VO_{2peak} *and performance [in a time-to-exhaustion test] (7).*'

These examples indicate that several of the principles are being violated by experts in the field: classification of levels of evidence are not consistently applied, evidence that is available is not always recognised, and clinical relevancy is not determined based on validated (surrogate) markers. These issues should be overcome in order to provide a rational and evidence-based framework for doping research, and we feel that for a major part this is feasible. There are evident differences between therapeutics and doping (e.g. in terms of populations and existing regulation), and some of the described principles or methods might be more difficult or need more attention in the doping setting. But the fact that studies are underpowered to see small effects does not make evidence from n=1 trials (individual experience) stronger, nor can we accept an observed effect on surrogate markers to equal an effect on clinical outcome when this link is not sufficiently proven. In this thesis we therefore describe the initial setup for an updated doping research framework, and offer suggestions to reduce the gap that exists between this discipline and therapeutics. In the first chapters the foundation under doping is evaluated, where we investigate which evidence is available for pharmacological performance enhancement in sports and we describe a study investigating such effects for recombinant human erythropoietin in an evidence-based manner. A separate chapter describes that in the same study, we investigated the harmful effects of this treatment in athletes, showing that such studies could form the evidence-base for the benefit-risk assessment of a doping agent. Furthermore, we document the analysis of a commonly used surrogate marker and its relation to the clinical outcome of endurance performance to provide knowledge about the link between this marker and clinical effect. Finally, we discuss the detection of substances, a problem that is perhaps even more relevant to doping research than to therapeutics. Reliable and validated assays and related procedures here are paramount, and in the last two chapters we provide suggestions for improvement of several currently applied procedures and add to the available knowledge. In all chapters, state-of-the-art clinical pharmacological methodology is applied to address a particular doping problem, such as for example population pharmacokinetic modelling in the chapter on salbutamol detection. The combination of chapters shows that there is an overall lack of evidence in doping research, but also shows the feasibility of applying accepted standards from therapeutics to provide this necessary evidence. With that, this thesis aims to contribute to further evolution of this research field, just as the field of therapeutics has recently evolved towards evidence-based medicine.

REFERENCES

1 Heymans C. Pharmacology in old and modern medicine. *Annu Rev Pharmacol*. 1967;7:1-14.
2 Brownstein MJ. A brief history of opiates, opioid peptides, and opioid receptors. *Proc Natl Acad Sci U S A*. 1993;90(12):5391-5393.
3 Stolberg VB. The use of coca: prehistory, history, and ethnography. *J Ethn Subst Abuse*. 2011;10(2):126-146.
4 Crocq MA. Historical and cultural aspects of man's relationship with addictive drugs. *Dialogues Clin Neurosci*. 2007;9(4):355-361.
5 Yesalis CE. History of doping in sport. In: Bahrke MS, Yesalis CE, eds. *Performance enhancing substances in sport and exercise*. Champaign: Human Kinetics; 2002:1-20.
6 Chalmers I, Dickersin K, Chalmers TC. Getting to grips with Archie Cochrane's agenda. *BMJ*. 1992;305(6857):786-788.
7 Guyatt GH, Oxman AD, Vist GE, et al. GRADE: an emerging consensus on rating quality of evidence and strength of recommendations. *BMJ*. 2008;336(7650):924-926.
8 OCEBM_Levels_of_Evidence_Working_Group*. "The Oxford Levels of Evidence 2". Oxford Centre for Evidence-Based Medicine; https://www.cebm.net/index.aspx?o=5653. Accessed 29-10-2018).
9 Cohen AF. Developing drug prototypes: pharmacology replaces safety and tolerability? *Nat Rev Drug Discov*. 2010;9(11):856-865.
10 World_Anti-Doping_Agency. Anti-doping testing figures. 2012-2017; https://www.wada-ama.org/en/resources/laboratories/anti-doping-testing-figures-report. Accessed August 8, 2018.
11 World_Anti-Doping_Agency. World Anti-Doping Agency Code. 2015; https://www.wada-ama.org/en/resources/the-code/world-anti-doping-code. Accessed August 8, 2018.
12 Dickinson M. Study raises doubts about Chris Froome's salbutamol test. *The Times*. May 7, 2018;Cycling.
13 Sperlich B, Wallmann-Sperlich B, Zinner C, Von Stauffenberg V, Losert H, Holmberg HC. Functional High-Intensity Circuit Training Improves Body Composition, Peak Oxygen Uptake, Strength, and Alters Certain Dimensions of Quality of Life in Overweight Women. *Front Physiol*. 2017;8:172.

FIGURE 1 DIAGRAM OF TWO IMPORTANT FACTORS IN DRUG DEVELOPMENT Proving a clinical effect is done most rationally and efficiently by moving drugs to the upper right corner: obtaining a strong link between the molecular mechanism and the clinical effect, as well as measuring markers that relate to the clinical effect. With permission from Adam F. Cohen in Nature Reviews[9]

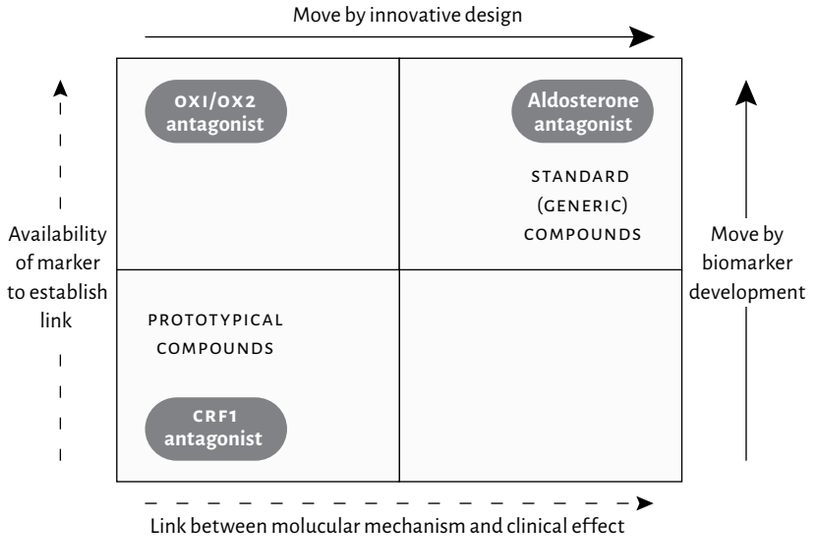


TABLE 1 OXFORD CENTRE FOR EVIDENCE-BASED MEDICINE 2011 LEVELS OF EVIDENCE

Question	Step 1 (Level 1*)	Step 2 (Level 2*)	Step 3 (Level 3*)	Step 4 (Level 4*)	Step 5 (Level 5)
Does this intervention help? (Treatment Benefits)	Systematic review of randomized trials or n-of-1 trials	Randomized trial or observational study with dramatic effect	Non-randomized controlled cohort / follow-up study**	Case-series, case-control studies, or historically controlled studies**	Mechanism-based reasoning

** Level may be graded down on the basis of study quality, imprecision, indirectness (study PICO does not match questions PICO), because of inconsistency between studies, or because the absolute effect size is very small; Level may be graded up if there is a large or very large effect size.*
*** As always, a systematic review is generally better than an individual study.*

2
REVIEW OF WADA PROHIBITED SUBSTANCES: LIMITED EVIDENCE FOR PERFORMANCE-ENHANCING EFFECTS

SPORTS MEDICINE 2018

Jules AAC Heuberger
Adam F Cohen

ABSTRACT

The World Anti-Doping Agency is responsible for maintaining a Prohibited List that describes the use of substances and methods that are prohibited for athletes. The list currently contains 23 substance classes, and an important reason for the existence of this list is to prevent unfair competition due to pharmacologically enhanced performance. The aim of this review was to give an overview of the available evidence for performance enhancement of these substance classes. We searched the scientific literature through PubMed for studies and reviews evaluating the effects of substance classes on performance. Findings from double-blind, randomised controlled trials were considered as evidence for (the absence of) effects if they were performed in trained subjects measuring relevant performance outcomes. Only 5 of 23 substance classes show evidence of having the ability to enhance actual sports performance, i.e. anabolic agents, beta-2-agonists, stimulants, glucocorticoids and beta-blockers. One additional class, growth hormone, has similar evidence but only in untrained subjects. The observed effects all relate to strength or sprint performance (and accuracy for beta-blockers); there are no studies showing positive effects on reliable markers of endurance performance. For 11 classes, no well-designed studies are available, and, for the remaining six classes, there is evidence of an absence of a positive effect. In conclusion, for the majority of substance classes, no convincing evidence for performance enhancement is available, while, for the remaining classes, the evidence is based on a total of only 266 subjects from 11 studies.

INTRODUCTION

The mission of the World Anti-Doping Agency (WADA) is to lead a collaborative worldwide movement for doping-free sport, and its activities focus on the responsibilities given by the World Anti-Doping Code.[1] One of these responsibilities is to publish an annual Prohibited List, which identifies the substances and methods prohibited in- and out-of-competition, and in particular sports.[2] This list is compiled by the List Expert Group and Health, Medical and Research Committee of the WADA, in consultation with scientific, medical and anti-doping experts, using criteria described in the World Anti-Doping Code. This describes that a substance or method shall be considered to be placed on the Prohibited List if the substance or method meets any two of the following three criteria:[1]

1. Medical or other scientific evidence, pharmacological effect or experience that the substance or method, alone or in combination with other substances or methods, has the potential to enhance, or enhances, sport performance
2. Medical or other scientific evidence, pharmacological effect or experience that the use of the substance or method represents an actual or potential health risk to the athlete
3. WADA's determination that the use of the substance or method violates the spirit of sport, described in the introduction to the World Anti-Doping Code.

The third criterion is clearly most subjective and is more a fundamental and philosophical question than a scientific one.[3] However, the remaining two criteria do mention the availability of scientific evidence, indicating that the decision for placing substances and methods on the Prohibited List could be evidence-based. So how strong is this evidence for the listed substances? In this review, we specifically focus on the evidence for performance enhancement, although there could be other reasons athletes use prohibited substances, including masking or diminishing the side effects of other prohibited substances. Several reviews are available focusing on performance effects of different categories on the Prohibited List[4–7]; however, the current review aims to provide a comprehensive and up-to-date overview of the evidence for performance enhancement of all categories of substances in- and out-of-competition on the 2018 Prohibited List, applying standards considered appropriate in clinical therapeutics.

METHODOLOGICAL CONSIDERATIONS

The 2018 Prohibited List was used as framework for this review.[2] We searched the scientific literature for studies and reviews evaluating the clinical effects of the different substances and categories of substances on performance using PubMed as the search engine. Scientific articles with no date restriction and with combinations of the following keywords were evaluated for their relevance by title and abstract: 'athletes', 'performance', 'sport', 'doping' and 'trained', in combination with a specific prohibited compound or category (e.g. 'terbutaline' or 'beta-2 agonist'). Reference lists of identified publications were searched for additional relevant publications. Performance was interpreted according to the broadest sports-related definition, including strength (power) and endurance. Although the criterion in The WADA Code states that evidence for *the potential* to enhance performance is sufficient to place a substance on the Prohibited List, in this review clinical pharmacological evidence for actual performance enhancement was considered essential to determine that a substance or category of substances has a positive effect on performance. In other words, similar to any other therapeutic review, to make an evidence-based conclusion that there are performance enhancing effects the level of evidence should preferably be high (level 1), meaning that evidence should come from double-blind randomised controlled trials (or meta-analysis based on randomised controlled trials).[8] This was also taken into account when evaluating the search results, although there will inevitably be cases where information has to be inferred from other, less reliable evidence.

In addition, ideally these trials should measure relevant performance outcomes and so we defined which outcomes should be considered most relevant. Here we apply the same standard as for clinical trials, where proven effects on clinical outcome are accepted as most reliable, and effects on surrogate markers that have a proven link to that clinical outcome are accepted, as for example described by the US Food and Drug Administration.[9] When translated to sport performance, the most relevant outcome measure is the 'actual' performance of the sport itself, such as for example muscle strength for weight lifting or running time for distance running. However, surrogate markers that describe an important aspect of the performance might be acceptable, but conclusions based on such markers can only be reliable if there is a proven high correlation with the actual performance. For endurance performance for example, although maximal oxygen consumption (VO_{2max}) is often used and has been shown to be a prerequisite for performance,[10,11] its predictive value for endurance performance within a group of athletes is very limited.[12,13] Moreover, it seems that successful endurance athletes reach a plateau in VO_{2max} despite continuing to improve performance,[14–16] thereby questioning whether increasing VO_{2max} by any means would have an impact on performance, at least in highly trained subjects. And finally, there has been critique on the use of the maximal exercise test that generates the VO_{2max} marker to accurately evaluate athletic potential in general, as it does not resemble normal exercise.[17] It is therefore unclear whether a pharmacological effect on VO_{2max} (or other maximal exercise test markers) translates into an effect on performance per se, making it a marker with insufficient predictive value. Another test that is not very reliable in measuring effects on actual performance is the time to exhaustion test. Such a test has been shown to have low reproducibility, especially compared to time trials that continue for a predetermined amount of time or work.[18,19] Moreover, there is no clear evidence of their correlation with actual performance, except for absence of a correlation with Ironman performance in one study.[20] Possibly this is because sports disciplines do not rely on time to exhaustion principles, but rather on pacing to a finish line or time. In summary, there currently are no widely recognised laboratory markers for (aerobic) endurance performance, leaving tests for actual endurance performance (e.g. a time trial) as the most reliable available measure. Markers for sprint performance on the other hand, for example as measured by a Wingate test, do resemble actual performance such as sprinting in cycling and this surrogate marker has been shown to correlate with other performance types as well,[21,22] which is why we considered it to be a relevant marker.

Finally, the training status of the study participants is a relevant factor when interpreting the outcome. The aim of preventing performance advantages through doping, as described in the WADA Code, is most (although admittedly not solely) relevant in high level (and in particular professional) sports, due to the attention, fame and commercial considerations involved in that level of sports. Clinical studies should reflect the 'target population', which in this case would be elite and professional athletes. However, because of doping/WADA regulations, it was/is very challenging or even impossible to conduct intervention studies of banned substances in such a population. For this reason, we considered studies in (highly)

trained athletes most relevant, so that observed effects apply to this level of athletes, and that extrapolation of observed effects in this population to the performance of professional athletes was most valid. However, data in less well trained subjects may also be of value, and was also reviewed. Determining the training level of subjects was based on commonly used markers for performance where possible. For the level of training in endurance performance, VO_{2max} and maximal power output (P_{max}) were used. Three categories were defined somewhat arbitrarily (without taking the type of maximal exercise testing protocol into consideration): untrained (VO_{2max} < 55 mL/min/kg and / or P_{max} <3.5 W/kg); trained (VO_{2max} ≥ 55 and < 65 mL/min/kg and / or P_{max} ≥3.5 and <5.0 W/kg); and highly trained (VO_{2max} ≥ 65 mL/min/kg and / or P_{max} ≥5.0 W/kg). For strength training it was more difficult to objectively categorize study populations as available measurements varied widely between included studies. Therefore, subjects were categorised as trained or untrained based on the description in the article of whether subjects had been actively engaged in resistance training.

FINDINGS

Prohibited at all times

S0 Non-approved substances

Any pharmacological substance that has no current approval by any governmental regulatory health authority for human therapeutic use belongs in this category, making the category very broad. Substances in this category could be drugs under pre-clinical or clinical development, discontinued drugs, designer drugs, or substances approved for veterinary use only. In any case, they will be substances that (currently) lack solid evidence for (beneficial) effects in humans in general, and therefore, in practically all cases, lack evidence for enhancement of performance in particular.

S1 Anabolic agents

Anabolic agents, or anabolic-androgenic steroids (AAS), are synthetic derivatives of testosterone which have attracted attention as doping substances due to their potential to increase protein synthesis and decrease protein breakdown (anabolic effects) and increase muscle growth (androgenic effects) by activating the androgen receptor. A very thorough review evaluated the evidence for effects of AAS on performance in 2004.[23] Upon inspection of the studies covered in the review, there are various studies with a randomised, double-blind, controlled design that investigated effects on strength. The most recent of these studies show clear effects of AAS on different strength outcomes in strength trained men, alone and combined with strength training.[24–26] One of these studies showed with an elegant design that high dose testosterone (600 mg/week) both with and without strength training significantly increased bench-press and squatting power with on average 10-20% compared to the respective placebo condition.[24] Well-designed studies investigating effects on endurance performance, or related measures such as VO_{2max}, covered in the review are older and more sparse. These show no treatment-induced improvements, although they did not show an effect on strength either, indicating the sample size or dose might be too small to detect effects.[27,28] However, since the review one additional randomised, placebo controlled trial has become available which also showed a lack of effect on endurance performance markers of a month of AAS in doses similar to those that showed strength effects.[29] Additionally, this study showed evidence that there is no effect of AAS treatment on recovery. In summary, high dose AAS appears to increase strength, but not endurance performance. The evidence on strength effects is based upon 3 studies with in total 91 volunteers.

S2 Peptide hormones, growth factors, related substances and mimetics

ERYTHROPOIETINS AND AGENTS AFFECTING ERYTHROPOIESIS

These agents are aimed at increasing red blood cell volume through inducing erythropoiesis and thereby potentially enhancing performance. Interestingly, for "natural" increases in red blood cell volume through altitude training, the evidence for performance enhancing effects is not fully convincing to begin with.[30]

Erythropoietin-Receptor Agonists

Erythropoietin-receptor agonists, such as recombinant human erythropoietins (rHuEPO), stimulate erythropoiesis and thereby increase haemoglobin levels, which potentially increases oxygen carrying capacity and thereby improves endurance performance. A systematic review of the literature by Heuberger *et al* concluded

however that there was a lack of evidence for efficacy on endurance performance.[31] Of the then 13 available reviewed studies, only five had a placebo-controlled and double-blind design,[32–36] all showing similar effects of rHuEPO in both trained and untrained subjects: in all studies VO_{2max} increased by approximately 7%, while maximal power output was evaluated in two of the studies and increased by 7% as well.[33,36] Finally, time to exhaustion improved by 22% in untrained[33] and 9.4% in trained subjects.[32] Two subsequent randomised, placebo-controlled trials also showed increases in VO_{2max}, maximal power output and time to exhaustion of 5%, 6% and 58%, respectively, in trained subjects,[37] and an increase in VO_{2max} of 6%, but no increase in time to exhaustion, in untrained subjects.[38] None of these studies showed however, whether these effects on surrogate biomarkers impacted actual performance. Because of this lack of information, a double-blind, randomised, placebo-controlled study in trained cyclists followed and showed that clinically more relevant tests such as a time trial and uphill road race were not affected by rHuEPO treatment.[39] Although there was again an effect of rHuEPO on maximal exercise test variables including VO_{2max} and maximal power output (increase of 5 and 3% respectively), there is no evidence that these erythropoietin-induced effects improve actual cycling performance in trained cyclists. The absence of an effect on these measures most related to competitive (cycling) performance in athletes is insightful, but one should be cautious about extrapolating these findings to all performance types in elite athletes; not all performance aspects of endurance have been studied, and the target population has not been included. In any case, there is no evidence available that rHuEPO enhances time trial, climbing or other race performance in athletes.

Hypoxia-inducible factor activating agents

Hypoxia-inducible factor (HIF) activating agents have a direct effect on erythropoietin production by stimulating erythropoietin gene expression and thereby the same rationale for potential performance enhancing effects applies as for direct rHuEPO administration. As shown for rHuEPO in the section above, increases in erythropoietin and the accompanying increases in haemoglobin have not been shown to improve endurance performance in trained subjects. Moreover, evidence for the effects of HIF activating agents is even more sparse. Cobalt has been observed to increase erythropoiesis in anaemic patients.[40,41] No trials have been performed evaluating these effects on erythropoiesis or performance in healthy volunteers, let alone athletes, as can be seen in the review by Ebert and Jelkmann.[42] More recently, a study claimed to show effects of xenon on erythropoietin production in healthy volunteers,[43] but the statistics of the study have been criticised.[44] Small molecule HIFs are in clinical development, but have not yet been approved for clinical use. The published clinical studies show that these compounds produce modest increases in erythropoietin in both anaemic patients and healthy volunteers,[45,46] however, currently there are no studies evaluating effects on performance of healthy or trained subjects.

GATA inhibitors

By inhibiting GATA, an erythropoietin gene expression inhibitor, a similar effect as for the HIF activating agents could be expected. There are, however, no published clinical studies of the effects of these compounds, the mechanism has only been proven pre-clinically.[47,48]

Transforming growth factor beta inhibitors

Erythropoietin induction by transforming growth factor beta (TGF)-beta inhibition is a very recent development in the possible treatment of anaemia, and in particular for myelodysplastic syndromes. Luspatercept and sotatercept have been shown to increase haemoglobin levels in such patients,[49,50] but there is no evidence of any related effects on performance in healthy or trained individuals.

Innate repair receptor agonists

Innate repair receptor agonists are non-erythropoietic derivatives of rHuEPO that have been developed for their potential tissue-protective properties and to date have been evaluated in only a few clinical trials. One published placebo controlled trial indicated carbamylated erythropoietin was safe and well-tolerated,[51] but no evidence of performance effects is available.

PEPTIDE HORMONES AND HORMONE MODULATORS

Chorionic Gonadotrophin and Luteinizing Hormone and their releasing factors

Chorionic gonadotrophin (CG) and luteinizing hormone (LH) are hormones that bind to the same receptor (LHCG-receptor), which has several functions in the

reproductive system. In females, follicular maturation, ovulation and luteal function are influenced through stimulation of the receptor in the ovary, in males the receptor is located in the testis and stimulates testosterone production. There is no indication that the effects in females can positively influence performance,[52] but the increase in testosterone in males may give similar effects as described for the anabolic agents: a single intramuscular injection of 6000 IU of CG for example, increased testosterone levels approximately by 40 nmol/L in healthy men.[53] This is half the increase observed after a 10-week treatment with 600 mg testosterone enanthate (an anabolic steroid), which has been shown to increase bench press and squat muscle strength.[24] There are no studies however that have investigated the effects of CG or LH on any sports performance measures.

Corticotrophins and their releasing factors

Adrenocorticotropic hormone (ACTH) is involved in the hypothalamic-pituitary-adrenal axis and is released in response to stress, leading to increases in cortisol. Through this cortisol response, free fatty acids are released, potentially sparing glycogen, which is then assumed to benefit endurance performance. In addition, ACTH stimulates glucocorticoid secretion (see section S9 Glucocorticoids). A double-blind, placebo controlled cross-over study in 16 trained cyclists showed however, that although a 1 mg ACTH depot dose decreased the feeling of fatigue during a submaximal effort, it did not improve maximal performance in a maximal exercise test, nor did it affect recovery between two consecutive tests.[54] Similarly, 20 km time trial performance was not affected by 0.25 mg ACTH intramuscular injections in a double-blind, placebo-controlled cross-over study in 8 (highly) trained male cyclists.[55] Perceived fatigue was not decreased by ACTH in this study. As these are the only studies performed, we conclude there is no evidence of beneficial effects of ACTH or its releasing factors on actual performance.

Growth Hormone, its fragments and releasing factors

Growth hormone (GH) use in adults with GH deficiency results in reduced body fat, increased lean body mass and increased fitness and strength[56] and has therefore attracted attention as a potential performance enhancing drug. This mechanism is mainly mediated by insulin-like growth factor-1 (IGF-1). A systematic review evaluated effects on strength or endurance performance.[57] For strength, two double-blind studies were identified that showed no effects of GH on muscle strength of different muscles compared to placebo when combined with strength training in untrained[58] and trained strength athletes.[59] Endurance performance was evaluated in two double-blind studies: multiple dosing of GH did not have an effect on VO_{2max} and maximal power output compared to placebo in trained subjects.[60] A single dose of GH increased plasma lactate levels during submaximal cycling exercise compared to placebo in seven highly trained cyclists in a cross-over design.[61] Such single administrations of GH therefore rather seem to decrease endurance performance, underlined by the fact that 3 out of 7 cyclists in this study had difficulties completing the cycling trial when treated with GH, compared to none on placebo treatment. Following the review, one randomised, placebo-controlled, blinded trial with 8 weeks of daily GH treatment confirmed these findings and showed no effects on strength or VO_{2max}. In this study, there was however an increase in sprint performance in a 30 second maximal sprint test (Wingate test) of approximately 1 kJ (or a 3.9% relative increase in the combined male and female group and a 5.5% relative increase for the male group only), which was slightly larger when GH was co-administered with weekly testosterone doses.[62] It should be noted that GH also increased the incidence of swelling, joint pain and paraesthesia in this study, indicating these gains are not without downsides and possible risks. Additionally, subjects were untrained for endurance, therefore it is difficult to know how this effect on sprint performance extrapolates to elite athletes.

GROWTH FACTORS AND GROWTH FACTOR MODULATORS

Blood platelets can release growth factors for example when triggered by signs of injury. These could potentially be used for treating sports injuries,[63] but are thought to also give benefit in healthy athletes. For most of these factors however, including fibroblast, hepatocyte, mechano, platelet-derived and vascular-endothelial growth factors and thymosin-beta 4, there are no studies of the effects on performance. The only studies available have evaluated the safety or efficacy of these products in healthy volunteers and patients.[64–66] There is one exception that has been investigated as an ergogenic aid, which is insulin-like growth factor-1 (IGF-1). IGF-1 is thought to possess ergogenic effects mainly through the anabolic pathway that is shared with GH. A randomised, double-blind, placebo-controlled study investigated the effects of a recombinant human insulin growth factor (IGF)-I/IGF binding

protein 3 complex (rhIGF-I/rhIGFBP-3) in untrained persons on body composition and aerobic performance.[67] No effects on body composition were observed, but an increase in VO_{2max} was reported for both a low (30 mg/d) and high (60 mg/d) dose. The conclusion that IGF-1 therefore improves aerobic fitness should be interpreted with care however: firstly, changes in outcome parameters were only analysed within each group, and not compared to the placebo group, which would have been the appropriate analysis in such a study design. Secondly, even if the observed effect on VO_{2max} is truly caused by IGF-1, it is unclear if this has an impact on actual performance. Unfortunately, no performance parameter such as running speed on the treadmill test was reported, nor was a test performed mimicking actual sports performance. This, in addition to the fact that participants were untrained, makes it impossible to interpret what these findings mean for performance of elite athletes.

S3 Beta-2 agonists

Beta-2 agonists are used in the treatment of asthma as they act as bronchodilators through their relaxing effect on the smooth muscles of the lung via the beta-2 adrenergic receptor. In addition, they have an effect on muscle tissue through this pathway, and both actions have been implied to possess performance enhancing effects. Several extensive reviews have evaluated the evidence for this. Pluim *et al*[68] concluded in 2011 in a systematic review based on a meta-analysis of randomised controlled trials that there are no positive effects of inhaled beta-2 agonists on endurance, strength or sprint performance, and that there was insufficient evidence to draw conclusion about systemic beta-2 agonist use. In 2015, Cairns *et al*[69] had more systemic dosing studies at their disposal, and concluded in their review that only high-dose systemic beta-2 agonists (at a serum concentration of ~0.1 µmol/L) have a positive effect on muscle strength and peak sprint power. This is based on the observation that after oral administration of 20-25 mg terbutaline, sarcoplasmic reticulum rates of Ca^{2+} release and uptake were increased, together with maximal voluntary isometric contraction (+6%) and peak twitch force (+11%), in a placebo-controlled randomised crossover design in highly trained and trained men.[70] No effects on time to exhaustion were observed. High dose (15mg) inhaled terbutaline reached similar serum concentrations (~0.1 µmol/L) in another double-blinded randomised crossover trial and increased quadriceps muscle strength by 8.4%. In addition, Wingate peak and mean power were increased by 2.2% and 3.3% respectively and Wingate total work by 3% compared to placebo in trained males, but time trial performance was not affected.[71] A double-blind, randomised, and placebo-controlled study in highly trained athletes published after the review by Cairns *et al*[69] showed that single and 2-week dosing of 8 mg salbutamol had no effect on body mass, VO_{2max}, incremental peak power output, time to exhaustion, maximal voluntary isometric contraction or isometric endurance. There was however a significant increase in Wingate peak power of 4% and 6% for single and multiple dosing, respectively, similar to the inhaled terbutaline study.[72] The statistical analysis in this study did not include a comparison to the placebo treatment, but because there was no significant effect in the placebo group observed, the increase in the salbutamol group seemed to be a true effect. In all three studies, subjects experienced mild side effects, namely tremor and tachycardia. Only one additional study showed effects of inhaled administered beta-2 agonists. In this case the effect was only seen on one very specific task of which the clinical relevance is questionable, namely quadriceps endurance in highly trained endurance athletes,[73] and so the vast majority of evidence shows no ergogenic effects of inhaled beta-2 agonists. Overall, these findings indicate that only high beta-2 agonist concentrations, which are mainly achieved by systemic administration, can improve performance, but only in strength and very short disciplines requiring high power development, as represented by the Wingate test, and at the cost of tremor and tachycardia. This evidence is based upon 3 studies with a total of 39 volunteers.

S4 Hormone and metabolic modulators

AROMATASE INHIBITORS

Aromatase inhibitors lead to reduced enzyme activity for the conversion of androgens to oestrogens. This in turn leads to lowered oestrogen levels, and thereby via inhibition of negative feedback on the hypothalamus to higher testosterone levels. This increase has been shown to be approximately 15 nmol/L in healthy males for exemestane.[74] As for CG and LH, there are no trials investigating the effects of these aromatase inhibitors on performance, and the only indication of potential effects is an increase in testosterone, which is roughly 25% that observed after AAS treatment leading to increased muscle strength.[24] Evidence is therefore similarly weak as for CG and LH.

SELECTIVE ESTROGEN RECEPTOR MODULATORS

The evidence basis for selective estrogen receptor modulators (SERMs) is similar to that for aromatase inhibitors. SERMs, such as tamoxifen and raloxifen, are clinically used for their estrogenic and antiestrogenic effects in different tissues. This induces increases in pituitary gonadotrophin secretion and consequently increases in testosterone levels in men, seemingly somewhat smaller than for aromatase inhibitors.[75] There are no studies investigating the effects of SERMs on performance.

OTHER ANTI-ESTROGENIC SUBSTANCES

The examples mentioned in the 2018 Prohibited List in this category, clomiphene and cyclofenil, are older SERMs (although perhaps less selective than, for example, tamoxifen). As effects are similar to compounds described in the previous section and there are no studies into performance enhancement,[76,77] the conclusion about the evidence for performance effects is the same: there is no evidence available. Another substance in this category, fulvestrant, is a selective estrogen receptor degrader with no effects that could clearly enhance performance and no evidence that it does so.

AGENTS MODIFYING MYOSTATIN FUNCTION(S)

Myostatin is a negative regulator of muscle growth and therefore lowering its levels or inhibition of its action could potentially increase muscle size and improve performance. Although muscle growth is observed in some pre-clinical studies, it is questionable if this also results in increased strength, as reviewed by Fedoruk and Rupert.[78] In addition, there are currently no approved drugs (developed for diseases with muscle weakness or wasting) in this class yet,[79] and so there is currently no evidence of effects on performance in athletes.

METABOLIC MODULATORS

There are several substance types in the metabolic modulators category. Peroxisome proliferator-activated receptor-delta (PPAR-delta) agonists and AMP-activated protein kinase (AMPK) activators might enhance performance through their effects on energy expenditure and substrate utilization. In mice, a PPAR-delta agonist as well as an AMPK agonist (i.e. 5- aminoimidazole -4-carboxamide ribonucleotide (AICAR)) increased running endurance.[80] There are however currently no approved PPAR-delta agonists,[81] neither is there evidence for performance enhancement in humans. Similarly, specific AMPK activators are not approved (such as AICAR), although there are approved drugs that have an AMPK activating effect, e.g. metformin. Clinical studies evaluating effects on performance in healthy subjects are however sparse, as reviewed by Niederberger *et al*.[82] This review cites two studies evaluating metformin effects in healthy volunteers; one is a multiple dose double-blind, placebo controlled cross-over trial in healthy subjects.[83] The blinding of this study was described as not being optimal (due to taste and gastrointestinal side effects), randomisation is not described and there was no baseline measurement for each treatment, making the conclusions less robust. Nevertheless, there was no positive effect observed on performance markers. Moreover, a small but significant decrease in VO_{2max} and time to exhaustion in the maximal test was found in the metformin treatment group. The second study is a randomised, double-blind, placebo-controlled single dose cross-over study which showed no difference between treatments, although this study also did not include a baseline measurement.[84] In both studies, participants were untrained.

With regards to insulin Kuipers and van Dugteren indicated that based on several observations, this drug is not expected to have a physiologically significant effect on muscle growth, even in combination with glucose and/or amino acids.[85] There are however, no studies published assessing the effects of insulin on performance.

Finally, inhibitors of fatty acid oxidation belong to this category. Meldonium is classified as a partial inhibitor of fatty acid oxidation, but in a recent editorial, Greenblatt and Greenblatt concluded that there are no studies available that have evaluated the performance enhancing properties of meldonium in trained subjects.[86] Another inhibitor of free fatty acid oxidation, trimetazidine, was reported to improve maximal walking distance in patients with peripheral arterial disease,[87] but there is no evidence of such an effect on exercise performance in healthy or trained individuals.

S5 Diuretics and masking agents

The category of diuretics and masking agents is not necessarily on the Prohibited List for its potential to enhance performance. Masking agents are supposed to interfere with analytical testing of markers or other substances on the Prohibited

List. Diuretics increase urine production and by this effect are thought to dilute, and therefore interfere with detection of, banned substances in urine. This increased water excretion caused by diuretics might also improve performance, as it can quickly reduce weight which might give a competitive advantage. In sports with weight classes for example, this effect could place athletes in a lighter category, and in speed or endurance sports lighter athletes might have an advantage. Cadwallander *et al*[88] reviewed the effects of diuretics, but it should be noted that some of the studies were not placebo-controlled, and only used a control condition. Although it could be argued that the diuretic effect would have de-blinding effects anyway, the results should be interpreted with caution. Caldwell *et al*[89] showed that two doses of approximately 60 mg furosemide decreased work load during a maximal exercise test, and decreased VO_{2max} compared to baseline measurements but not compared with controls, in untrained subjects. Armstrong *et al*[90] found that trained runners had an impaired running time in 1500, 5000 and 10000 meter races after 40 mg of furosemide, a difference versus control that was significant at the two longest distances. A third study did not find an effect of a 1000 mg infusion of acetazolamide on 30 second peak or average cycling power, although it did seem to decrease peak VO_2 uptake during this test.[91] Another study evaluated the effects of a single dose of 500 mg acetazolamide in a quasi-randomised, double-blind, placebo-controlled cross-over study and found that there was no effect on VO_{2max} but time to exhaustion was reduced by 29% in a continuous exercise to exhaustion.[92] Finally, a double-blind, placebo-controlled cross-over study in untrained subjects investigated the effects of four doses of 250 mg acetazolamide every 8 hours and found a decrease in VO_{2max} and maximal power output.[93] An additional study that was not covered in the review by Cadwallander *et al*[88] showed that in a randomised, double-blind, placebo-controlled, cross-over study 250 mg acetazolamide three times a day for two days did not significantly affect knee extension maximum voluntary contraction at the beginning of the test or at exhaustion in untrained subjects.[94] It did however decrease endurance performance. Overall, not all study designs were sufficiently robust and most included untrained subjects, so definite conclusions cannot be made about the performance enhancing properties of diuretics. But given the available studies, if anything, the evidence indicates that athletic performance is negatively affected by diuretics.

M1-3 Prohibited methods

There are several non-pharmacological interventions that are prohibited at all times, termed prohibited methods. These are manipulation of blood and blood components (e.g. blood transfusion), chemical and physical manipulation (e.g. tampering with a sample or intravenous infusions of fluid) and gene doping. As this review focuses on pharmacological interventions, evidence for effects on performance of these categories is not discussed here.

Prohibited in-competition

S6 Stimulants

Stimulants are thought to potentially improve performance via the effects on neurotransmitter levels in the brain, predominantly dopamine and norepinephrine. Research into effects of stimulants on performance has mainly focused on a few drug classes. Amphetamines such as amphetamine sulphate[95] showed positive effects on muscle strength (knee extension strength +23%), acceleration (+4%) and time to exhaustion (+5%) in untrained subjects. Similarly, methylphenidate[96] improved time to exhaustion (+29%) in highly trained subjects. VO_{2max} was not affected in either study and endurance performance (such as a time trial) was not investigated in these studies. Of note, the former study used no baseline correction (i.e. amphetamine performance was directly compared to placebo performance in the randomised cross-over design) and for the latter study it is unclear whether it was (double-)blinded, which may both make the results less robust. Another study with a higher dose of methylphenidate showed no effect on time trial performance in normal temperature, but there was an improvement of 15% average power output compared to placebo in the heat (30 degrees) in trained subjects.[97] Levomethamphetamine was investigated for its effect on time trial performance in young participants and showed no change.[98]

Ephedrine, pseudoephedrine and phenylpropanolamine have a similar mechanism of action to amphetamines. Two studies investigating the effects of ephedrine showed positive effects. One study found an effect on peak Wingate sprint power (+0.6%) but not on time to exhaustion[99] in untrained subjects and another study found an improvement in a type of time to exhaustion test

in trained strength athletes, namely leg and bench press repetitions (+30% and +8%, respectively).[100] One positive study for pseudoephedrine used a dose of 180 mg which increased knee extension strength by 9% and peak Wingate sprint performance by 3%, but not bench press power in strength trained subjects.[101] Later publications also showed that low doses of pseudoephedrine used clinically did not affect 5000m run time in highly trained runners[102] or peak power or total work during a Wingate test in trained subjects[103]; only high doses improved performance, with 1500m run time decreasing by 2% in highly trained runners.[104] The authors of this latter study therefore concluded that high pseudoephedrine doses are needed for performance effects.

For another well-known stimulant that is on the Prohibited List, cocaine, there are no well-designed studies evaluating effects on performance.

Overall, studies of the effects of these stimulants show varying results, making it unclear whether they improve performance, as was concluded in a review published by Clarkson and Thompson in 1997.[6] In certain conditions and performance tests, they may modestly improve performance if administered in sufficiently high doses, but there is not sufficient conclusive evidence to determine how they affect most actual sports performance types. The available evidence consists of the results of 2 studies involving a total of 29 volunteers.

S7 Narcotics

The narcotics category consists of strong analgesics, all belonging to the opioids. Although surprisingly not all opioids are currently banned (e.g. tramadol is allowed), substances like morphine and its analogues and fentanyl and its derivatives are. Although analgesic effects might enhance performance, common side effects of opioids, including nausea, sedation and respiratory depression, would equally argue against any beneficial effects. One study showed that intrathecal injection of fentanyl did not impact average power output during a 5-km cycling time trial in trained cyclists.[105] However, power output during the first half of the time trial was increased, and then decreased during the second half compared to placebo. The authors attributed this to attenuated afferent feedback from exercising muscles, which is then followed by excessive development of fatigue, and overall deterioration of the ability to "dose" their effort. Besides this report there are no convincing clinical studies of the effects of narcotics on sports performance, leaving evidence for either positive or negative effects on performance lacking, as was also concluded by the authors of a recent review.[106]

S8 Cannabinoids

Cannabinoids are known to affect perceptual function, and Huestis[107] concluded in a review of (non-sport) performance that this leads to decreased ability to concentrate and maintain attention. In addition, this review concluded that cannabinoids impair information processing and reaction time, all of which would probably negatively affect sports performance, as concluded in a more recent review.[108] Around the same time Huestis *et al*[109] argued in a review that although there are indications that in some settings cannabis has a detrimental effect on performance, in other settings known effects of cannabis might be beneficial. Examples include sports where vision or muscle relaxation are important, or when anxiety or fear impair the potential of the athlete. There are, however, very few scientific data available on the effects of cannabinoids on sports performance itself, a conclusion that was also reached in two recent reviews.[110,111] One double-blind, randomised, placebo-controlled cross-over study is available which showed tetrahydrocannabinol (THC) had no effect on hand grip strength and decreased performance in a specific type of submaximal bicycle test compared to placebo in healthy untrained males.[112] This shows there is no evidence for performance enhancement of cannabinoids.

S9 Glucocorticoids

Glucocorticoids act on metabolism and the immune system, and through that mechanism potentially affect performance. For this reason systemic doses are prohibited in competition. A recent review showed there are varying results of glucocorticoid treatment in performance tests.[113] The two available controlled studies evaluating maximal exercise test variables failed to show effects on VO_{2max} and ventilatory threshold of five days of dexamethasone in untrained subjects[114] and on maximal power output of four weeks of budesonide treatment in trained subjects.[115] Effects on short intense exercise were evaluated in three studies. In untrained men, one-legged knee-extensor exercise time to exhaustion was not

affected by five days of dexamethasone.[116] In contrast, using a similar dosing scheme another study did find an increase in one-legged knee-extensor exercise time to exhaustion of 29% and running distance in a certain type of maximal exercise test, namely 20-m shuttle-run test of 19%.[117] Sprint performance over 30 metres was not affected in this study. The authors of the latter study postulated a lack of statistical power in the former study was the cause of this apparent discrepancy in outcomes between the two studies. A third study evaluated the effects of a single prednisone administration on one-legged hopping, and found an 11% improvement in maximal force of the first bout, but not on subsequent bouts or time to exhaustion in any of the bouts.[118] It should be noted however that there was no baseline measurement done on the study day and so effects of inter-occasion variability cannot be excluded. Similar to these short intense exercise studies, results from studies investigating types of cycling performance are equivocal. A single dose of 20 mg prednisolone did not affect cycling time to exhaustion in trained males, alone or in combination with 4 mg salbutamol,[119] a finding that was confirmed in a very similar study.[120] However, a multiple dose of 60 mg prednisolone daily for seven days did increase cycling time to exhaustion in trained males by 28 minutes (62%), although this performance was not controlled with a baseline measurement.[121] An almost identical study that did include a baseline measurement showed an increase of 91% (50.9 minutes) in cycling time to exhaustion with the same dosing regimen, combined with intense training in untrained subjects.[122] Although the statistical comparison was made between baseline measurement and post-treatment, and not additionally to the placebo measurements, it seems likely that this is a true effect as there was no change in the placebo treatment. Another study confirmed these findings in untrained females treated with 50 mg prednisone daily for a week, which showed a 39% increase (18.5 minutes) in cycling time to exhaustion.[123] It should be noted however that it is unclear how time to exhaustion relates to real life endurance performance, which is usually not until exhaustion but until a finish line. In summary, there is conflicting evidence on the effectiveness of glucocorticoids for improving different performance types. However, there seems to be an effect on specific strength tests and shuttle run time, and multiple, but not single doses, seem to improve time to exhaustion in moderately trained subjects. At the same time, only one study with 10 subjects showed an effect on a relevant performance surrogate marker, namely one-legged hopping maximal force.

P1. Prohibited in particular sports

This category covers substances prohibited in particular sports (i.e. archery, automobile, billiards, darts, golf, shooting, skiing/snowboarding and underwater sports) and contains only the group of beta-blockers. This group of substances inhibits beta-adrenergic receptors, thereby reducing heart rate, anxiety and tremulousness, which could potentially enhance performance in sports where precision and accuracy are vital. There is one double-blind, randomised, placebo-controlled cross-over study available evaluating the effect of metoprolol on shooting performance in amateur marksmen.[124] The study showed that on average, participants improved their shooting on metoprolol compared to placebo, which was especially the case in the more skilled marksmen. It seems, therefore, that beta-blockers do improve shooting performance, and possibly other precision and accuracy sports included in this category as well, based on one study of 33 subjects.

CONCLUSION

Of all 23 specific substance classes defined in the 2018 Prohibited List, only five classes show evidence of having the ability to enhance actual sports performance (see Table 1). Anabolic agents can increase muscle strength at supratherapeutic doses, beta-2 agonists can increase muscle strength and peak sprint power at high concentrations, some stimulants increase muscle strength, peak sprint power and decrease in 1500 m run time, glucocorticoids can improve muscle strength and beta-blockers can improve accuracy. In addition, for one more class there is evidence of performance enhancement, but only in untrained subjects: growth hormone can improve sprint performance. Importantly, there is no robust evidence for any of the substance classes on the Prohibited List for the ability to improve endurance performance. Glucocorticoids improve time to exhaustion, but it is unclear whether this relates to actual endurance performance. Erythropoietin-receptor agonists improve VO_{2max} and maximal power output, but the available evidence shows no effect on actual endurance performance. What also becomes clear from this overview is that for most substance classes (11 out of 23) there are no well-designed studies available that evaluate effects on performance (in trained subjects), meaning there is absence of level 1 evidence. Physicians involved

in administering such substances in particular are performing practices similar to off-license prescribing, and prescribing without evidence is considered bad medical practice. In contrast, for another six substance classes, well-designed studies are available which show some evidence of absence of (relevant) effects on performance. Overall, this review therefore shows that for the majority of substance classes (17 out of 23) there is no convincing evidence that they enhance performance of athletes. Moreover, in regard to the other five classes that are prohibited in all sports that do have evidence-based effects, it is unproven whether such effects are relevant or useful in many types of sport (e.g. endurance sports) as they only improve specific performance tasks (mainly strength and sprint power). Although these aspects of course do play some role in many sport disciplines (such as for example athletics or soccer), it is not very clear whether these effects would also impact actual performance in those disciplines. In any case there are no studies investigating this as we have shown. These findings together seem discordant with the general perception that substances on the Prohibited List by definition improve performance (to a great extent). This is especially evident when it is considered that a total of only 266 subjects (from 11 studies) form the clinical pharmacological level 1 evidence base for performance enhancement, the main reason for anti-doping efforts. Although the WADA Code only requires evidence for *the potential* to enhance performance, and there are two other criteria that can be applied to make a substance prohibited, we conclude there is a lack of high level evidence for improvement of actual performance based on this review. Undertaking more high quality clinical research to provide the level 1 evidence base for the current Prohibited List could fill some of these gaps. Some of this research could be impossible due to practical or ethical objections, but the current level of randomised evidence is low and there appear to be many areas where such research is possible. Furthermore if there is clear evidence that there are no performance enhancing effects of a certain class, athletes should be informed of this. This could potentially lead to fewer athletes being tempted to use these substances. Finally, such steps would lead to a more transparent and high level evidence-based fight against doping, and possibly reduce the efforts and resources needed to test for abuse.

REFERENCES

1 The World Anti-doping agency. World anti-doping code 2015. 2015.
2 The World Anti-doping agency. Standard prohibited list 2018. 2018;(January).
3 de Hon O. The redundancy of the concept of 'spirit of sport' in discussions on the prohibited list of doping substances. *Int J Sport Policy Polit*. 2017;9(4):667-676. doi:10.1080/19406940.2017.1348380.
4 Momaya A, Fawal M, Estes R. Performance-enhancing substances in sports: a review of the literature. *Sports Med*. 2015;45(4):517-531. doi:10.1007/s40279-015-0308-9.
5 Bird SR, Goebel C, Burke LM, Greaves RF. Doping in sport and exercise: anabolic, ergogenic, health and clinical issues. *Ann Clin Biochem*. 2016;53(2):196-221. doi:10.1177/0004563215609952.
6 Clarkson PM, Thompson HS. Drugs and sport. Research findings and limitations. *Sports Med*. 1997;24(6):366-384. doi:10.2165/00007256-199724060-00003.
7 Birzniece V. Doping in sport: effects, harm and misconceptions. *Intern Med J*. 2015;45(3):239-248. doi:10.1111/imj.12629.
8 Burns PB, Rohrich RJ, Chung KC. The levels of evidence and their role in evidence-based medicine. *Plast Reconstr Surg*. 2011;128(1):305-310. doi:10.1097/PRS.0b013e318219c171.
9 FDA. FDA facts: biomarkers and surrogate endpoints. U.S. Food and Drug Administration. https://www.fda.gov/aboutfda/innovation/ucm512503.htm. Published 2017. Accessed July 30, 2018.
10 Pollock ML. Submaximal and maximal working capacity of elite distance runners. Part I: Cardiorespiratory aspects. *Ann N Y Acad Sci*. 1977;301(1):310-322. doi:10.1111/j.1749-6632.1977.tb38209.x.
11 Joyner MJ, Coyle EF. Endurance exercise performance: The physiology of champions. *J Physiol*. 2008;586(1):35-44. doi:10.1113/jphysiol.2007.143834.
12 Sjödin B, Svedenhag J. Applied physiology of marathon running. *Sports Med*. 1985;2:83-99.
13 Lucia A, Hoyos J, Santalla A, Rez MP, Chicharro JL. Kinetics of VO2 in professional cyclists. *Med Sci Sport Exerc*. 2002;34(2):320-325.
14 Jones AM. The physiology of the world record holder for the women's marathon. *Int J Sports Sci Coach*. 2006;1(2):101-116. doi:10.1260/174795406777641258.
15 Coyle EF. Improved muscular efficiency displayed as Tour de France champion matures. *J Appl Physiol*. 2005;98(6):2191-2196. doi:10.1152/japplphysiol.00216.2005.
16 Jones AM. A five year physiological case study of an Olympic runner. *Br J Sports Med*. 1998;32(1):39-43. doi:10.1136/bjsm.32.1.39.
17 Noakes T. Testing for maximum oxygen consumption has produced a brainless model of human exercise performance. *Br J Sports Med*. 2008;42(7):551-555. doi:10.1136/bjsm.2008.046821.
18 Jeukendrup A, Saris WH, Brouns F, Kester AD. A new validated endurance performance test. *Med Sci Sports Exerc*. 1996;28:266-270.
19 Laursen PB, Francis GT, Abbiss CR, Newton MJ, Nosaka K. Reliability of time-to-exhaustion versus time-trial running tests in runners. *Med Sci Sports Exerc*. 2007;39(8):1374-1379. doi:10.1249/mss.0b013e3180601 0f5.
20 Laursen PB, Rhodes EC, Langill RH, McKenzie DC, Taunton JE. Relationship of exercise test variables to cycling performance in an ironman triathlon. *Eur J Appl Physiol*. 2002;87(4-5):433-440. doi:10.1007/s00421-002-0659-4.
21 van Someren K, S Palmer G. Prediction of 200-m sprint kayaking performance. *Can J Appl Physiol*. 2003;28:505-517.
22 Riechman SE, Zoeller RF, Balasekaran G, Goss FL, Robertson RJ. Prediction of 2000 m indoor rowing performance using a 30 s sprint and maximal oxygen uptake. *J Sports Sci*. 2002;20(9):681-687.
23 Hartgens F, Kuipers H. Effects of androgenic anabolic steroids in athletes . *Sports Med*. 2017;34(8):1-2.
24 Bhasin S, Storer TW, Berman N, et al. The effects of supraphysiologic doses of testosterone on muscle size and strength in normal men. *N Engl J Med*. 1996;335(1):1-7. doi:10.1056/NEMJ199607043350101.
25 Giorgi A, Weatherby RP, Murphy PW. Muscular strength, body composition and health responses to the use of testosterone enanthate: a double blind study. *J Sci Med Sport*. 1999;2(4):341-355. doi:10.1016/S1440-2440(99)80007-3.
26 Friedl K, Dettori J, Jr CH. Comparison of the effects of high dose testosterone and 19-nortestosterone to a replacement dose of testosterone on strength and body composition in normal men. *J Steroid Biochem Mol Biol*. 1991;40(4):607-612.
27 Stromme SB, Meen HD, Aakvaag A. Effects of an androgenic-anabolic steroid on strength development and plasma testosterone levels in normal males. *Med Sci Sports Exerc*. 1974;6(3):203-207.
28 Hervey GR, Hutchinson I, Knibbs A V., et al. "Anabolic" effects of methandienone in men undergoing athletic training. *Lancet*. 1976;308(7988):699-702. doi:10.1016/S0140-6736(76)90001-5.
29 Baume N, Schumacher YO, Sottas PE, et al. Effect of multiple oral doses of androgenic anabolic steroids on endurance performance and serum indices of physical stress in healthy male subjects. *Eur J Appl Physiol*. 2006;98(4):329-340. doi:10.1007/s00421-006-0271-0.
30 Lundby C, Millet GP, Calbet JA, Bärtsch P, Subudhi AW. Does 'altitude training' increase exercise performance in elite athletes? *Br J Sports Med*. 2012;46(11):792 LP-795.
31 Heuberger JAAC, Cohen Tervaert JM, Schepers FML, et al. Erythropoietin doping in cycling: Lack of evidence for efficacy and a negative risk-benefit. *Br J Clin Pharmacol*. 2013;75(6). doi:10.1111/bcp.12034.
32 Birkeland K, Stray-Gundersen J, Hemmersbach P, Hallén J, Haug E, Bahr R. Effect of rHEPO administration on serum levels of sTfR and cycling performance. *Med Sci Sport Exerc*. 2000;32:1238-1243.
33 Wilkerson DP, Rittweger J, Berger NJ, Naish PF, Jones AM. Influence of recombinant human erythropoietin treatment on pulmonary O_2 uptake kinetics during exercise in humans. *J Physiol*. 568(2):639-652. doi:10.1113/jphysiol.2005.089920.
34 Ninot G, Connes P, Caillaud C. Effects of recombinant human erythropoietin injections on physical self in endurance athletes. *J Sports Sci*. 2006;24:383-391.
35 Parisotto R, Gore CJ, Emslie KR, et al. A novel method utilising markers of altered erythropoiesis for the detection of recombinant human erythropoietin abuse in athletes. *Haematologica*. 2000;85(6):564 LP-572.
36 Connes P, Perrey S, Varray A, Préfaut C, Caillaud C. Faster oxygen uptake kinetics at the onset of submaximal cycling exercise following 4 weeks recombinant human erythropoietin (r-HuEPO) treatment. *Pflügers Arch*. 2003;447(2):231-238. doi:10.1007/s00424-003-1174-0.

37 Annaheim S, Jacob M, Krafft A, Breymann C, Rehm M, Boutellier U. RHEPO improves time to exhaustion by non-hematopoietic factors in humans. 116(3):623-633. doi:10.1007/s00421-015-3322-6.

38 Clark B, Woolford SM, Eastwood A, Sharpe K, Barnes PG, Gore CJ. Temporal changes in physiology and haematology in response to high- and micro-doses of recombinant human erythropoietin. *Drug Test Anal*. 2017;9(10):1561-1571. doi:10.1002/dta.2176.

39 Heuberger JAAC, Rotmans JI, Gal P, et al. Effects of erythropoietin on cycling performance of well trained cyclists: a double-blind, randomised, placebo-controlled trial. 2017;3026(17):1-13.

40 Bowie E, Hurley P. Cobalt chloride in the treatment of refractory anaemia in patients undergoing long-term haemodialysis. *Aust N Z J Med*. 5(4):306-313. doi:10.1111/j.1445-5994.1975.tb03263.x.

41 Edwards MS, Curtis JR. Use of cobaltous chloride in anæmia of maintenance-hæmodialysis patients. *Lancet*. 1971;298(7724):582-583. doi:10.1016/S0140-6736(71)80127-7.

42 Ebert B, Jelkmann W. Intolerability of cobalt salt as erythropoietic agent. *Drug Test Anal*. 2014;6(3):185-189. doi:10.1002/dta.1528.

43 Stoppe C, Ney J, Brenke M, et al. Sub-anesthetic xenon increases erythropoietin levels in humans: a randomized controlled trial. *Sports Med*. 2016;46(11):1753-1766. doi:10.1007/s40279-016-0505-1.

44 Balachandran A, Streiner DL, Signorile JF. Comment on "Sub-anesthetic xenon increases erythropoietin levels in humans: a randomized controlled trial." *Sports Med*. 2017;47(2):379. doi:10.1007/s40279-016-0659-x.

45 Gupta N, Wish JB. Hypoxia-inducible factor prolyl hydroxylase inhibitors: a potential new treatment for anemia in patients with CKD. *Am J Kidney Dis*. 2017;69(6):815-826. doi:10.1053/j.ajkd.2016.12.011.

46 Locatelli F, Fishbane S, Block GA, Macdougall IC. Targeting hypoxia-inducible factors for the treatment of anemia in chronic kidney disease patients. *Am J Nephrol*. 2017;45(3):187-199.

47 Macdougall IC. New anemia therapies: translating novel strategies from bench to bedside. *Am J Kidney Dis*. 2012;59(3):444-451. doi:10.1053/J.AJKD.2011.11.013.

48 Imagawa S, Matsumoto K, Horie M, et al. Does K-11706 enhance performance and why? *Int J Sport Med*. 2007;28(11):928-933. doi:10.1055/s-2007-964988.

49 Komrokji R, Garcia-Manero G, Ades L, et al. Sotatercept with long-term extension for the treatment of anaemia in patients with lower-risk myelodysplastic syndromes: a phase 2, dose-ranging trial. *Lancet Haematol*. 2018;5(2):e63-e72. doi:10.1016/S2352-3026(18)30002-4.

50 Platzbecker U, Germing U, Götze KS, et al. Luspatercept for the treatment of anaemia in patients with lower-risk myelodysplastic syndromes (PACE-MDS): a multicentre, open-label phase 2 dose-finding study with long-term extension study. *Lancet Oncol*. 2017;18(10):1338-1347. doi:10.1016/S1470-2045(17)30615-0.

51 Boesch S, Nachbauer W, Mariotti C, et al. Safety and tolerability of carbamylated erythropoietin in Friedreich's ataxia. *Mov Disord*. 2014;29(7):935-939. doi:10.1002/mds.25836.

52 Handelsman DJ. The rationale for banning human chorionic gonadotropin and estrogen blockers in sport. *J Clin Endocrinol Metab*. 2006;91(5):1646-1653. doi:10.1210/jc.2005-2569.

53 Padrón RS, Wischusen J, Hudson B, Burger HG, Kretser DMD. Prolonged biphasic response of plasma testosterone to single intramuscular injections of human chorionic gonadotropin. *J Clin Endocrinol Metab*. 1980;50(6):1100-1104. doi:10.1210/jcem-50-6-1100.

54 Soetens E, De Meirleir K, E Hueting J. No influence of ACTH on maximal performance. *Psychopharmacology (Berl)*. 1995;118:260-266.

55 Baume N, Steel G, Edwards T, Thorstensen E, Miller BF. No variation of physical performance and perceived exertion after adrenal gland stimulation by synthetic ACTH (Synacthen®) in cyclists. *Eur J Appl Physiol*. 2008;104(4):589-600. doi:10.1007/s00421-008-0802-y.

56 Woodhouse LJ, Mukherjee A, Shalet SM, Ezzat S. The influence of growth hormone status on physical impairments, functional limitations, and health-related quality of life in adults. *Endocr Rev*. 2006;27(3):287-317. doi:10.1210/er.2004-0022.

57 Liu H, Bravata DM, Olkin I, et al. Review systematic review : the effects of growth hormone on athletic performance. *Ann Intern Med*. 2008;148(10):747-758. doi:148:747-758.

58 Yarasheski KE, Campbell JA, Smith K, Rennie MJ, Holloszy JO, Bier DM. Effect of growth hormone and resistance exercise on muscle growth in young men. *Am J Physiol Metab*. 1992;262(3):E261-E267. doi:10.1152/ajpendo.1992.262.3.E261.

59 Deyssig R, Frisch H, Blum W, Waldhor T. Effect of growth hormone treatment on hormonal parameters, body composition and strength in athletes. *Acta Endocrinol (Copenh)*. 1993;128(4):313-318.

60 Berggren A, Ehrnborg C, Rosén T, Ellegård L, Bengtsson BÅ, Caidahl K. Short-term administration of supraphysiological recombinant human growth hormone (GH) does not increase maximum endurance exercise capacity in healthy, active young men and women with normal GH-insulin-like growth factor I axes. *J Clin Endocrinol Metab*. 2005;90(6):3268-3273. doi:10.1210/jc.2004-1209.

61 Lange KHW, Larsson B, Flyvbjerg A, et al. Acute growth hormone administration causes exaggerated increases in plasma lactate and glycerol during moderate to high intensity bicycling in trained young men. *J Clin Endocrinol Metab*. 2002;87(11):4966-4975. doi:10.1210/jc.2001-011797.

62 Meinhardt U, Nelson AE, Hansen JL, Birzniece V, Clifford D, Leung K. The effects of growth hormone on body composition and physical performance in recreational athletes. *Ann Intern Med*. 2010;152:568-577. doi:10.7326/0003-4819-152-9-201005040-00007.

63 Creaney L, Hamilton B. Growth factor delivery methods in the management of sports injuries: the state of play. *Br J Sports Med*. 2008;42(5):314 LP-320.

64 Morishita R, Makino H, Aoki M, et al. Phase I/IIa clinical trial of therapeutic angiogenesis using hepatocyte growth factor gene transfer to treat critical limb ischemia. *Arterioscler Thromb Vasc Biol*. 2011;31(3):713-720. doi:10.1161/ATVBAHA.110.219550.

65 Steed DL. Clinical evaluation of recombinant human platelet -derived growth factor for the treatment of lower extremity diabetic ulcers. *J Vasc Surg*. 2018;21(1):71-81. doi:10.1016/S0741-5214(95)70245-8.

66 Ruff D, Crockford D, Girardi G, Zhang Y. A randomized, placebo-controlled, single and multiple dose study of intravenous thymosin β4 in healthy volunteers. *Ann N Y Acad Sci*. 2010;1194(1):223-229. doi:10.1111/j.1749-6632.2010.05474.x.

67 Guha N, Nevitt SP, Francis M, et al. The effects of recombinant human insulin-like growth factor-I/insulin-like growth factor binding protein-3 administration on body bomposition and physical fitness in recreational athletes. *J Clin Endocrinol Metab*. 2015;100(8):3126-3131. doi:10.1210/jc.2015-1996.

68 Pluim BM, de Hon O, Staal JB, et al. β_2-Agonists and physical performance: a systematic review and meta-analysis of randomized controlled trials. *Sports Med*. 2011;41(1):39-57. doi:10.2165/11537540-000000000-00000.

69 Cairns SP, Borrani F. β-Adrenergic modulation of skeletal muscle contraction: key role of excitation–contraction coupling. *J Physiol*. 2015;593(21):4713-4727. doi:10.1113/JP270909.

70 Hostrup M, Kalsen A, Ørtenblad N, et al. β2-Adrenergic stimulation enhances Ca2+ release and contractile properties of skeletal muscles, and counteracts exercise-induced reductions in Na+-K+-ATPase Vmax in trained men. *J Physiol*. 2014;592(24):5445-5459. doi:10.1113/jphysiol.2014.277095.

71 Hostrup M, Kalsen A, Bangsbo J, Hemmersbach P, Karlsson S, Backer V. High-dose inhaled terbutaline increases muscle strength and enhances maximal sprint performance in trained men. *Eur J Appl Physiol*. 2014;114(12):2499-2508. doi:10.1007/s00421-014-2970-2.

72 Hostrup M, Kalsen A, Auchenberg M, Bangsbo J, Backer V. Effects of acute and 2-week administration of oral salbutamol on exercise performance and muscle strength in athletes. *Scand J Med Sci Sports*. 2016;26(1):8-16. doi:10.1111/sms.12298.

73 Decorte N, Bachasson D, Guinot M, et al. Impact of salbutamol on neuromuscular function in endurance athletes. *Med Sci Sports Exerc*. 2013;45(March). doi:10.1249/MSS.0b013e3182951d2d.

74 Mauras N, Lima J, Patel D, et al. Pharmacokinetics and dose finding of a potent aromatase inhibitor, aromasin (exemestane), in young males. *J Clin Endocrinol Metab*. 2003;88(12):5951-5956. doi:10.1210/jc.2003-031279.

75 Duschek EJ, Gooren L, Netelenbos J. Effects of raloxifene on gonadotrophins, sex hormones, bone turnover and lipids in healthy elderly men. *Eur J Endocrinol*. 2004;150:539-546.

76 Guay AT, Bansal S, Heatley GJ. Effect of raising endogenous testosterone levels in impotent men with secondary hypogonadism: double blind placebo-controlled trial with clomiphene citrate. *J Clin Endocrinol Metab*. 1995;80(12):3546-3552. doi:10.1210/jcem.80.12.8530597.

77 Tenover J, Dahl K, Hseuh A, Lim P, Matsumoto A BW. Serum bioactive and immunoreactive follicle-stimulating hormone levels and the response to clomiphene in healthy young and elderly men. *J Clin Endocrinol Metab*. 1987;64(6):1103-1108. doi:10.1210/jcem-64-6-1103.

78 Fedoruk MN, Rupert JL. Myostatin inhibition: a potential performance enhancement strategy? *Scand J Med Sci Sports*. 2008;18(2):123-131. doi:10.1111/j.1600-0838.2007.00759.x.

79 Saitoh M, Ishida J, Ebner N, D Anker S, von Haehling S. Myostatin inhibitors as pharmacological treatment for muscle wasting and muscular dystrophy. *JCSM Clin Reports*. 2017;2.

80 Narkar VA, Downes M, Yu RT, et al. AMPK and PPARδ agonists are exercise mimetics. *Cell*. 2008;134(3):405-415. doi:https://doi.org/10.1016/j.cell.2008.06.051.

81 Tan CK, Zhuang Y, Wahli W. Synthetic and natural peroxisome proliferator-activated receptor (PPAR) agonists as candidates for the therapy of the metabolic syndrome. *Expert Opin Ther Targets*. 2017;21(3):333-348. doi:10.1080/14728222.2017.1280467.

82 Niederberger E, S King T, Russe O, Geisslinger G. Activation of AMPK and its impact on exercise capacity. *Sports Med*. 2015;45.

83 Braun B, Eze P, Stephens BR, et al. Impact of metformin on peak aerobic capacity. *Appl Physiol Nutr Metab*. 2008;33(1):61-67. doi:10.1139/H07-144.

84 Johnson ST, Robert C, Bell GJ, Bell RZ, Lewanczuk RZ, Boulé NG. Acute effect of metformin on exercise capacity in active males. *Diabetes, Obes Metab*. 2008;10(9):747-754. doi:10.1111/j.1463-1326.2007.00805.x.

85 Kuipers H, Ruijsch Van Dugteren G. Letter to the editors - The prohibited list and cheating in sport [1]. *Int J Sports Med*. 2006;27(1):80-82. doi:10.1055/s-2005-873069.

86 Greenblatt HK, Greenblatt DJ. Meldonium (Mildronate): a performance-enhancing drug? *Clin Pharmacol Drug Dev*. 2016;5(3):167-169. doi:10.1002/cpdd.264.

87 Vitale C, Marazzi G, Pelliccia F, et al. Trimetazidine improves exercise performance in patients with peripheral arterial disease. *Pharmacol Res*. 2011;63(4):278-283. doi:https://doi.org/10.1016/j.phrs.2011.01.003.

88 Cadwallader AB, de la Torre X, Tieri A, Botrè F. The abuse of diuretics as performance-enhancing drugs and masking agents in sport doping: pharmacology, toxicology and analysis. *Br J Pharmacol*. 2010;161(1):1-16. doi:10.1111/j.1476-5381.2010.00789.x.

89 Caldwell JE, Ahonen E, Nousiainen U. Differential effects of sauna-, diuretic-, and exercise-induced hypohydration. *J Appl Physiol*. 1984;57(4):1018-1023. doi:10.1152/jappl.1984.57.4.1018.

90 Armstrong LE, Costill DL, Fink WJ. Influence of diuretic-induced dehydration on competitive running performance. *Med Sci Sports Exerc*. 1985;17(4):456-461.

91 Kowalchuk JM, Heigenhauser GJ, Sutton JR, Jones NL. Effect of acetazolamide on gas exchange and acid-base control after maximal exercise. *J Appl Physiol*. 1992;72(1):278-287.

92 Stager JM, Tucker A, Cordain L, Engebretsen BJ, Brechue WF, Matulich CC. Normoxic and acute hypoxic exercise tolerance in man following acetazolamide. *Med Sci Sports Exerc*. 1990;22(2):178-184.

93 Schoene RB, Bates PW, Larson EB, Pierson DJ. Effect of acetazolamide on normoxic and hypoxic exercise in humans at sea level. *J Appl Physiol Respir Env Exerc Physiol*. 1983;55(6):1772-1776.

94 Fulco CS, Muza SR, Ditzler D, Lammi E, Lewis SF, Cymerman A. Effect of acetazolamide on leg endurance exercise at sea level and simulated altitude. *Clin Sci*. 2006;110(6):683-692. doi:10.1042/CS20050233.

95 Chandler J V, Blair SN. The effect of amphetamines on selected physiological components related to athletic success. *Med Sci Sports Exerc*. 1980;12(1):65-69.

96 Swart J, Lamberts RP, Lambert MI, et al. Exercising with reserve: evidence that the central nervous system regulates prolonged exercise performance. *Br J Sports Med*. 2009;43(10):782-788. doi:10.1136/bjsm.2008.055889.

97 Roelands B, Hasegawa H, Watson P, et al. The effects of acute dopamine reuptake inhibition on performance. *Med Sci Sports Exerc*. 2008;40:879-885.

98 Dufka F, Galloway G, Baggott M, Mendelson J. The effects of inhaled L-methamphetamine on athletic performance while riding a

stationary bike: a randomised placebo-controlled trial. *Br J Sports Med*. 2009;43(11):832-835. doi:10.1136/bjsm.2008.048348.
99 Bell DG, Jacobs IRA, Ellerington K. Effect of caffeine and ephedrine ingestion on anaerobic exercise performance. *Med Sci Sports Exerc*. 2001;33(8):1399-1403.
100 Jacobs I, Pasternak H, Bell DG. Effects of ephedrine, caffeine, and their combination on muscular endurance. *Med Sci Sports Exerc*. 2003;35(6):987-994. doi:10.1249/01.MSS.0000069916.49903.70.
101 Gill ND, Shield A, Blazevich AJ, Zhou S, Weatherby RP. Muscular and cardiorespiratory effects of pseudoephedrine in human athletes. *Br J Clin Pharmacol*. 2000;50(3):205-213. doi:10.1046/j.1365-2125.2000.00252.x.
102 Chester N, Reilly T, Mottram D. Physiological, subjective and performance effects of pseudoephedrine and phenylpropanolamine during endurance running exercise. *Int J Sports Med*. 2003;24:3-8.
103 Hodges ANH, Lynn BM, Bula JE, Donaldson MG, Dagenais MO, McKenzie DC. Effects of pseudoephedrine on maximal cycling power and submaximal cycling efficiency. / Effets de la pseudo-ephedrine sur la puissance maximale et sur l ' efficacite d ' un exercice maximal et d ' un autre sous-maximal sur bicyclette ergometrique. *Med Sci Sport Exerc*. 2003;35(8):1316-1319.
104 Hodges K, Hancock S, Currell K, Hamilton B, Jeukendrup AE. Pseudoephedrine enhances performance in 1500m runners. *Med Sci Sport Exerc*. 2006;38(2):329-334. doi:10.1249/01.mss.0000183201.79330.9c.
105 Amann M, Proctor LT, Sebranek JJ, Pegelow DF, Dempsey JA. Opioid-mediated muscle afferents inhibit central motor drive and limit peripheral muscle fatigue development in humans. *J Physiol*. 2009;587(1):271-283. doi:10.1113/jphysiol.2008.163303.
106 Holgado D, Hopker J, Sanabria D, Zabala M. Analgesics and sport performance: beyond the pain-modulating effects. *PM&R*. 2018;10(1):72-82. doi:https://doi.org/10.1016/j.pmrj.2017.07.068.
107 Huestis MA. Cannabis (Marijuana) — Effects on human behavior and performance. *Forensic Sci Rev*. 2002;14(1-2):15-60.
108 Hilderbrand RL. High-performance sport, marijuana, and cannabimimetics. *J Anal Toxicol*. 2011;35(9):624-637. doi:10.1093/anatox/35.9.624.
109 Huestis MA, Mazzoni I, Rabin O. Cannabis in sport: anti-doping perspective. *Sports Med*. 2011;41(11):949-966. doi:10.2165/11591430-000000000-00000.
110 Trinh K V, Diep D, Robson H. Marijuana and its effects on athletic performance: a systematic review. *Clin J Sport Med*. 2018;28(4):350-357.
111 Kennedy MC. Cannabis: Exercise performance and sport. A systematic review. *J Sci Med Sport*. 2017;20(9):825-829. doi:https://doi.org/10.1016/j.jsams.2017.03.012.
112 Steadward RD, Singh M. The effects of smoking marihuana on physical performance. *Med Sci Sports*. 1975;7(4):309-311.
113 Collomp K, Arlettaz A, Buisson C, Lecoq A-M, Mongongu C. Glucocorticoid administration in athletes: Performance, metabolism and detection. *Steroids*. 2016;115:193-202. doi:https://doi.org/10.1016/j.steroids.2016.09.008.
114 Marquet P, Lac G, Chassain AP, Habrioux G, Gallen FX. Dexamethasone in resting and exercising men. I. Effects on bioenergetics, minerals, and related hormones. *J Appl Physiol*. 1999;87(1):175-182.
115 Kuipers H, Van't Hullenaar GAC, Pluim BM, et al. Four weeks' corticosteroid inhalation does not augment maximal power output in endurance athletes. *Br J Sports Med*. 2008;42(11):568-571. doi:10.1136/bjsm.2007.042572.
116 Nordsborg N, Ovesen J, Thomassen M, et al. Effect of dexamethasone on skeletal muscle Na^+,K^+ pump subunit specific expression and K^+ homeostasis during exercise in humans. *J Physiol*. 2008;586(5):1447-1459. doi:10.1113/jphysiol.2007.143073.
117 Casuso RA, Melskens L, Bruhn T, Secher NH, Nordsborg NB. Glucocorticoids improve high-intensity exercise performance in humans. *Eur J Appl Physiol*. 2014;114(2):419-424. doi:10.1007/s00421-013-2784-7.
118 Zorgati H, Prieur F, Vergniaud T, et al. Ergogenic and metabolic effects of oral glucocorticoid intake during repeated bouts of high-intensity exercise. *Steroids*. 2014;86:10-15. doi:10.1016/j.steroids.2014.04.008.
119 Arlettaz A, Collomp K, Portier H, Lecoq AM, Pelle A, De Ceaurriz J. Effects of acute prednisolone intake during intense submaximal exercise. *Int J Sports Med*. 2006;27(9):673-679. doi:10.1055/s-2005-872826.
120 Arlettaz A, Collomp K, Portier H, et al. Effects of acute prednisolone administration on exercise endurance and metabolism. *Br J Sports Med*. 2008;42(4):250-254. doi:10.1136/bjsm.2007.039040.
121 Arlettaz A, Portier H, Lecoq AM, Rieth N, De Ceaurriz J, Collomp K. Effects of short-term prednisolone intake during submaximal exercise. *Med Sci Sports Exerc*. 2007;39(9):1672-1678. doi:10.1249/mss.0b013e3180dc992c.
122 Collomp K, Arlettaz A, Portier H, et al. Short-term glucocorticoid intake combined with intense training on performance and hormonal responses. *Br J Sports Med*. 2008;42(12):983 LP-988.
123 Le Panse B, Thomasson R, Jollin L, et al. Short-term glucocorticoid intake improves exercise endurance in healthy recreationally trained women. *Eur J Appl Physiol*. 2009;107(4):437-443. doi:10.1007/s00421-009-1149-8.
124 Kruse P, Ladefoged J, Nielsen U, Paulev PE, Sørensen JP. Beta-blockade used in precision sports: effect on pistol shooting performance. *J Appl Physiol*. 1986;61:417-420.

TABLE 1 OVERVIEW OF ALL SUBSTANCE CLASSES AND EVIDENCE FOR PERFORMANCE ENHANCEMENT

Substance class	Well-designed studies?	Studies with trained athletes?	Relevant performance parameters showing improvement	Number of trained athletes in studies with relevant performance parameters
S0. Non-approved substances				
	No	NA	NA	NA
S1. Anabolic agents				
	Yes	Yes	Muscle strength when combined with strength training • +10-20% bench-press and squatting power [24] • +12% bench press power [25] • +23% and +12% elbow flexion and knee extension [26]	Total: 91 • 40 [24] • 21 [25] • 30 [26]
S2. Peptide hormones, growth factors, related substances and mimetics				
Erythropoietin-receptor agonists	Yes	Yes	No evidence for effects on relevant endurance parameters, only on VO_{2max}, maximal power output and time to exhaustion	Total: 161 • 20 [32] • 11 [34] • 27 [35] • 16 [36] • 40 [37] • 47 [39]
Hypoxia-inducible factor activating agents	No	NA	NA	NA
GATA inhibitors	No	NA	NA	NA
TGF-beta inhibitors	No	NA	NA	NA
Innate repair receptor agonists	No	NA	NA	NA
Chorionic gonadotrophin and luteinizing hormone and their releasing factors	No	NA	NA	NA
Corticotrophins and their releasing factors	Yes	Yes	No	Total: 24 • 16 [54] • 8 [55]
Growth hormone, its fragments and releasing factors	Yes	No	Sprint performance • +3.9% Wingate sprint capacity [62] in untrained subjects	Total: 123 • 22 [59] • 30 [60] • 7 [61] • 64 untrained [62]
Growth factors and growth factor modulators	Yes	No	NA	NA

(continuation on next page)

TABLE 1 *(continuation of previous page)*

S3. Beta-2 agonists				
	Yes	Yes	Peak sprint power, muscle strength • +6% maximal voluntary isometric contraction, +11% peak twitch force [70] • +8.4% quadriceps muscle strength, +2.2% and +3.3% Wingate peak and mean power, +3% Wingate total work [71] • +4% (single dose) and +6% (multiple dose) Wingate peak power [72]	Total: 39 • 10 [70] • 9 [71] • 20 [72]
S4. Hormone and metabolic modulators				
Aromatase inhibitors	No	NA	NA	NA
Selective estrogen receptor modulators	No	NA	NA	NA
Other anti-estrogenic substances	No	NA	NA	NA
Agents modifying myostatin function(s)	No	NA	NA	NA
Metabolic modulators	No	NA	NA	NA
S5. Diuretics and masking agents				
	Yes	Yes	No	Total: 70 • 62 [89] • 8 [90]
S6. Stimulants				
	Yes	Yes	Muscle strength, sprint performance, 1500 meter run time • +9% knee extension strength, +3% peak Wingate sprint performance [101] • 2% decrease 1500 meter run time [104]	Total: 58 • 9 [97] • 22 [101] • 9 [102] • 11 [103] • 7 [104]
S7. Narcotics				
	Yes	Yes	No	Total: 8 • 8 [105]
S8. Cannabinoids				
	Yes	No	NA	NA
S9. Glucocorticoids				
	Yes	Yes	One-legged hopping force • +11% one-legged hopping maximal force [118]	Total: 10 • 10 [118]
P1. Beta-blockers (prohibited in particular sports)				
	Yes	Yes	Shooting performance • +13.4% shooting performance [124]	Total: 33 33 [124]

NA, not applicable; TGF, transforming growth factor; VO_{2max}, maximal oxygen consumption

3 ERYTHROPOIETIN DOPING IN CYCLING: LACK OF EVIDENCE FOR EFFICACY AND A NEGATIVE RISK–BENEFIT

BRITISH JOURNAL OF CLINICAL PHARMACOLOGY 2013, 75(6):1406-1421

Jules AAC Heuberger, Joost M Cohen Tervaert, Femke ML Schepers, AD Bastiaan Vliegenthart, Joris I Rotmans, Johannes MA Daniels, Jacobus Burggraaf and Adam F Cohen

ABSTRACT

Imagine a medicine that is expected to have very limited effects based upon knowledge of its pharmacology and (patho)physiology and that is studied in the wrong population, with low-quality studies that use a surrogate end-point that relates to the clinical end-point in a partial manner at most. Such a medicine would surely not be recommended. The use of recombinant human erythropoietin (rHuEPO) to enhance performance in cycling is very common. A qualitative systematic review of the available literature was performed to examine the evidence for the ergogenic properties of this drug, which is normally used to treat anaemia in chronic renal failure patients. The results of this literature search show that there is no scientific basis from which to conclude that rHuEPO has performance-enhancing properties in elite cyclists. The reported studies have many shortcomings regarding translation of the results to professional cycling endurance performance. Additionally, the possibly harmful side-effects have not been adequately researched for this population but appear to be worrying, at least. The use of rHuEPO in cycling is rife but scientifically unsupported by evidence, and its use in sports is medical malpractice. What its use would have been, if the involved team physicians had been trained in clinical pharmacology and had investigated this properly, remains a matter of speculation. A single well-controlled trial in athletes in real-life circumstances would give a better indication of the real advantages and risk factors of rHuEPO use, but it would be an oversimplification to suggest that this would eradicate its use.

INTRODUCTION

Sport is big business

The summer of 2012 was an intensive summer of sports. From all these events, it is clear that sports play a very important role in our society as it brings people together, gives pleasure, keeps people healthy and can bring professional athletes fame and honour.

Sport has grown to be so important that large amounts of money are now involved and the will and pressure to win have steadily increased. Cheating has therefore become a threat to all sports, with some sports being more susceptible to it than others. Cheating by use of medicines has understandably taken place outside the realm of clinical pharmacology and evidence based medicine. We question if this is correct, as uncontrolled use of a substance induces risks for the users, irrespective of such a substance being used legally or illegally. In this review we will focus on the use of recombinant human erythropoietin (rHuEPO) in cycling, a sport that has had many reports of cheating, culminating in the last decennia, with many suspicions and suspensions. We will address the question if the current available evidence even warrants the widespread use of this substance. Many of the big champions in cycling have been associated with, or suspended for use of (blood) doping. In the Tour de France of 1998 the entire Festina team, as well as the TVM team, were taken out of the race on suspicion of rHuEPO use. This Tour was later given the name 'Tour du Dopage' and many confessions of systematic doping (i.e. rHuEPO) use throughout the peloton were given. In spite of this, later champions in the Tour de France, Giro d'Italia and Vuelta a España have also been suspended because of proof of blood doping, but the Code of Silence also called 'omerta', was never broken. Seven years after the last of seven consecutive Tour de France wins, one of the most successful road cyclists ever, Lance Armstrong, has been suspended for life by the United States Anti-Doping Agency (USADA) on charges of doping (e.g. rHuEPO) use and trafficking in the biggest doping case ever, backed by confessions of many of his teammates.[1]

Knowledge of both the effects and side effects of rHuEPO in this population is essential, especially with so many misconceptions among the people involved. Firstly, if the effects are not pronounced, the motives for misuse will be less strong. Secondly, even if the effects are pronounced, knowledge of the potentially

dangerous side effects needs to be communicated to the cyclists, who are likely to be under severe pressure to use performance enhancing agents, together with the coaches and physicians supervising them.[1]

Physiology of erythropoietin

Erythropoietin (EPO) is a (glyco)protein that is mainly involved in erythropoiesis, the (re-)generation of erythrocytes, or red blood cells. Red blood cells are cells without a nucleus and transport oxygen through the blood. Due to a lack of ability to repair themselves without a nucleus and other cellular machinery, erythrocytes have a life span of approximately 120 days in the circulation and after that need to be replaced.[2] The spleen removes the old erythrocytes (2-3 million every second) and to keep oxygen carrying capacity of the blood at a steady level, constant erythropoiesis is necessary.[2] Erythropoiesis starts in the bone marrow, where red blood cells originate from pluripotent stem cells.[3] These stem cells continuously make identical copies of themselves and in that way create progenitor cells for, among others, erythrocytic cells.[3] These cells go through different stages, one of which is the burst-forming unit-erythroid (BFU-e). This cell type matures into a colony-forming unit-erythroid (CFU-e), which in turn forms the proerythroblast, which divides four times into 16 reticulocytes, later developing into mature red blood cells.[3] The first report of a factor influencing this red blood cell production was by Carnot and Deflandre,[4] who called it "hemopoietine". This factor, now called erythropoietin, is a hormone of 165 amino acids with four glycosylation sites and a molecular weight of 30,400 and a carbohydrate content of 40%.[5,6] Under normal (non-hypoxic) conditions the concentration in blood is relatively constant at 5 pmol/L, essential to stimulate cells in the bone marrow to produce new erythrocytes, compensating for the physiological demise of erythrocytes.[3] This level is equal to ~20 mU/mL when erythropoietin is quantified as 'international units' (IU), assuming a specific activity of 130,000 IU/mg. The cells that are the main target for the hormone are the CFU-e's and proerythroblasts, containing the highest density of erythropoietin receptors (EpoR).[7] The main effect of EPO is on CFU-e's, as it promotes survival of these cells.[8] One of the pathways involved in this process, activated by EPO, is the cell proliferation pathway of Ras/MAP kinase.[9,10] After binding of EPO to its receptor dimerization of two EpoR molecules occurs, starting the intracellular signalling leading to proliferation of CFU-e's.[11,12]

Production and metabolism

The main EPO producing organ in humans is the kidney,[13,14] where peritubular interstitial cells are the cells governing this production.[15,16] Production of EPO is highly regulated, and it has been shown that baseline EPO levels can increase up to a 1,000-fold in low blood oxygen content, for example in severe anaemia.[3] EPO production is therefore highly dependent on blood oxygen levels, with hypoxia increasing EPO production, irrespective of the cause of reduced tissue oxygen supply.[3] This process takes approximately 1.5-2 hours before EPO levels start increasing linearly, reflecting the time of signal transduction and hormone synthesis and secretion. Peak EPO concentrations after hypoxia are reached within 48 hours, with concentration being dependent on the severity of hypoxia.[3] However, only moderately elevated serum concentrations of EPO seem to be sufficient to maintain an increased erythropoiesis rate.[17] The proposed oxygen-sensing mechanism regulating erythropoietin production involves the hypoxia-inducible factor (HIF), a transcription factor.[18] HIF expression is seen in hypoxia exposed cells within 30 minutes,[19] after which the heterodimeric protein travels to the nucleus to activate the EPO enhancer,[20] inducing EPO transcription. In the presence of oxygen this factor is hydroxylated, suppressing the activity and promoting degradation.[21] Another pathway involved in EPO production is the kinase C pathway, activated through adenosine. This non-HIF transcriptional factor also increases EPOmRNA expression.[22] GATA-2 inhibits the EPO promoter, and is a third regulational pathway of EPO. GATA inhibitors can therefore also enhance EPO production.[23] After hypoxia-induced erythropoietin production a rise in red blood cells and haematocrit (Hct) is seen after 60-70 hours,[24] corresponding to the time course of CFU-e differentiation into mature erythrocytes.[25] This also is in line with the observation that endogenous erythropoietin has an estimated half-life much shorter than that, of approximately 5.2 hours.[26] So the clearance of EPO is, like many other glycoproteins, rather slow. This is mainly due to the terminal sialic acid residues, preventing galactose receptor binding, internalization and degradation in the liver. Indeed, it has been shown that desialated EPO results in rapid hepatic clearance,[27] but this process almost does not occur *in vivo* and therefore plays only a minor role in endogenous EPO clearance.[28] Also renal clearance plays a minor role, as the disappearance rate does not change markedly in the anephric state.[29] The major elimination route for erythropoietin

seems to be EpoR mediated uptake and degradation.[30] As Widness *et al*[31] showed, bone marrow ablation after myoablative conditioning led to a decrease in EPO elimination. Similar observations were made in irradiated dogs after hypoxia-plasma,[32] and the opposite was seen in patients with hyperactive marrow due to haemolytic anaemia.[33] This mechanism in turn would indicate that elimination of EPO is related to its affinity to and residence time at the receptor.

Recombinant erythropoietin in disease

As EPO plays an important role in regulating erythropoiesis, a major step in medicine was taken when recombinant EPO was first produced by Lin *et al*[34] and Jacobs *et al*[35] in Chinese hamster ovary cells, later optimised for clinical use in patients with renal anaemia. Trials with the first recombinant human EPO (rHuEPO) showed a correction of anaemia in end-stage renal disease[36] and rHuEPO was approved by the FDA for human use in patients with chronic renal failure in 1989.[22] These first recombinant forms of EPO (called epoietin alfa, e.g. Eprex®) are identical to endogenous human EPO with regards to the amino acid backbone and four glycosylation sites, although some differences in molecular composition of the N-glycans have been found.[37] Half-lives are quite similar to endogenous EPO (4-9 hours),[38] which is also the case for second generation rHuEPO, epoietin beta (e.g. Neorecormon®).[39] The same holds true for a later generation of recombinant EPO produced in human cells, epoietin delta (Dynepo), recently taken off the market in 2009.[40] Other forms of EPO, darbepoietin-alfa (NESP/Aranesp) and Mircera (CERA) have a longer half-life due to differences in amino-acid sequence, hyperglycosylation (NESP; $t_{1/2}$ = 24-26 hours[41]) and incorporation of a large polymer chain (CERA; $t_{1/2}$ = 6 days[42]). All these forms of recombinant human erythropoietin can help patients with chronic renal failure (CRF) to overcome the insufficient production of EPO due to the damaged kidneys and maintain steady-state erythropoiesis.

...and in sport. But does it work?

The treatment immediately also got the attention of athletes. As rHuEPO increases red blood cell mass and exercise capacity in anaemic patients, it might have the same effect in the athlete's body, thereby enhancing performance. With this rationale athletes started using rHuEPO, and the use of rHuEPO was put on the International Olympic Committee's (IOC) list of prohibited substances already in 1990. Now the list has been expanded to all "Erythropoiesis-Stimulating Agents (ESAs) [e.g. erythropoietin (EPO), darbepoetin (dEPO), hypoxia-inducible factor (HIF) stabilisers, methoxy polyethylene glycol-epoietin beta (CERA), peginesatide (Hematide)]".[43] The World Anti-Doping Agency (WADA) defines blood doping as "... the misuse of certain techniques and/or substances to increase one's red blood cell mass, which allows the body to transport more oxygen to muscles and therefore increase stamina and performance."[43] But do rHuEPO and other ESAs actually increase red blood cell mass in world-class cyclists and does this result in increased stamina and performance? First we look into the factors that determine stamina and endurance performance, especially in elite cycling and then the effects of rHuEPO on these parameters are reviewed.

What is endurance performance?

Main determining factors

The main determinants of aerobic endurance performance according to a model by Pate and Kriska[44] are maximal oxygen uptake (VO_{2max}), work economy (C) and the lactate threshold (LT). These three factors are now generally accepted as key factors in endurance performance[45–47] and are supported by findings in different studies on VO_{2max},[48,49] C[47,50,51] and LT.[47,48,50,51] A fourth factor, the lactate turn point (LTP), has also gotten some attention.[52]

VO_{2MAX} is a prerequisite but not a sole determining factor

VO_{2max}, the maximal oxygen uptake, has traditionally been regarded as the most important measure in endurance performance. According to Fick's Law it is dependent on cardiac output and the arterio-venous oxygen difference. These in turn, are mainly dependent on total blood volume (BV), the main limiting factor of stroke volume, and total body haemoglobin. However, lung diffusing capacity, heart rate, distribution of the blood volume to working skeletal muscles and arterial O_2 extraction contribute to VO_{2max} as well, as reviewed by Joyner *et al*[45] and Bassett *et al*[53] and reported by other researchers.[54,55] Heinicke *et al*[54] demonstrated the relationship between VO_{2max} and BV and total body haemoglobin in endurance

disciplines. Training can improve many of the mentioned factors to increase VO_{2max}, such as increasing blood volume,[56] and indeed, VO_{2max} values of champion endurance athletes are 50-100% greater than those observed in normally active healthy young subjects.[45] That an increase in VO_{2max} has a great potential to increase endurance performance was already shown by Buick *et al*[57] and Brien *et al*.[58] After autologous red blood cell reinfusion elevating haemoglobin and haematocrit levels in well-trained runners, running performance was significantly increased. Ekblom *et al*[59] cites another article by Celsing *et al*[60] to show that a haemoglobin increase irrespective of baseline haemoglobin levels will increase maximal aerobic power and therefore performance. However, the last statement in this paper by Ekblom *et al* is at least as important, where the authors warn to extrapolate this finding to the physically fit athlete, as in these subjects other factors than haemoglobin and maximal aerobic power seem to play a role in performance. Later research emphasised this warning, as VO_{2max} was found not to be the only determinant of endurance performance and more emphasis has recently come to the other two factors described by Pate and Kriska. VO_{2max}, although a prerequisite to perform at a high level,[48] has a very limited predictive value for endurance performance within a group of high-performance athletes for example.[61–67] Also, it has been shown that although successful endurance athletes reached a high VO_{2max} after initial years of training, later on they stayed at a plateau in their VO_{2max} but despite that kept improving performance[47,68,69] (note that one of these reports[69] is about Armstrong). Research into training for endurance performance shows the same trend: moderately trained athletes are able to improve VO_{2max} (as well as LT and C) by interval and/or intensive training,[70,71] whereas these training regimens do not improve VO_{2max} in well-trained athletes, but mainly improve the economy and lactate threshold,[50,72] possibly caused by improving buffering capacity.[73]

It is more than the VO_{2MAX}

Correspondingly, it is not VO_{2max}, but power output at submaximal intensities such as the first (VT1) and second (VT2) ventilation threshold, or respiratory compensation point (RCP) that significantly differ between elite (i.e. amateur) and professional cyclists.[64,74] All these findings indicate that other factors than VO_{2max} play an important role in professional and world-class cyclists. For example, when a published model[75] predicting endurance performance is used to predict the 1-hour cycling world record as described by Padilla *et al*,[76] predictions are far from the observed results. Based on the VO_{2max} and body mass of the studied subject, Miguel Indurain, the distance covered in 1 hour would have been 43.645 km predicted by the model, whereas the actual world record was set at 53.040 km/h. Calculating back from this record, the model would predict an impossible VO_{2max} of 10.3 L/min, where ranges for world-class athletes are 5-6 L/min.[67,77,78] This and another model[79] both identify VO_{2max} as the most important determinant for endurance performance and describe the relationship as being proportionally curvilinear, meaning that the better the athlete is trained, a similar increase in VO_{2max} leads to a proportionally smaller increase in performance. This also demonstrates that in world-class athletes, an increase in VO_{2max} will have only limited effect on performance. As the model by Nevill *et al* is not accurate to predict 1-hour performance in world-class cyclists, this suggests other factors than VO_{2max} play important roles in endurance performance at this level of performance.

Lactate Threshold

Therefore we now first take a look at the importance of lactate threshold in endurance athletes. Lactate threshold (LT), similar to the first ventilatory threshold (VT1 or VT), is the intensity of work or VO_2 at which the blood lactate concentration gradually starts to increase.[80] Aerobic enzyme activity is a major determinant of this LT, reflected by a decline in activity during a period of detraining accompanying a reduction in LT.[81] Because LT reflects an onset of anaerobic metabolism and the coinciding metabolic alterations (see a review by Joyner *et al*[45] and Bassett *et al*[53] for more detail on the mechanisms), this in turn determines the fraction of maximal aerobic power that can be sustained for an extended period. Several studies show that the VO_2 at this lactate threshold is highly related to performance, more so than VO_{2max}.[45,47,63–65,67,68,82] Elite cyclist are reported to be able to reach lactate thresholds between 300 and 400 W,[63,77,83] or 70-85% VO_{2max} (3.5-4.7 L/min).[65,67] Lactate threshold therefore reflects a balance between the rate of lactate production in the muscles and the rate of lactate efflux to the blood and clearance from the blood. In this balance another independent factor appears to play a role in endurance performance; difference in performance (time to fatigue) in cyclists with similar VO_{2max} can be explained by $\%VO_{2max}$ at LT, but an additional increase in performance in some athletes seems to be related to a high muscle capillary

density.[45,65] A similar correlation between endurance performance and capillary density was found in another study by Coyle *et al*,[67] and Anderson *et al*[84] found that capillary density increases with training. This might indicate these athletes have a higher capacity to remove and recycle muscle fatiguing metabolites allowing muscles to better tolerate lactic acid production and anaerobic metabolism,[85] or maintain/elongate mean transit time of the blood to increase oxygen extraction.[86]

Lactate Turn Point

Furthermore, another factor related to lactate, although less frequently used to measure endurance performance, is the lactate turn point (LTP),[52] or similar measures called respiratory compensation point (RCP),[87] second ventilatory threshold (VT2) or the onset of blood lactate accumulation (OBLA).[88] These factors represent a level of high work intensity at which lactate levels show a sudden and sustained rise and hypocapnic hyperventilation occurs.[63,68] This value is notably high in professional cyclists and an important factor during extreme endurance events.[64,83] A relationship between RCP and endurance performance has been reported,[63,89] with world-class cyclists having values up to 430-505 W,[63,76,83,90] or 90% of VO_{2max}.

Economy

The last important factor contributing to endurance performance is assumed to be completely independent of the previously mentioned factors, and is called work economy or efficiency (C). It is referred to as the ratio between work output (speed, power) and oxygen cost. Running economy is commonly defined as the steady-rate VO_2 in millilitres per minute per kilogram at a standard velocity, cycling economy as the caloric expenditure at a given work rate. A number of physiological and biomechanical factors seem to influence C in trained or elite athletes. These include metabolic adaptations within the muscle such as increased mitochondria and oxidative enzymes, the ability of the muscles to store and release elastic energy by increasing the stiffness of the muscles, and more efficient mechanics leading to less energy wasted on braking forces and excessive vertical oscillation.[44] Work economy has been shown to be a discriminator of endurance performance independently of VO_{2max} in runners[48,68,91–93] and cyclists,[63,69,78] becoming more important than VO_{2max} once a certain level of fitness is reached.[63] A possible explanation for differences in economy is the composition of the working muscles, where higher economy implies an improved efficiency of ATP turnover within muscle fibres during contraction.[94] Different muscle fibre types have different efficiencies; type I fibres (slow twitch) are most efficient, then type IIa fibres are recruited and lastly type IIb fibres (fast twitch). Several observations indicate work economy and endurance performance are related to the percentage type I fibres.[67,94,95] Training can induce changes from type IIb to IIa, and type IIa to type I in animals,[96] and possibly in humans,[67,97] supporting the finding that training can improve work economy, C.

Other factors

Besides these main determinants several other factors were also reported to be influencing endurance performance. Heart rate for example, although values corresponding to physiological markers such as LT and VT2 remain stable,[68,83] shows a rightward shift in its relationship with running speed[68] with chronic endurance training. This could be related to enlargement of the heart volume due to endurance training,[98,99] increasing stroke volume and allowing a reduced heart rate for the same cardiac output. Breathing pattern is another factor influencing cycling performance, as professional cyclists have been reported to lack a tachypnoeic shift at high workloads, indicating a more efficient use of their respiratory muscles.[100] Also the quantity of muscle mass recruited for sustained power production can influence performance, as elite cyclists can use 20-25% more muscle mass in endurance tests, therefore reducing the stress and power production per fibre.[65,101] Additionally, peak power output has been shown to be a predictor of performance in a time trial[102] and power to weight ratios contribute to climbing performance in cycling.[103] Lastly, two world-class endurance performance athletes showed to have an extremely low peak blood lactate concentration, which might indicate a mechanism for their outstanding performances[68,69] (note that one of these reports is about Armstrong[69]).

In summary, endurance performance mainly depends on an athlete's VO_{2max}, LT, LTP and C; VO_{2max} and LT/LTP interact to determine how long a rate of aerobic and anaerobic metabolism can be sustained and economy then determines how much speed or power can be achieved at a given amount of energy consumption. However, the contribution of each of these factors differs between different levels of training. Moderately trained athletes can easily improve all factors, whereas

increasing performance in elite athletes mainly seems to be governed by changes in LT, LTP and C. Additional factors, including capillary density, heart rate and heart volume, muscle mass and breathing pattern can influence endurance performance.

METHODS

Studying the effects of rHuEPO on endurance performance

Search strategy

Several studies have addressed the effects of rHuEPO with regard to endurance performance in subjects other than patients. A literature search was conducted in PubMed to identify these papers, using combinations of the key words 'erythropoietin', 'athletic performance', 'physical endurance', 'doping in sports' and 'athletes' for the primary search. Literature references in key papers were examined manually to identify additional papers. We did not attempt to derive quantitative systematic conclusions from a meta-analysis; therefore, this could be termed a qualitative systematic review.

RESULTS

Study population mismatch with professional cyclists

There are quite some studies looking at the effects of rHuEPO with regards to endurance performance in subjects other than patients. Some studies included "(endurance trained) recreational athletes" or "well trained individuals", others "healthy normal subjects". This brings already the first problem when interpreting the observations and results in these studies. As no standard has been used to classify the cycling abilities of the subjects, such as proposed by Jeukendrup *et al*,[77] subjects vary in baseline endurance performance and fitness level within a study and between studies. The level of training of the used subjects is poorly reported, but when trying to use the methods for cycling classification from Jeukendrup *et al*[77] on the scarce information reported, based on maximal power output and VO_{2max} (absolute and per kg body weight) subjects would be placed either in untrained cyclists (or healthy recreationally active subjects)[104–110] or trained cyclists.[111–115] Based on the reported information subjects in one study could not be classified.[116] Although these maximal parameters are not optimal to distinguish between top-level cyclists (as discussed above), it is clear that the studied subjects are not at all at a competing level of cycling performance. Moreover, this points out a very problematic aspect, which is that the studies do not use well-trained cyclists, let alone (for obvious reasons) elite or world-class cyclists, who, according to Jeukendrup *et al* would have VO_{2max} values above 70 mL/min/kg (5 L/min) and power outputs above 5 W/kg (See also Figure 1 to compare the studied subjects with these reference values). The only study using subjects with power outputs above this threshold with ~5.7 W/kg is by Connes *et al*,[111] but on the other hand their VO_{2max} is only ~64 mL/min/kg. As has been described for endurance athletes earlier in this review, it could well be that the studied subjects did not reach a plateau in VO_{2max} yet, which will prove an important flaw for the interpretation of the results of the studies. It is also well known that cyclists classified as well-trained or higher have different physiological characteristics in the factors comprising endurance performance than trained or untrained cyclists.[45,77,117] Lucia *et al*[63] showed for example, that the VO_2 kinetics are very different even between well-trained cyclists and world-class cyclists. Additionally, this classification shows there are major discrepancies between the groups in training status, which makes comparison difficult. As is the case for all research, including research on performance enhancement, Hopkins *et al*[118] stated that "the results of a research study apply with reasonable certainty only to populations that have similar characteristics to the sample under study. Elite athletes almost certainly have genetic endowment, training history, and training programs that differ from those of sub-elite athletes. A treatment may therefore produce different effects on performance in these two groups. It follows that the subjects in a study have to be elite athletes for the results to apply convincingly to elite athletes." Therefore it cannot be assumed that effects found in these rHuEPO studies on healthy untrained or trained individuals automatically apply to well-trained, elite and world-class cyclists.

Recombinant human erythropoietin dosing

The doses of rHuEPO in all studies vary, but all are subcutaneous injections, most in a similar range of 150 IU/kg per week, see Table 1. Almost all studies used forms of rHuEPO with half-lives similar to endogenous EPO, namely Eprex®,[109–111,113,114,116]

Neorecormon®,[104–107] Recormon®[115] or it was not reported.[112] Only one study used rHuEPO with a longer half-life, NESP.[108] Another problem with evaluating the results of these studies is that only 8 studies[105,106,109–111,113,115,116] out of 13 were placebo controlled. As endurance performance can change significantly due to for example training, it is crucial to control for these effects with a placebo treated group. Moreover, unfortunately only 5[106,109,111,113,115] of these studies were reported to be double-blinded, controlling for any bias due to treatment which is of possible major influence on the exercise tests performed in the studies. As the study using NESP as rHuEPO treatment is not placebo controlled and does not measure any performance parameters during normoxia, it is difficult to draw conclusions about the effects of this form of rHuEPO on endurance performance. Moreover, the newest form of rHuEPO, CERA, to our knowledge has not been studied for effects on endurance performance in athletes yet at all.

Haematological effects of rHuEPO

Although doses differ somewhat across the studies, most studies report similar magnitude of effects in haematological parameters with these doses. There are reports that during rHuEPO administration reticulocyte numbers are increased twofold in the lower doses[109,110] to threefold in the higher doses,[114,116] and drop below baseline approximately 7-14 days after rHuEPO treatment is ceased.[109,110,114,116] EPO concentrations also drop below baseline after rHuEPO treatment is stopped.[109,116] Another effect that is seen in most studies is the rise in [Hb] and Hct. Increases between 4.6% and 17.4%, and 8.3% and 19% are reported for [Hb] and Hct respectively (Table 1), with no obvious differences between training statuses of the athletes. These levels are reported to return to baseline within one month after cessation of treatment.[109] An increase in haematocrit could lead to an increase in oxygen carrying capacity, however, does this enhance performance? Hct is not a good marker of performance, as endurance athletes usually have lower Hct values than untrained subjects due to plasma volume expansion.[119] Additionally, it is a very variable measure and affected by different circumstances.[120] On top of that, increases of Hct cause an increase in viscosity of the blood,[121,122] which might hamper performance due to reductions in blood flow and increased heart muscle work. Considering that during exercise a decrease in plasma volume increases Hct even more,[120] and dehydration, hyperthermia and altitude possibly exaggerate this effect in 3-week races, it is not obvious what the effects of this rise in Hct will have in professional cyclists. Even more so because the rHuEPO treatment not only increases haemoglobin concentration and Hct, but at the same time decreases plasma volume, thereby resulting in almost no effect on, or a slight decrease in, BV.[123] The use of rHuEPO therefore interestingly possibly counteracts the plasma volume expansion of endurance training.[56] Despite all these observations however, the combination of effects seems to increase the performance parameter VO_{2max}, at least in the studied subjects under laboratory conditions.

Effects on VO_{2MAX}

The most important question in this review is then whether these effects on haematological parameters translate into an effect on performance. The different parameters that determine endurance performance were discussed previously, but unfortunately most studies only examine one of these parameters, being VO_{2max}. In the reported studies this parameter is increased in the rHuEPO treated subjects, with a relatively constant value for all studies, independent of training status of the subjects, between 7% and 9.7% (Table 1). Absolute values of VO_{2max} and treatment effects can be seen in Figure 1. This increase in VO_{2max} has been reported to be accompanied by an increase in power output.[105,106,110,111,114] This, in turn, resulted in an increase in performance a time-to-exhaustion test of 22%[106] and 54.3%[105] in untrained subjects and a smaller increase of 9.4% (and 1.5% in placebo)[115] and 16.6%,[112] in trained subjects. Importantly this surrogate parameter is measured in a test lasting about 20 minutes and leading to exhaustion, quite different from the required ~5 hour performance in a cycling race.

Does it translate to cycling performance?

As mentioned earlier, VO_{2max} is poorly related to cycle performance[64,74] and Lucia *et al*[124] even questions whether VO_{2max} is the limiting factor for maximal endurance performance in some 50% of professional cyclists due to a lack of plateau in VO_{2max} during an exercise to exhaustion test. Additionally, time to exhaustion protocols like the ones used here are subject to high variability and therefore poorly reproducible,[125,126] whereas time trial protocols would give a better performance evaluation,[125] also eliminating the influence of wrongly extrapolating laboratory test setting results to race-events.[118] The use of rHuEPO in these subjects clearly has

an effect on VO_{2max}, which might improve performance at peak intensity during severe exercise, although evidence for this is rather "soft". Apart from the uncertainty whether these same effects can be observed in well-trained or elite cyclists, surprisingly little is known from these studies about effects on submaximal intensities. This might be of major importance when looking at the nature of cycling. Long exercise times during consecutive days with the finish line as a known endpoint (contrary to the "open end" of time to exhaustion tests) makes it crucial for cyclists to dose their power during a race. This combined with (team) tactics, the terrain and the effects of drag force make it that cyclists only work a small amount of time at their peak intensities, or even above intensities where lactate accumulation occurs (i.e. VT2/OBLA/RCP/LTP). Investigations in world-class cyclists show that during 3-week races the subjects' HR is above such an intensity (HR_{obla}) only 3.6% (119 sec) of the time climbing a "Hors Categorie" climb (hardest climb), even less so during first and second category climbs, 2.6% (45 sec) and 2.5% (22 sec) respectively.[90] Similar low percentages were reported by Lucia *et al*[127] for total race time with HR above the RCP (at 90% VO_{2max}) during the Tour de France or Vuelta a España, 2.7% (149 min) and 3.3% (166 min) respectively. For time trials a difference in time spent with a HR above OBLA was found between different type of time trials, with prologue, short, long and uphill time trials recording 59, 38, 3.5 and 0% for cyclists going all-out.[128] HR values corresponding to physiological markers of performance (e.g. LT, VT2) have been shown to be stable during the course of a training year of professional cyclists.[83]

Other endurance performance parameters unstudied

For the major part of a race, cyclists therefore exercise well below their VO_{2max} levels, but this parameter surprisingly still has gotten the most attention when looking at rHuEPO effects. Some of the studies that did look at other parameters observed no change in the VO_2-kinetics[105,106,110,129] or VO_2 at submaximal exercise,[112] despite the increased oxygen carrying capacity due to the increase in [Hb] and Hct. This would mean that the oxygen carrying capacity of the blood does not determine VO_2 kinetics, but that this is regulated and limited by factors in the muscles rather than oxygen supply. This would also indicate that there is no change in LT in these subjects resulting from rHuEPO treatment as shown by Wilkerson *et al*,[106] who found no effect on gas exchange threshold (GET), a measure closely related to LT, due to the rHuEPO treatment. However, other researchers did find an increase in VT of 14.3%,[114] although this trial was not placebo controlled, so training and placebo effects cannot be accounted for. Another group[111] using a placebo controlled blinded study also found an increase in VT of 12.6%. No conclusive evidence for effects of rHuEPO on LT/VT is therefore available, with evidence on another important lactate parameter, LTP/OBLA/VT2/RCP, completely absent. It is important to elucidate the effects of rHuEPO on these parameters, as performance in cycling is much better related to these factors.[64,74] Time trial world record performance (1-hour world record) for example, seems to be best correlated to and predicted by the speed or power output at OBLA.[76] Other groups also report that performance in longer time trials is highly correlated to power output at OBLA[130] or power output at LT,[131] or with VO_2 at VT1[132] or LT.[67] In >50km time trials during the Tour the France, performance was correlated with power output at VT1.[133] In these time trials VO_{2max} is not related to performance, which was only demonstrated in shorter time trials (20 min).[131] Lastly, also uphill cycling has been correlated best to power outputs at LT or OBLA.[130] This means that the most determining disciplines for the general classification in stage-races in professional cycling are correlated to submaximal exercise parameters.

In the reviewed rHuEPO studies the last important endurance performance factor, economy, was only measured by one group[105] and did not change after rHuEPO treatment. This would be expected from the non-haematological, biomechanical factors that determine economy as discussed previously. On the other hand, there is some evidence that prolonged exposure to rHuEPO in healthy subjects may induce changes in the human skeletal tissue shown by an increase in the relative amount of the slow myosin light chain (MLC) (Type I fibres) while decreasing the amount of fast MLC (Type II fibres), possibly leading to improved economy.[134] However, more evidence is needed here as well to draw conclusions about effects of rHuEPO on economy. Especially because Lance Armstrong, accused of having the biggest doping (e.g. rHuEPO) network in the history of sports, was reported to have, although in a questionable study, a high muscular efficiency partly contributing to his world-class performance.[69]

Some other parameters, such as blood lactate, end-exercise HR and HR kinetics were investigated and reported as not altered by rHuEPO treatment,[106] although other studies indicate a non-significant drop in blood lactate[110] and HR[110,111] or significant in heart rate,[114] although only at submaximal exercise.[112] A significant

drop in blood lactate at rest and 10 min into a time to exhaustion (TTE) test, but not at exhaustion[105] was seen. Blood volume was also not affected.[112,123] One blinded study looked at the effect of rHuEPO on perception on physical self and reported a positive effect on perceived physical condition and strength.[113] Although animals overexpressing EPO had 14% higher muscle volume and a 25% increase in muscle vascularisation, this did not translate to increased muscle force or stamina.[135] Moreover, in healthy males no effects of prolonged rHuEPO treatment on capillarization or muscle fibre hypertrophy were reported in a publication[136] from the same study performed by Thomsen *et al*.[105]

Alternative mechanisms by which EPO works?

It may be argued that focusing on direct endurance measures does not take into account possible mechanisms by which rHuEPO causes better recovery after exercise. rHuEPO may have anti-inflammatory effects and may mitigate ischemia-reperfusion related damage,[137–140] which could potentially improve recovery. It has been suggested that EPO and its receptor function as a paracrine/autocrine system to mediate the protection of tissues subjected to (metabolic) stress.[141] However, these effects have not been confirmed in properly designed clinical trials. In fact, most clinical trials focusing on the tissue protective effects of rHuEPO have shown adverse rather than beneficial effects. Serious untoward effects have also been shown in rHuEPO-treated patients with stroke, myocardial infarction, or acute kidney injury, and in surgical patients (as reviewed by Patel *et al*[142]). These findings appear to be compatible with a pro-coagulant state induced by rHuEPO and possibly also with an augmentation of acute inflammatory reactions by the drug.[143] The data therefore do not suggest substantial effects on recovery of muscle injury during exercise.

Thus, except on VO_{2max}, no coherent or reproducible findings have been reported for both erythropoietic and non-erythropoietic effects of rHuEPO, rendering the evidence too weak to support any conclusion about effects on performance in professional cyclists.

Lack of scientific evidence

Given that (I) most of the research with rHuEPO on endurance performance has focused on a parameter for maximal exercise, VO_{2max}, (II) the factors that make professional and world-class cyclists unique are not VO_{2max}, but LT, RCP and C, (III) endurance performance in professional cycling such as in time trials is best correlated with submaximal exercise factors (e.g. LT, VT1, OBLA, RCP), (IV) only small parts of professional cycling races are cycled at severe or maximal intensities (above OBLA/RCP) and (V) the characteristics of the study populations have been significantly different from that of the population that is suspected of using rHuEPO, it cannot be concluded that rHuEPO use in professional cyclists (or even elite cyclists) will enhance cycling performance.

A more scientific approach needed

Summarizing, the available literature lacks the appropriate information, validity and robustness to conclude that rHuEPO enhances world-class cycling performance. To be able to make such statements, more thorough research needs to be conducted looking at the effects of rHuEPO on submaximal performance parameters and the cycling economy, preferably in a population with cycling performance abilities as close as possible to those of professional cyclists and under conditions closely resembling racing conditions and the required performance duration. As long as the effect on endurance performance in professional cycling is not clear, putting the treatment on the prohibited list falsely implies a proven ergogenic effect, possibly stimulating its abuse,[144] although it should also be recognised that there is no convincing evidence that any drug works in this context.

Adverse effects of rHuEPO in athletes

Apart from creating a level playground for all athletes by banning and trying to prevent doping use, doping is also forbidden to protect the athletes from using possibly harmful substances. The presented rHuEPO studies in healthy or trained subjects do not focus on negative side-effects of the treatment however. What some of these studies did observe is a significant rise in systolic blood pressure (SBP) at submaximal exercise.[112] A second publication[129] from the study described by Thomsen *et al*[105] reported a rise in systolic, diastolic and mean blood pressure at rest or maximal exercise, as well as a rise in systolic and mean blood pressure at submaximal exercise. Although others do not observe a blood pressure rise at rest,[106,114,116] a rise in blood pressure (either at rest or (sub-)maximal exercise) could be a possible threatening side-effect of rHuEPO use in healthy athletes. However,

the numbers of subjects and treatment times in the presented studies are too small to detect any (rare) adverse events. To get information on this, larger studies, namely patient studies, must be consulted, although it must be kept in mind that results of these studies do not per se translate to well-trained athletes. One of such patient studies was prematurely discontinued due to increased incidence of thrombotic events in rHuEPO treated metastatic breast cancer patients.[145] Other trials and meta-analyses showed a similar trend in different groups of patients treated with rHuEPO compared to placebo.[146–148] It should be noted however that these studies used ~4 times higher doses of rHuEPO (usually in the range of 40.000 IU or 600 IU/kg per week) compared to the endurance performance studies in healthy subjects. The increased blood viscosity in treated anaemic patients,[122,149] the earlier described rise in blood pressure and enhanced coagulation,[150] endothelial activation and platelet reactivity[151] and inflammation[152] after rHuEPO treatment have been mentioned to be involved in these thrombotic events. On top of these rHuEPO effects, acute exercise also enhances coagulation,[153] although less pronounced in trained than in untrained subjects. And because stroke volume and blood volume are reduced in acute exercise, haematocrit is increased,[120] which is even more pronounced in dehydrated and hyperthermic exercise conditions.[154,155] This combination of factors might increase the risk of thrombotic events in endurance performance athletes using rHuEPO. One of such adverse events could manifest for example after enhanced Hct levels lower cerebral blood flow and therefore oxygen supply to the brain, which in turn might predispose to cerebral infarction.[156] These thrombotic risks are underlined by a case report by Lage *et al*[157] where a professional cyclist presented with cerebral sinus thrombosis, thereafter confessing to 3 months of 2000 IU rHuEPO use every two days, in combination with 15 days of growth hormone and continuous high doses of vitamin A and E. Additionally, high haematocrit values could cause heart failure, myocardial infarction, seizures and pulmonary embolism.[158,159] Another association that has been reported for this treatment, caused by the induced hypertension, is hypertensive posterior encephalopathy.[160] A complication of rHuEPO use in patients that is life-threatening, is the onset of red cell aplasia, a very rare side-effect mainly linked to anti-erythropoietin antibody formation due to Eprex® use.[161] The improper handling and storage in illicit use in sport might enhance the risks of this and other immunogenic complications.[162,163]

And lastly, rHuEPO use has also been connected to promoting tumour growth and angiogenesis in tumours, however Fandrey *et al*[164] report that there is no evidence for such an involvement yet.

Summarizing, as only case reports have been presented on negative effects of rHuEPO use in cyclists, most information available is from patient-studies. These studies indicate that rHuEPO has several cardio-vascular effects, raising the risk of thrombotic events, encephalopathy and possibly some other complications. These risks might be even higher in cycling taking into account the circumstances in this sport which could compound these risks. Also, the needed secrecy in sports might lead to bad handling and storage of the rHuEPO, possibly elevating the risks of side-effects such as red cell aplasia.

CONCLUSION

Cyclists and rHuEPO: a risky choice to what advantage?

As the case of the United States Anti-Doping Agency *versus* Armstrong proves again, recombinant Human EPO has been used by many professional (including champion) cyclists. Given that it increases Hct, it is thought to enhance performance in professional cycling and therefore has been put on the list of prohibited substances of the International Olympic Committee. As rHuEPO is on this list, cyclists caught were breaking the rules and should be punished for doing so. However, this review shows that only very weak scientific evidence exists about the effects of rHuEPO on cycling performance. Sport physicians and cyclists should be informed about the dangers of the use of such a substance, as already proposed by Kuipers about doping in general.[144] Neither scientific basis for performance-enhancing properties, nor possible harmful side-effects have been provided for athletes or trainees.

The situation for rHuEPO use in athletes is analogous to the many forms of non-evidence-based treatments that exist in medical practice and which, by common opinion, should be refuted or confirmed by good clinical trials with real-life end-points. A single well-controlled trial in athletes during real-life circumstances would give a better indication of the real advantages and risk factors of rHuEPO use, but it would be an oversimplification to suppose that this would eradicate its use, even if no benefit were to be seen with increased biomarkers of risk.

High-quality scientific evidence is always preferable to the current situation, in which athletes risk their career and health with irrational use of a substance. If the size of the athletic benefit could be shown to be large (which, on the basis of the evidence presented in this review, is unlikely), it would support the enormous and apparently largely ineffective efforts currently made to detect and prevent the use of rHuEPO. If the effect were to be small, these efforts could be directed elsewhere.

REFERENCES

1 United States Anti-Doping Agency. *Reasoned decision of the united states anti-doping agency on disqualification and ineligibility - united states anti-doping agency,* Claimant, v. Lance Armstrong, Respondent. 2012.

2 Sackmann E. *Biological Membranes Architecture and Function. Handbook of Biological Physics* 1995.

3 Ratcliffe P, Eckardt K, Bauer C. Hypoxia, erythropoietin gene expression, and erythropoiesis.In: *Handbook Of Physiology-Environmental Physiology* 1996: 1125-1153.

4 Carnot P, DeFlandre C. Sur l'activite hemopoietique de serum au cours de la regeneration du sang. *C R Acad Sci (Paris)* 1906.

5 Imai N, Kawamura A, Higuchi M, Oh-Eda M, Orita T, Kawaguchi T, Ochi N. Physicochemical and biological comparison of recombinant human erythropoietin with human urinary erythropoietin. *J Biochem.* 1990; 107: 352-359.

6 Recny MA, Scoble HA, Kim Y. Structural characterization of natural human urinary and recombinant DNA-derived erythropoietin. Identification of des-arginine 166 erythropoietin. *J Biol Chem.* 1987; 262: 17156-17163.

7 Broudy VC, Lin N, Brice M, Nakamoto B, Papayannopoulou T. Erythropoietin receptor characteristics on primary human erythroid cells. *Blood* 1991; 77: 2583-2590.

8 Koury MJ, Bondurant MC. Erythropoietin retards DNA breakdown and prevents programmed death in erythroid progenitor cells. *Science* 1990; 248: 378-381.

9 Gobert S, Duprez V, Lacombe C, Gisselbrecht S, Mayeux P. The signal transduction pathway of erythropoietin involves three forms of mitogen-activated protein (MAP) kinase in UT7 erythroleukemia cells. *Eur.J Biochem.* 1995; 234: 75-83.

10 Miura Y, Miura O, Ihle JN, Aoki N. Activation of the mitogen-activated protein kinase pathway by the erythropoietin receptor. *J Biol Chem.* 1994; 269: 29962-29969.

11 Elliott S, Pham E, Macdougall IC. Erythropoietins: a common mechanism of action. *Exp Hematol.* 2008; 36: 1573-1584.

12 Remy I, Wilson IA, Michnick SW. Erythropoietin receptor activation by a ligand-induced conformation change. *Science* 1999; 283: 990-993.

13 Jacobson LO, Goldwasser E, Fried W, Plzak L. Role of the kidney in erythropoiesis. 1957. *J Am Soc Nephrol.* 2000; 11: 589-590.

14 Kuratowska Z, Lewartowski B, Michalski E. Studies on the production of erythropoietin by isolated organs. *Blood* 1961.

15 Fisher JW, Koury S, Ducey T, Mendel S. Erythropoietin production by interstitial cells of hypoxic monkey kidneys. *Br J Haematol.* 1996; 95: 27-32.

16 Koury ST, Bondurant MC, Koury MJ, Semenza GL. Localization of cells producing erythropoietin in murine liver by in situ hybridization. *Blood* 1991; 77: 2497-2503.

17 Cotes PM, Dore CJ, Yin JA, Lewis SM, Messinezy M, Pearson TC, Reid C. Determination of serum immunoreactive erythropoietin in the investigation of erythrocytosis. *N.Engl.J Med* 1986; 315: 283-287.

18 Semenza GL. Regulation of mammalian O2 homeostasis by hypoxia-inducible factor 1. *Annu.Rev Cell Dev.Biol* 1999; 15: 551-578.

19 Wang GL, Semenza GL. General involvement of hypoxia-inducible factor 1 in transcriptional response to hypoxia. *Proc Natl.Acad Sci USA* 1993; 90: 4304-4308.

20 Haase VH. Hypoxic regulation of erythropoiesis and iron metabolism. *Am J Physiol.Renal Physiol.* 2010; 299: F1-13.

21 Bruegge K, Jelkmann W, Metzen E. Hydroxylation of hypoxia-inducible transcription factors and chemical compounds targeting the HIF-alpha hydroxylases. *Curr.Med Chem.* 2007; 14: 1853-1862.

22 Fisher JW. Erythropoietin: physiology and pharmacology update. *Exp Biol Med (Maywood.)* 2003; 228: 1-14.

23 Imagawa S, Nakano Y, Obara N, Suzuki N, Doi T, Kodama T, Nagasawa T, Yamamoto M. A GATA-specific inhibitor (K-7174) rescues anemia induced by IL-1beta, TNF-alpha, or L-NMMA. *FASEB J* 2003; 17: 1742-1744.

24 Koury ST, Koury MJ, Bondurant MC, Caro J, Graber SE. Quantitation of erythropoietin-producing cells in kidneys of mice by in situ hybridization: correlation with hematocrit, renal erythropoietin mRNA, and serum erythropoietin concentration. *Blood* 1989; 74: 645-651.

25 Koury MJ, Sawyer ST, Bondurant MC. Splenic erythroblasts in anemia-inducing Friend disease: a source of cells for studies of erythropoietin-mediated differentiation. *J Cell Physiol.* 1984; 121: 526-532.

26 Eckardt KU, Boutellier U, Kurtz A, Schopen M, Koller EA, Bauer C. Rate of erythropoietin formation in humans in response to acute hypobaric hypoxia. *J Appl.Physiol.* 1989; 66: 1785-1788.

27 Goldwasser E, Kung CK, Eliason J. On the mechanism of erythropoietin-induced differentiation. 13. The role of sialic acid in erythropoietin action. *J Biol Chem.* 1974; 249: 4202-4206.

28 Dinkelaar RB, Engels EY, Hart AA, Schoemaker LP, Bosch E, Chamuleau RA. Metabolic studies on erythropoietin (EP): II. The role of liver and kidney in the metabolism of Ep. *Exp Hematol.* 1981; 9: 796-803.

29 Steinberg SE, Garcia JF, Matzke GR, Mladenovic J. Erythropoietin kinetics in rats: generation and clearance. *Blood* 1986; 67: 646-649.

30 Nalbant D, Saleh M, Goldman FD, Widness JA, Veng-Pedersen P. Evidence of receptor-mediated elimination of erythropoietin by analysis of erythropoietin receptor mRNA expression in bone marrow and erythropoietin clearance during anemia. *J Pharmacol Exp Ther* 2010; 333: 528-532.

31 Widness JA, Schmidt RL, Hohl RJ, Goldman FD, Al-Huniti NH, Freise KJ, Veng-Pedersen P. Change in erythropoietin pharmacokinetics following hematopoietic transplantation. *Clin Pharmacol Ther* 2007; 81: 873-879.

32 Stohlman F, Jr. Observations on the physiology of erythropoietin and its role in the regulation of red cell production. *Ann.N.Y.Acad Sci* 1959; 77: 710-724.

33 Alexanian R. Erythropoietin excretion in bone marrow failure and hemolytic anemia. *J Lab.Clin Med* 1973; 82: 438-445.

34 Lin FK, Suggs S, Lin CH, Browne JK, Smalling R, Egrie JC, Chen KK, Fox GM, Martin F, Stabinsky Z, . Cloning and expression of the human erythropoietin gene. *Proc Natl.Acad Sci USA* 1985; 82: 7580-7584.

35 Jacobs K, Shoemaker C, Rudersdorf R, Neill SD, Kaufman RJ, Mufson A, Seehra J, Jones SS, Hewick R, Fritsch EF, . Isolation and characterization of genomic and cDNA clones of human erythropoietin. *Nature* 1985; 313: 806-810.

36 Eschbach JW, Egrie JC, Downing MR, Browne JK, Adamson JW. Correction of the anemia of end-stage renal disease with

recombinant human erythropoietin. Results of a combined phase I and II clinical trial. *N.Engl.J Med* 1987; 316: 73-78.
37 Sasaki H, Bothner B, Dell A, Fukuda M. Carbohydrate structure of erythropoietin expressed in Chinese hamster ovary cells by a human erythropoietin cDNA. *J Biol Chem*. 1987; 262: 12059-12076.
38 Egrie J, Eschbach JW, Adamson JW, McGuire T. Pharmacokinetics of recombinant human erythropoietin administered to hemodialysis (HD) patients. *Kidney International* 1988.
39 European Medicines Agency. NeoRecormon, INN-Epoetin beta. 1-4-2004.
40 European Medicines Agency. Dynepo is epoetin delta. 1-3-2002.
41 Elliott S, Lorenzini T, Asher S, Aoki K, Brankow D, Buck L, Busse L, Chang D, Fuller J, Grant J, Hernday N, Hokum M, Hu S, Knudten A, Levin N, Komorowski R, Martin F, Navarro R, Osslund T, Rogers G, Rogers N, Trail G, Egrie J. Enhancement of therapeutic protein in vivo activities through glycoengineering. *Nat Biotechnol*. 2003; 21: 414-421.
42 Macdougall IC. CERA (Continuous Erythropoietin Receptor Activator): a new erythropoiesis-stimulating agent for the treatment of anemia. *Curr.Hematol.Rep.* 2005; 4: 436-440.
43 World Anti-Doping Agency. 2012.
44 Pate RR, Kriska A. Physiological basis of the sex difference in cardiorespiratory endurance. *Sports.Med* 1984; 1: 87-98.
45 Joyner MJ, Coyle EF. Endurance exercise performance: the physiology of champions. *J Physiol*. 2008; 586: 35-44.
46 Coyle EF. Physiological determinants of endurance exercise performance. *J Sci Med Sport*. 1999; 2: 181-189.
47 Jones AM. A five year physiological case study of an Olympic runner. *Br J Sports.Med* 1998; 32: 39-43.
48 Pollock ML. Submaximal and maximal working capacity of elite distance runners. Part I: Cardiorespiratory aspects. *Ann.N.Y.Acad.Sci*. 1977; 301: 310-322.
49 Saltin B, Astrand PO. Maximal oxygen uptake in athletes. *J Appl.Physiol*. 1967; 23: 353-358.
50 Enoksen E, Shalfawi SA, Tonnessen E. The effect of high- vs. low-intensity training on aerobic capacity in well-trained male middle-distance runners. *J Strength.Cond.Res* 2011; 25: 812-818.
51 Helgerud J, Engen LC, Wisloff U, Hoff J. Aerobic endurance training improves soccer performance. *Med Sci.Sports.Exerc.* 2001; 33: 1925-1931.
52 Davis HA, Bassett J, Hughes P, Gass GC. Anaerobic threshold and lactate turnpoint. *Eur.J Appl.Physiol.Occup.Physiol*. 1983; 50: 383-392.
53 Bassett DR, Jr., Howley ET. Limiting factors for maximum oxygen uptake and determinants of endurance performance. *Med Sci Sports Exerc.* 2000; 32: 70-84.
54 Heinicke K, Wolfarth B, Winchenbach P, Biermann B, Schmid A, Huber G, Friedmann B, Schmidt W. Blood volume and hemoglobin mass in elite athletes of different disciplines. *Int J Sports Med* 2001; 22: 504-512.
55 Wagner PD. New ideas on limitations to VO2max. *Exerc.Sport.Sci Rev* 2000; 28: 10-14.
56 Convertino VA. Blood volume: its adaptation to endurance training. *Med Sci Sports Exerc.* 1991; 23: 1338-1348.
57 Buick FJ, Gledhill N, Froese AB, Spriet L, Meyers EC. Effect of induced erythrocythemia on aerobic work capacity. *J Appl.Physiol*. 1980; 48: 636-642.
58 Brien AJ, Simon TL. The effects of red blood cell infusion on 10-km race time. *JAMA* 1987; 257: 2761-2765.
59 Ekblom B. Blood doping and erythropoietin. The effects of variation in hemoglobin concentration and other related factors on physical performance. *Am J Sports.Med* 1996; 24: S40-S42.
60 Celsing F, Svedenhag J, Pihlstedt P, Ekblom B. Effects of anaemia and stepwise-induced polycythaemia on maximal aerobic power in individuals with high and low haemoglobin concentrations 1. *Acta Physiol Scand*. 1987; 129: 47-54.
61 Sjodin B, Svedenhag J. Applied physiology of marathon running. *Sports Med* 1985; 2: 83-99.
62 Londeree BR. The use of laboratory test results with long distance runners. *Sports Med* 1986; 3: 201-213.
63 Lucia A, Hoyos J, Santalla A, Perez M, Chicharro JL. Kinetics of VO(2) in professional cyclists. *Med Sci.Sports.Exerc.* 2002; 34: 320-325.
64 Lucia A, Pardo J, Durantez A, Hoyos J, Chicharro JL. Physiological differences between professional and elite road cyclists. *Int J Sports Med* 1998; 19: 342-348.
65 Coyle EF, Coggan AR, Hopper MK, Walters TJ. Determinants of endurance in well-trained cyclists. *J Appl.Physiol*. 1988; 64: 2622-2630.
66 Wilber RL, Zawadzki KM, Kearney JT, Shannon MP, Disalvo D. Physiological profiles of elite off-road and road cyclists. *Med Sci Sports Exerc.* 1997; 29: 1090-1094.
67 Coyle EF, Feltner ME, Kautz SA, Hamilton MT, Montain SJ, Baylor AM, Abraham LD, Petrek GW. Physiological and biomechanical factors associated with elite endurance cycling performance. *Med Sci Sports Exerc.* 1991; 23: 93-107.
68 Jones AM. The physiology of the Wolrd Record Holder for the Women's Marathon. *Int J Sports Science & Coaching* 2006; 1: 101-116.
69 Coyle EF. Improved muscular efficiency displayed as Tour de France champion matures. *J Appl.Physiol*. 2005; 98: 2191-2196.
70 Helgerud J, Hoydal K, Wang E, Karlsen T, Berg P, Bjerkaas M, Simonsen T, Helgesen C, Hjorth N, Bach R, Hoff J. Aerobic high-intensity intervals improve VO2max more than moderate training. *Med Sci.Sports.Exerc.* 2007; 39: 665-671.
71 Esfarjani F, Laursen PB. Manipulating high-intensity interval training: effects on VO2max, the lactate threshold and 3000 m running performance in moderately trained males. *J Sci.Med Sport*. 2007; 10: 27-35.
72 Billat VL, Flechet B, Petit B, Muriaux G, Koralsztein JP. Interval training at VO2max: effects on aerobic performance and overtraining markers. *Med Sci.Sports.Exerc.* 1999; 31: 156-163.
73 Weston AR, Myburgh KH, Lindsay FH, Dennis SC, Noakes TD, Hawley JA. Skeletal muscle buffering capacity and endurance performance after high-intensity interval training by well-trained cyclists. *Eur.J Appl.Physiol.Occup.Physiol*. 1997; 75: 7-13.
74 Chicharro JL, Hoyos J, Lucia A. Effects of endurance training on the isocapnic buffering and hypocapnic hyperventilation phases in professional cyclists. *Br J Sports Med* 2000; 34: 450-455.
75 Nevill AM, Jobson SA, Palmer GS, Olds TS. Scaling maximal oxygen uptake to predict cycling time-trial performance in the field: a non-linear approach. *Eur.J Appl.Physiol*. 2005; 94: 705-710.
76 Padilla S, Mujika I, Angulo F, Goiriena JJ. Scientific approach to the 1-h cycling world record: a case study. *J Appl.Physiol*. 2000; 89: 1522-1527.
77 Jeukendrup AE, Craig NP, Hawley JA. The bioenergetics of World Class Cycling. *J Sci Med Sport*. 2000; 3: 414-433
78 Lucia A, Hoyos J, Perez M, Santalla A, Chicharro JL. Inverse relationship between VO2max and economy/efficiency in world-class cyclists. *Med Sci.Sports Exerc.* 2002; 34: 2079-2084.
79 Ingham SA, Whyte GP, Pedlar C, Bailey DM, Dunman N, Nevill AM. Determinants of 800-m and 1500-m running performance using allometric models. *Med Sci.Sports.Exerc.* 2008; 40: 345-350.
80 di Prampero PE, Atchou G, Bruckner JC, Moia C. The energetics of endurance running. *Eur.J Appl.Physiol.Occup.Physiol*. 1986; 55: 259-266.
81 Coyle EF, Martin WH, III, Bloomfield SA, Lowry OH, Holloszy JO. Effects of detraining on responses to submaximal exercise. *J Appl.Physiol*. 1985; 59: 853-859.
82 Kumagai S, Tanaka K, Matsuura Y, Matsuzaka A, Hirakoba K, Asano K. Relationships of the anaerobic threshold with the 5 km, 10 km, and 10 mile races. *Eur.J Appl.Physiol.Occup.Physiol*. 1982; 49: 13-23.
83 Lucia A, Hoyos J, Perez M, Chicharro JL. Heart rate and performance parameters in elite cyclists: a longitudinal study. *Med Sci Sports Exerc.* 2000; 32: 1777-1782.
84 Andersen P, Henriksson J. Capillary supply of the quadriceps femoris muscle of man: adaptive response to exercise. *J Physiol*. 1977; 270: 677-690.
85 Tesch PA, Wright JE. Recovery from short term intense exercise: its relation to capillary supply and blood lactate concentration. *Eur.J Appl.Physiol.Occup.Physiol*. 1983; 52: 98-103.
86 Saltin B. Hemodynamic adaptations to exercise. *Am J Cardiol*. 1985; 55: 42D-47D.
87 Beaver WL, Wasserman K, Whipp BJ. A new method for detecting anaerobic threshold by gas exchange. *J Appl.Physiol*. 1986; 60: 2020-2027.
88 Sjodin B, Jacobs I. Onset of blood lactate accumulation and marathon running performance. *Int J Sports Med* 1981; 2: 23-26.
89 Lucia A, Hoyos J, Chicharro JL. Physiology of professional road cycling. *Sports Med* 2001; 31: 325-337.
90 Padilla S, Mujika I, Santisteban J, Impellizzeri FM, Goiriena JJ. Exercise intensity and load during uphill cycling in professional 3-week races. *Eur.J Appl.Physiol*. 2008; 102: 431-438.
91 Conley DL, Krahenbuhl GS. Running economy and distance running performance of highly trained athletes. *Med Sci Sports Exerc.* 1980; 12: 357-360.
92 Saltin B, Kim CK, Terrados N, Larsen H, Svedenhag J, Rolf CJ. Morphology, enzyme activities and buffer capacity in leg muscles of Kenyan and Scandinavian runners. *Scand.J Med Sci Sports* 1995; 5: 222-230.
93 Bassett DR, Jr., Howley ET. Maximal oxygen uptake: 'classical' versus 'contemporary' viewpoints. *Med Sci Sports Exerc.* 1997; 29: 591-603.
94 Coyle EF, Sidossis LS, Horowitz JF, Beltz JD. Cycling efficiency is related to the percentage of type I muscle fibers. *Med Sci Sports Exerc.* 1992; 24: 782-788.
95 Horowitz JF, Sidossis LS, Coyle EF. High efficiency of type I muscle fibers improves performance. *Int J Sports Med* 1994; 15: 152-157.
96 Green HJ, Klug GA, Reichmann H, Seedorf U, Wiehrer W, Pette D. Exercise-induced fibre type transitions with regard to myosin, parvalbumin, and sarcoplasmic reticulum in muscles of the rat. *Pflugers.Arch*. 1984; 400: 432-438.
97 Tesch PA, Karlsson J. Muscle fiber types and size in trained and untrained muscles of elite athletes. *J Appl.Physiol*. 1985; 59: 1716-1720.
98 Longhurst JC, Kelly AR, Gonyea WJ, Mitchell JH. Cardiovascular responses to static exercise in distance runners and weight lifters. *J Appl.Physiol*. 1980; 49: 676-683.
99 Roskamm H. Optimum patterns of exercise for healthy adults. *Can.Med Assoc.J* 1967; 96: 895-900.
100 Lucia A, Carvajal A, Calderon FJ, Alfonso A, Chicharro JL. Breathing pattern in highly competitive cyclists during incremental exercise. *Eur.J Appl.Physiol.Occup.Physiol*. 1999; 79: 512-521.
101 Coyle EF. Integration of the physiological factors determining endurance performance ability. *Exerc.Sport.Sci Rev* 1995; 23: 25-63.
102 Hawley JA, Noakes TD. Peak power output predicts maximal oxygen uptake and performance time in trained cyclists. *Eur.J Appl.Physiol.Occup.Physiol*. 1992; 65: 79-83.
103 Swain DP. The influence of body mass in endurance bicycling. *Med Sci Sports Exerc.* 1994; 26: 58-63.
104 Lundby C, Achman-Andersen NJ, Thomsen JJ, Norgaard AM, Robach P. Testing for recombinant human erythropoietin in urine: problems associated with current anti-doping testing. *J Appl.Physiol*. 2008; 105: 417-419.
105 Thomsen JJ, Rentsch RL, Robach P, Calbet JA, Boushel R, Rasmussen P, Juel C, Lundby C. Prolonged administration of recombinant human erythropoietin increases submaximal performance more than maximal aerobic capacity. *Eur.J Appl.Physiol*. 2007; 101: 481-486.
106 Wilkerson DP, Rittweger J, Berger NJ, Naish PF, Jones AM. Influence of recombinant human erythropoietin treatment on pulmonary O2 uptake kinetics during exercise in humans. *J Physiol*. 2005; 568: 639-652.
107 Rasmussen P, Foged EM, Krogh-Madsen R, Nielsen J, Nielsen TR, Olsen NV, Petersen NC, Sorensen TA, Secher NH, Lundby C. Effects of erythropoietin administration on cerebral metabolism and exercise capacity in men. *J Appl.Physiol*. 2010; 109: 476-483.
108 Lundby C, Damsgaard R. Exercise performance in hypoxia after novel erythropoiesis stimulating protein treatment. *Scand.J Med Sci Sports* 2006; 16: 35-40.
109 Parisotto R, Gore CJ, Emslie KR, Ashenden MJ, Brugnara C, Howe C, Martin DT, Trout GJ, Hahn AG. A novel method utilising markers of altered erythropoiesis for the detection of recombinant human erythropoietin abuse in athletes. *Haematologica*. 2000; 85: 564-572.
110 Russell G, Gore CJ, Ashenden MJ, Parisotto R, Hahn AG. Effects of prolonged low doses of recombinant human erythropoietin during submaximal and maximal exercise 1. *Eur.J Appl.Physiol*. 2002; 86: 442-449.
111 Connes P, Perrey S, Varray A, Prefaut C, Caillaud C. Faster oxygen uptake kinetics at the onset of submaximal cycling exercise following 4 weeks recombinant human erythropoietin (r-HuEPO) treatment. *Pflugers.Arch*. 2003; 447: 231-238.
112 Ekblom B, Berglund B. Effect of erythropoietin administration on maximal aerobic power. *Scand J Med Sci Sports* 1991; 88-93.
113 Ninot G, Connes P, Caillaud C. Effects of recombinant human erythropoietin injections on physical self in endurance athletes 1. *J Sports.Sci*. 2006; 24: 383-391.

114 Audran M, Gareau R, Matecki S, Durand F, Chenard C, Sicart MT, Marion B, Bressolle F. Effects of erythropoietin administration in training athletes and possible indirect detection in doping control. *Med Sci.Sports.Exerc.* 1999; 31: 639-645.

115 Birkeland KI, Stray-Gundersen J, Hemmersbach P, Hallen J, Haug E, Bahr R. Effect of rHEPO administration on serum levels of sTfR and cycling performance. *Med Sci.Sports.Exerc.* 2000; 32: 1238-1243.

116 Souillard A, Audran M, Bressolle F, Gareau R, Duvallet A, Chanal JL. Pharmacokinetics and pharmacodynamics of recombinant human erythropoietin in athletes. Blood sampling and doping control. *Br J Clin Pharmacol* 1996; 42: 355-364.

117 Faria EW, Parker DL, Faria IE. The science of cycling: physiology and training - part 1. *Sports Med* 2005; 35: 285-312.

118 Hopkins WG, Hawley JA, Burke LM. Design and analysis of research on sport performance enhancement. *Med Sci Sports Exerc.* 1999; 31: 472-485.

119 Eichner ER. Sports anemia, iron supplements, and blood doping. *Med Sci Sports Exerc.* 1992; 24: S315-S318.

120 Schmidt W, Biermann B, Winchenbach P, Lison S, Boning D. How valid is the determination of hematocrit values to detect blood manipulations? *Int.J Sports.Med* 2000; 21: 133-138.

121 Pirofsky B. The determination of blood viscosity in man by a method based on Poiseuille's law. *J Clin Invest.* 1953; 32: 292-298.

122 Shand BI, Buttimore AL, Lynn KL, Bailey RR, Robson RA. Effect of hemodialysis and recombinant human erythropoietin on determinants of blood viscosity. *Ren.Fail.* 1994; 16: 407-413.

123 Lundby C, Thomsen JJ, Boushel R, Koskolou M, Warberg J, Calbet JA, Robach P. Erythropoietin treatment elevates haemoglobin concentration by increasing red cell volume and depressing plasma volume. *J Physiol.* 2007; 578: 309-314.

124 Lucia A, Rabadan M, Hoyos J, Hernandez-Capilla M, Perez M, San Juan AF, Earnest CP, Chicharro JL. Frequency of the VO2max plateau phenomenon in world-class cyclists. *Int J Sports Med* 2006; 27: 984-992.

125 Jeukendrup A, Saris WH, Brouns F, Kester AD. A new validated endurance performance test. *Med Sci Sports Exerc.* 1996; 28: 266-270.

126 Krebs P.S., Powers S.K. Reliability of laboratory endurance tests. *Med.Sci.Sports Exerc.* 1989.

127 Lucia A, Hoyos J, Santalla A, Earnest C, Chicharro JL. Tour de France versus Vuelta a Espana: which is harder? *Med Sci Sports Exerc.* 2003; 35: 872-878.

128 Padilla S, Mujika I, Orbananos J, Angulo F. Exercise intensity during competition time trials in professional road cycling. *Med Sci Sports Exerc.* 2000; 32: 850-856.

129 Lundby C, Robach P, Boushel R, Thomsen JJ, Rasmussen P, Koskolou M, Calbet JA. Does recombinant human Epo increase exercise capacity by means other than augmenting oxygen transport? *J Appl.Physiol.* 2008; 105: 581-587.

130 Padilla S, Mujika I, Cuesta G, Goiriena JJ. Level ground and uphill cycling ability in professional road cycling. *Med Sci Sports Exerc.* 1999; 31: 878-885.

131 Bentley DJ, McNaughton LR, Thompson D, Vleck VE, Batterham AM. Peak power output, the lactate threshold, and time trial performance in cyclists. *Med Sci Sports Exerc.* 2001; 33: 2077-2081.

132 Hoogeveen AR, Hoogsteen GS. The ventilatory threshold, heart rate, and endurance performance: relationships in elite cyclists. *Int J Sports Med* 1999; 20: 114-117.

133 Lucia A, Hoyos J, Perez M, Santalla A, Earnest CP, Chicharro JL. Which laboratory variable is related with time trial performance time in the Tour de France? *Br J Sports Med* 2004; 38: 636-640.

134 Christensen B, Lundby C, Jessen N, Nielsen TS, Vestergaard PF, Moller N, Pilegaard H, Pedersen SB, Kopchick JJ, Jorgensen JO. Evaluation of functional erythropoietin receptor status in skeletal muscle in vivo: acute and prolonged studies in healthy human subjects. *PLoS One* 2012; 7: e31857.

135 Hojman P, Brolin C, Gissel H, Brandt C, Zerahn B, Pedersen BK, Gehl J. Erythropoietin over-expression protects against diet-induced obesity in mice through increased fat oxidation in muscles. *PLoS One* 2009; 4: e5894.

136 Lundby C, Hellsten Y, Jensen MB, Munch AS, Pilegaard H. Erythropoietin receptor in human skeletal muscle and the effects of acute and long-term injections with recombinant human erythropoietin on the skeletal muscle. *J Appl.Physiol.* 2008; 104: 1154-1160.

137 Brines M, Cerami A. Discovering erythropoietin's extra-hematopoietic functions: biology and clinical promise. *Kidney Int* 2006; 70: 246-250.

138 Brines M. The therapeutic potential of erythropoiesis-stimulating agents for tissue protection: a tale of two receptors. *Blood Purif.* 2010; 29: 86-92.

139 Rathod DB, Salahudeen AK. Nonerythropoietic properties of erythropoietin: implication for tissue protection. *J Investig.Med* 2011; 59: 1083-1085.

140 Vogel J, Gassmann M. Erythropoietic and non-erythropoietic functions of erythropoietin in mouse models. *J Physiol.* 2011; 589: 1259-1264.

141 Erbayraktar S, Yilmaz O, Gokmen N, Brines M. Erythropoietin is a multifunctional tissue-protective cytokine. *Curr Hematol.Rep.* 2003; 2: 465-470.

142 Patel NS, Collino M, Yaqoob MM, Thiemermann C. Erythropoietin in the intensive care unit: beyond treatment of anemia. *Ann. Intensive.Care* 2011; 1: 40.

143 Hojman P, Taudorf S, Lundby C, Pedersen BK. Erythropoietin augments the cytokine response to acute endotoxin-induced inflammation in humans. *Cytokine* 2009; 45: 154-157.

144 Kuipers H. Doping in sport--exercise scientists have to take responsibility. *Int.J Sports Med* 2001; 22: 545.

145 Rosenzweig MQ, Bender CM, Lucke JP, Yasko JM, Brufsky AM. The decision to prematurely terminate a trial of R-HuEPO due to thrombotic events. *J Pain Symptom Manage.* 2004; 27: 185-190.

146 Corwin HL, Gettinger A, Fabian TC, May A, Pearl RG, Heard S, An R, Bowers PJ, Burton P, Klausner MA, Corwin MJ. Efficacy and safety of epoetin alfa in critically ill patients. *N.Engl.J Med* 2007; 357: 965-976.

147 Bohlius J, Wilson J, Seidenfeld J, Piper M, Schwarzer G, Sandercock J, Trelle S, Weingart O, Bayliss S, Djulbegovic B, Bennett CL, Langensiepen S, Hyde C, Engert A. Recombinant human erythropoietins and cancer patients: updated meta-analysis of 57 studies including 9353 patients. *J Natl.Cancer Inst.* 2006; 98: 708-714.

148 Stowell CP, Jones SC, Enny C, Langholff W, Leitz G. An open-label, randomized, parallel-group study of perioperative epoetin alfa versus standard of care for blood conservation in major elective spinal surgery: safety analysis. *Spine (Phila Pa 1976.)* 2009; 34: 2479-2485.

149 Linde T, Sandhagen B, Danielson BG, Wikstrom B. Impaired erythrocyte fluidity during treatment of renal anaemia with erythropoietin. *J Intern Med* 1992; 231: 601-606.

150 Clyne N, Berglund B, Egberg N. Treatment with recombinant human erythropoietin induces a moderate rise in hematocrit and thrombin antithrombin in healthy subjects. *Thromb.Res.* 1995; 79: 125-129.

151 Stohlawetz PJ, Dzirlo L, Hergovich N, Lackner E, Mensik C, Eichler HG, Kabrna E, Geissler K, Jilma B. Effects of erythropoietin on platelet reactivity and thrombopoiesis in humans. *Blood* 2000; 95: 2983-2989.

152 Tobu M, Iqbal O, Fareed D, Chatha M, Hoppensteadt D, Bansal V, Fareed J. Erythropoietin-induced thrombosis as a result of increased inflammation and thrombin activatable fibrinolytic inhibitor. *Clin.Appl.Thromb.Hemost.* 2004; 10: 225-232.

153 Sinzinger H, Virgolini I. Effects of exercise on parameters of blood coagulation, platelet function and the prostaglandin system. *Sports Med* 1988; 6: 238-245.

154 Gonzalez-Alonso J, Mora-Rodriguez R, Below PR, Coyle EF. Dehydration markedly impairs cardiovascular function in hyperthermic endurance athletes during exercise. *J Appl. Physiol.* 1997; 82: 1229-1236.

155 Gonzalez-Alonso J, Mora-Rodriguez R, Coyle EF. Stroke volume during exercise: interaction of environment and hydration. *Am J Physiol.Heart Circ.Physiol.* 2000; 278: H321-H330.

156 Wade JP. Transport of oxygen to the brain in patients with elevated haematocrit values before and after venesection. *Brain* 1983; 106 (Pt 2): 513-523.

157 Lage JM, Panizo C, Masdeu J, Rocha E. Cyclist's doping associated with cerebral sinus thrombosis. *Neurology* 2002; 58: 665.

158 Jelkmann W. Use of recombinant human erythropoietin as an antianemic and performance enhancing drug. *Curr.Pharm. Biotechnol.* 2000; 1: 11-31.

159 Burch GE, DePasquale NP. Hematocrit, blood viscosity and myocardial infarction. *Am J Med* 1962.

160 Delanty N, Vaughan C, Frucht S, Stubgen P. Erythropoietin-associated hypertensive posterior leukoencephalopathy. *Neurology* 1997; 49: 686-689.

161 Cazzola M. Erythropoietin therapy: need for rationality and active surveillance. *Haematologica* 2003; 88: 601-605.

162 Casadevall N, Cournoyer D, Marsh J, Messner H, Pallister C, Parker-Williams J, Rossert J. Recommendations on haematological criteria for the diagnosis of epoetin-induced pure red cell aplasia. *Eur.J Haematol.* 2004; 73: 389-396.

163 Smalling R, Foote M, Molineux G, Swanson SJ, Elliott S. Drug-induced and antibody-mediated pure red cell aplasia: a review of literature and current knowledge. *Biotechnol.Annu. Rev* 2004; 10: 237-250.

164 Fandrey J, Dicato M. Examining the involvement of erythropoiesis-stimulating agents in tumor proliferation (erythropoietin receptors, receptor binding, signal transduction), angiogenesis, and venous thromboembolic events. *Oncologist.* 2009; 14 Suppl 1: 34-42.

FIGURE 1 MAXIMAL OXYGEN UPTAKE BEFORE AND AFTER TREATMENT WITH rHuEPO IN THE DIFFERENT STUDIES PER TREATED GROUP Maximal oxygen uptake (VO_{2max}) before (grey squares) and after (black triangles) treatment with rHuEPO in the different studies per treated group (bars representing SD). N is the number of subjects in each group, with an asterisk indicating that the article reported VO_{2max} values only in litres per minute, which has been converted to millilitres per kilogram per minute by dividing this value by mean weight of the group for comparison purposes (no SD is given for these studies because of this conversion). Studies above the horizontal dotted line were performed using subjects classified as untrained, while below the horizontal line the subjects were classified as trained cyclists. Vertical dashed lines represent minimal values of VO_{2max} for different classifications of cyclists as suggested by Jeukendrup et al.[77] (dashed line, trained; dotted line, well trained; dotdashed line, elite; and longdashed line, world class).

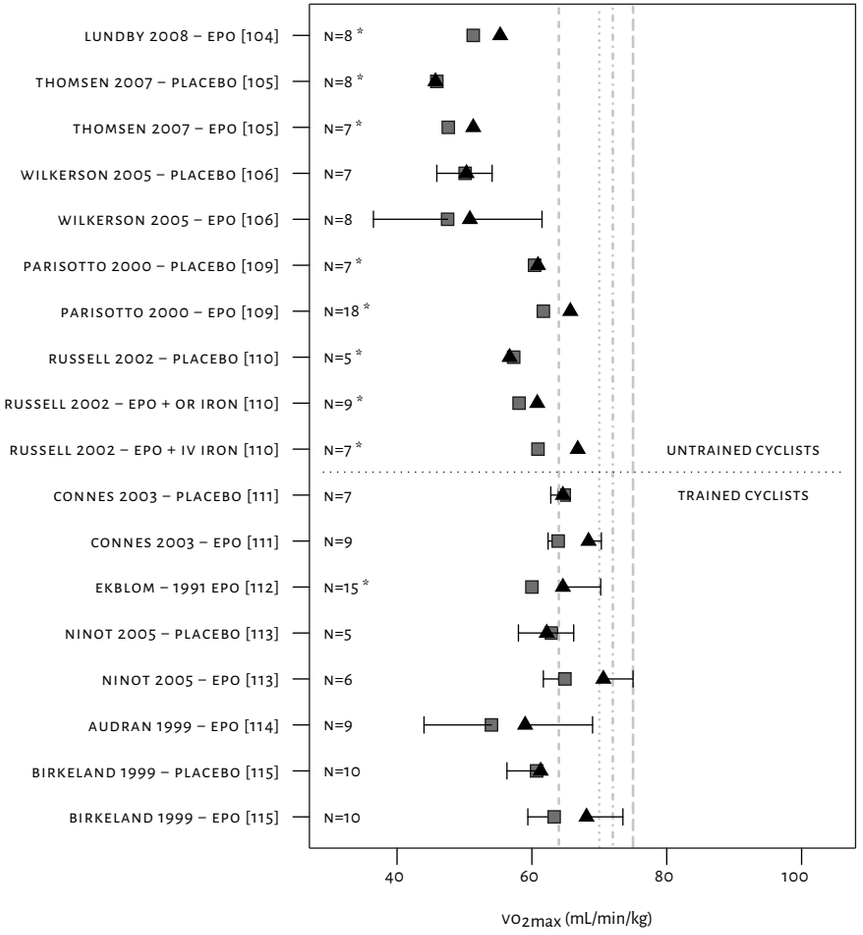


TABLE 1 OVERVIEW OF CHARACTERISTICS AND OUTCOMES OF THE STUDIES INVESTIGATING THE EFFECTS OF RECOMBINANT HUMAN ERYTHROPOIETIN ON ENDURANCE PERFORMANCE IN SUBJECTS OTHER THAN PATIENTS All effects are calculated based on the greatest difference found in the parameter when multiple measuring time points were reported. Abbreviations are as follows: Hb, haemoglobin; Hct, haematocrit; NA, not applicable; NESP, novel erythropoiesis stimulating protein (darbepoietin-alfa); VO_{2max}, maximal oxygen uptake.

Study	Type of subjects	Study set-up	Product	Dosing	Max. Hb increase (%)	Max. Hct increase (%)	Max. VO_{2max} increase (%)
Lundby *et al* 2008[104]	Untrained	Uncontrolled	Neorecormon®	5000 IU (~65 IU/kg) on alternating days for 14 days followed by once a week for 2 weeks	10.2	11.2	7.9
Thomsen *et al* 2007[105]	Untrained	Placebo	Neorecormon®	5000 IU (~60 IU/kg) on alternating days for 2 weeks, a dose on 3 consecutive days for one week and one dose a week for 12 weeks	11.1	10.7	9.1
Wilkerson *et al* 2005[106]	Untrained	Placebo + Blinded	Neorecormon®	150 IU/kg once a week for 4 weeks	7	12	7
Rasmussen *et al* 2010[107]	Untrained	Uncontrolled	Neorecormon®	5000 IU (~60 IU/kg) on alternating days for 2 weeks, a dose on 3 consecutive days for one week and one dose a week for 12 weeks	-	12	-
Lundby *et al* 2006[108]	Untrained	Uncontrolled	NESP	144 IU/kg (0.72 µg/kg) once a week for 4 weeks	17.4	16.4	-
Parisotto *et al* 2000[109]	Untrained	Placebo + Blinded	Eprex®	50 IU/kg three times a week over 25 days (in combination with ~100 mg iron either IM / OR)	7.4 / 12	-	6.3 / 6.9
Russell *et al* 2002[110]	Untrained	Placebo	Eprex®	50 IU/kg three times a week for 3 weeks and 20 IU/kg for 5 weeks	-	15	9.7
Connes *et al* 2003[111]	Trained	Placebo + Blinded	Eprex®	50 IU/kg three times a week for 4 weeks	9.6	8.3	7
Ekblom *et al* 1991[112]	Trained	Uncontrolled	NA	20 IU/kg three times a week for 6 weeks (or 4 weeks, and 40 IU/kg for the remaining 2 weeks)	-	11.7	8
Ninot *et al* 2006[113]	Trained	Placebo + Blinded	Eprex®	50 IU/kg three times a week for 4 weeks, followed by 20 IU/kg three times a week for 2 weeks	9.5	10.2	7
Audran *et al* 1999[114]	Trained	Uncontrolled	Eprex®	50 IU/kg daily for 26 days	9.3	11.5	9.3
Birkeland *et al* 2000[115]	Trained	Placebo + Blinded	Recormon®	5000 IU (181-232 IU/kg/week) three times a week	11.2	19	7
Souillard *et al* 1996[116]	Unknown	Placebo	Eprex®	200 IU/kg 5 times in 11 days	4.6	8.9	-

4
EFFECTS OF ERYTHROPOIETIN ON CYCLING PERFORMANCE OF WELL TRAINED CYCLISTS: A DOUBLE-BLIND, RANDOMISED, PLACEBO-CONTROLLED TRIAL

LANCET HAEMATOLOGY 2017;4(8):E374–86

Jules AAC Heuberger, Joris I Rotmans, Pim Gal, Frederik E Stuurman, Juliëtte van 't Westende, Titiaan E Post, Johannes MA Daniels, Matthijs Moerland, Peter LJ van Veldhoven, Marieke L de Kam, Herman Ram, Olivier de Hon, Jelle J Posthuma, Jacobus Burggraaf and Adam F Cohen

ABSTRACT

Substances that potentially enhance performance (e.g., recombinant human erythropoietin [rHuEPO]) are considered doping and are therefore forbidden in sports; however, the scientific evidence behind doping is frequently weak. We aimed to determine the effects of rHuEPO treatment in well trained cyclists on maximal, submaximal, and race performance and on safety, and to present a model clinical study for doping research on other substances. We did this double-blind, randomised, placebo-controlled trial at the Centre for Human Drug Research in Leiden (Netherlands). We enrolled healthy, well trained but non-professional male cyclists aged 18–50 years and randomly allocated (1:1) them to receive abdominal subcutaneous injections of rHuEPO (epoetin-beta; mean dose 6000 IU per week) or placebo (0.9% NaCl) for 8 weeks. Randomisation was stratified by age groups (18–34 years and 35–50 years), with a code generated by a statistician who was not masked to the study. The primary outcome was exercise performance, measured as maximal power output (P_{max}), maximal oxygen consumption VO_{2max}, and gross efficiency in maximal exercise tests with 25 W increments per 5 min, as lactate threshold and ventilatory threshold 1 (VT1) and 2 (VT2) at submaximal levels during the maximal exercise test, and as mean power, VO_2, and heart rate in the submaximal exercise tests at the highest mean power output for 45 min in a laboratory setting and in a race to the Mont Ventoux (France) summit, using intention-to-treat analyses. The trial is registered with the Dutch Trial Registry (Nederlands Trial Register), number NTR5643. Between March 7, 2016, and April 13, 2016, we randomly assigned 48 participants to the rHuEPO group (n=24) or the placebo group (n=24). Mean haemoglobin concentration (9.6 mmol/L versus 9.0 mmol/L [estimated difference 0.6, 95% CI 0.4 to 0.8]) and maximal power output (351.55 W versus 341.23 W [10.32, 3.47 to 17.17]), and VO_{2max} (60.121 mL/min per kg versus 57.415 mL/min per kg [2.707, 0.911 to 4.503]) in a maximal exercise test were higher in the rHuEPO group compared with the placebo group. Submaximal exercise test parameters mean power output (283.18 W versus 277.28 W [5.90, −0.87 to 12.67]) and VO_2 (50.288 mL/min per kg versus 49.642 mL/min per kg [0.646, −1.307 to 2.600]) at day 46, and Mont Ventoux race times (1 h 40 min 32 s versus 1 h 40 min 15 s [0.3%, −8.3 to 9.6]) did not differ between groups. All adverse events were grade 1–2 and were similar between both groups. No events of grade 3 or worse were observed. Although rHuEPO treatment improved a laboratory test of maximal exercise, the more clinically relevant submaximal exercise test performance and road race performance were not affected. This study shows that clinical studies with doping substances can be done adequately and safely and are relevant in determining effects of alleged performance-enhancing drugs.

INTRODUCTION

Use of drugs that potentially enhance performance (also called doping) is a major problem in many competitive sports, partly shown by the thousands of annual adverse analytical findings.[1] The 2017 prohibited list of drugs is substantial (>300 substances) and open-ended because all compounds that potentially enhance performance could be forbidden.[2] The list is not necessarily based on solid evidence, as shown by the criteria for including substances and methods section on the Prohibited List of the World Anti-Doping Agency (WADA) Code, which states as one of the criteria that there only needs to be "experience" that a substance "has the potential to enhance performance".[3] This criterium is probably driven by the assumption that scientific evidence cannot be obtained in many cases or generation of such evidence is too time-consuming and expensive. Moreover, in the time required to collect the evidence, the substance could be used, leading to unfair situations. Therefore, it is not surprising that the scientific evidence supporting the ban on substances to be used by athletes is scarce. A publicly known example of such a banned drug is recombinant human erythropoietin (rHuEPO), which has been under constant scrutiny by anti-doping authorities since its first alleged use in the late 1980s. Although relatively few athletes have been caught for rHuEPO abuse during their active careers, the attention for this banned substance has spiked recently because many professional cyclists competing in the 1990s and 2000s admitted to having used rHuEPO to improve their cycling performance.

rHuEPO induces erythropoiesis and thereby enhances blood haemoglobin concentrations, and it was assumed that this induction would result in increased muscle delivery of oxygen and hence improved exercise performance. However, the evidence for the performance-enhancing effects of rHuEPO in high-level competitive sports is rather scarce. The evidence constitutes of small, often uncontrolled studies,[4] in arguably unrepresentative populations and is often inappropriately expressed only in exercise parameters that mainly evaluate maximal exercise performance. These tests are often of incremental intensity or at a very high intensity, and therefore lead to exhaustion, usually within 20 min. By contrast, submaximal tests are at intensity levels that can be maintained for a long period of time (>20 min), which is the level at which cyclists perform most of the time. Therefore, both types of tests evaluate different types of performance parameters. Well powered studies on the effects of rHuEPO on submaximal exercise parameters in trained athletes are lacking. Additionally, studies[5] have reported that an increase in haematocrit and a subsequent increased blood viscosity is associated with a marked reduction in muscle oxygen delivery. Furthermore, elite athletes improve their work economy or submaximal performance, not their maximal oxygen consumption, when improving their performance over time, indicating that maximal oxygen consumption might not be a rate-limiting factor.[6] Finally, whether increasing haemoglobin beyond normal values is beneficial is unclear; data in patients with anaemia suggest that the goal of rHuEPO treatment should not be normalising haemoglobin concentrations because this results in an increased incidence of ischaemic stroke.[7] Furthermore, possible sudden deaths of professional cyclists related to rHuEPO were suggested in the late 1980s and early 1990s.

We decided to study rHuEPO as a model doping drug. The aim of this study was to evaluate the effect of rHuEPO in well trained cyclists on maximal and submaximal performance parameters in a laboratory setting and in a real-life road cycle race. Additionally, we evaluated if this trial design would be a practical approach to investigate other doping substances.

METHODS

Study design and participants

We designed a double-blind, randomised, placebo- controlled study of healthy male cyclists between ages 18 years and 50 years. We undertook the study at the Centre for Human Drug Research (CHDR) in Leiden (the Netherlands). Participants were recruited via advertisements, social media, newsletters of cycling clubs, and through the help of national sports associations. Main inclusion criteria were being fluent in Dutch and having a maximum power- to-weight ratio during the maximal exercise test at screening that exceeded 4 W/kg, normal exercise electrocardiogram (ECG), screening haemoglobin between 8.0 mmol/L and 9.8 mmol/L (equivalent to 12.8–15.7 g/dL), screening haematocrit below 48% and not being subject to anti-doping regulation or using medication that could potentially interact with the study drugs or study assessments. After passing a preliminary screening over the telephone, participants underwent a medical screening, followed by a maximal

exercise test to determine peak exercise performance. From the first dose to 3 months after the final dose, participants were not allowed to take part in sports events that were subject to anti-doping regulations. All participants gave written informed consent before any study-related activity.

The study was approved by the Independent Ethics Committee of the Foundation Evaluation of Ethics in Biomedical Research (Stichting Beoordeling Ethiek Biomedisch Onderzoek, Assen, Netherlands). The study is registered in the Dutch Trial Registry (Nederlands Trial Register), number NTR5643. Our study protocol is available online.

Randomisation and masking

Participants were randomly assigned (1:1) to either the rHuEPO group or the placebo group. To reduce potential variability between the groups due to age differences, a stratified randomisation was used with one block of participants aged 18–34 (inclusive) and another of participants aged 35–50 (inclusive). The randomisation code was generated by a statistician who was not masked to the study and was not involved in the execution of the study. Until study closure the treatment codes were only available to this statistician and the Leiden University Medical Centre (LUMC) pharmacy, that distributed the study agents. Participant enrollment was done by a physician who was masked to the study.

Procedures

Participants received weekly abdominal subcutaneous injections of epoetin-beta (NeoRecormon, Roche, Basel, Switzerland) or saline (0.9% NaCl) for 8 weeks. Target haemoglobin in the rHuEPO group was a 10–15% increase compared with the baseline haemoglobin concentration. Haemoglobin was measured with the HemoCue Hb 201+ analyser (Radiometer Benelux BV, Zoetermeer, Netherlands) and haematocrit with the Haematokrit 200 centrifuge (Hettich Benelux BV, Geldermalsen, Netherlands) before each dose administration and measurements were only available to personnel who were not masked to the study. All haematology samples were collected after participants were seated with their feet on the floor for at least 10 min. All participants in the rHuEPO group received 5000 IU per injection for the first four rHuEPO injections. If the haemoglobin concentration was below the target range, a physician who was not masked or related to the study modulated the dose to 6000 IU, 8000 IU, or 10 000 IU in the subsequent 4 weeks to reach the target range. When haemoglobin was in the target range during the treatment period, rHuEPO dose was adjusted to 2000 IU. For safety reasons, a placebo injection was administered if the haemoglobin concentration exceeded the upper limit of the haemoglobin range or if the haematocrit concentration was equal to or exceeded 52% (dose decision tree, Figure 1). Doses were given in maximum 1 mL injections and distributed over two syringes if this volume was exceeded.

The different doses were prepared by a technician not masked to the study from multidose vials containing a lyophilisate of 50 000 IU epoetin-beta and 10 mL solvent for solution for injection. rHuEPO and placebo were visually indistinguishable (both colourless solutions) and dose changes (changes in injected volume) were also randomly assigned to placebo participants by the statistician and pharmacy before the start of the study.

During the treatment period, all participants also received open-label daily oral doses of 200 mg ferrous fumarate (Pharmachemie BV, Haarlem, Netherlands) and 50 mg ascorbic acid (Pharmachemie BV), and received standard instructions about concomitant food intake. Intake of these supplements was recorded daily by the participant in a diary.

Participants were instructed to maintain their usual training programme throughout the study. The racing bikes of participants were equipped with a Single Leg Power Meter SGY-PM910H2 (Pioneer Europe, Antwerpen, Belgium) with Shimano Ultegra 6800 crank (Shimano, Osaka, Japan) to log training data on the bicycle during the entire study. Data of bicycle trainings were uploaded to the dedicated database Cyclo-Sphere. Additionally, participants recorded all exercise activity in a diary, including other sports or cycling done without the power meter.

Exercise tests were done on a Monark LC4r ergometer (COSMED, Rome, Italy). Gas exchange was measured by Quark CPET system (COSMED), with breath-by-breath sampling technology and integrated heart rate measurement, with one of two wireless heart rate straps (COSMED and Polar, Kempele, Finland). Data were collected on a dedicated computer with the Omnia Metabolic Modules software (COSMED). Before each test, the gas analysers and flow meter were calibrated. Maximal exercise tests were done during screening, at baseline (up to 14 days before

first dose), and during the treatment period at days 11, 25, 39, and 53 (participants could deviate by 1 day) after the first dose administration. Submaximal exercise tests were done at baseline (up to 14 days before first dose and at least three days after the baseline maximal test) and at 46 days (participants could deviate by 1 day) after the first dose administration.

For both the maximal and submaximal exercise tests, the start of the protocol dictated 1 min rest without pedalling, followed by a 2 min warm up at a pedalling workload of 75 W. In the maximal exercise protocol, a ramp test was done where the pedalling workload was increased after the warm up to 175 W, and increased by an additional 25 W every 5 min. Cadence had to be maintained between 70 rpm and 90 rpm. Exhaustion was reached when cadence could not be maintained above 70 rpm or when a participant terminated the test. Subsequently, a 3-min recovery with a pedalling workload of 50 W was initiated. Between 4 min 15 s and 4 min 45 s into each step and immediately after termination of the exercise test (at peak pedalling workload), blood was collected from an intravenous cannula in the right forearm to measure blood lactate concentrations with a Lactate Pro 2 meter (Arkray, Kyoto, Japan). The screening maximal exercise test was similar, except that lactate was not measured (blood was not collected) and an exercise ECG was monitored and recorded with a 12-lead ECG system (COSMED or Labtech Ltd, Debrecen, Hungary).

In the submaximal exercise protocol, pedalling workload was set at 80% of the maximal power reached during the baseline maximal exercise test. Participants were instructed to produce the highest mean power output during a 45-min period, attempting to mimic competitive cycling time trials. Participants could adjust the power on the bike by indicating with hand gestures to increase or decrease in power by steps of 10 W. Cadence had to be maintained between 70 rpm and 90 rpm and the test was stopped after 45 min, followed by 3 min recovery at 50 W. Blood was collected from an intravenous cannula at 10 min, 30 min, and 45 min to measure blood lactate concentrations.

Approximately 12 days (range 10–16) after the last dose participants competitively climbed Mont Ventoux (Vaucluse département, France) in an open course via Bédoin (France), bridging an altitude of 1610 m over 21.5 km, resulting in an average gradient of 7.5%. The race was preceded by a stage of 110 km in Provence (France; total elevation gain 1524 m) that was completed collectively (i.e., all participants finished the course to the foot of the Mont Ventoux in a closed pack). Before the 110 km stage and at the top of Mont Ventoux, blood was collected and all participants, including their bicycles, were weighed. After the race, participants were asked by personnel (masked to the study) whether they thought they had been treated with rHuEPO or placebo during the treatment period.

Vital signs were measured regularly and adverse events documented during every visit. Additionally, before and regularly during the treatment period blood samples were taken in which haematology, coagulation, and endothelial function markers were measured. A broad range of markers was measured to evaluate potential risk of rHuEPO treatment in well trained cyclists.

All data were stored in a clinical trial database (Promasys, Omnicomm Inc, Fort Lauderdale, FL, USA) and checked for accuracy and completeness. A masked data review was done before code-breaking and analysis, according to a standard procedure at our unit.

Outcomes

The primary outcome was exercise performance, under both maximal and submaximal conditions, as assessed with multiple measures. For the maximal exercise test, the primary outcome was measured as maximal power output (P_{max}), maximal oxygen consumption VO_{2max}, and gross efficiency. P_{max} was calculated with the following formula:

$$P_{max} = [\text{power of the last completed step}] + \left[\frac{\text{time (s) in the subsequent step} \times 25\text{W}}{300\text{ s}}\right]$$

The breath-by-breath dataset was averaged in epochs of 30 s and the VO_{2max} was determined. Gross efficiency was calculated by the following formulas:

$$\text{Gross efficiency} = \frac{\text{power}}{\text{energy expenditure}} \times 100$$

where energy expenditure was calculated at the last level when the respiratory quotient was less than 1.0 and the power step had lasted longer than 180 s using the formula:

$$\text{Energy expenditure} = [((3.869 \times VO_2) + (1.195 \times VCO_2)) \times \frac{4.186}{60}]$$

For the submaximal levels during the maximal exercise test, the primary outcomes were measured as the lactate threshold, determined with the modified Dmax method, and ventilatory threshold 1 (VT1) and 2 (VT2),[8] for which assessments were done by two masked staff members (JH and PG), in consensus.

We obtained secondary outcomes from the maximal exercise test (e.g., heart rate, respiratory min volume [VE], volume of CO_2 expired [VCO_2]) from the last completed 30 s average before the recovery phase. We calculated the mean power, VO_2, and heart rate primary outcomes during the 45-min submaximal exercise test on the basis of the breath-by-breath dataset. We calculated cycling economy, a primary outcome, as described previously using the following formula:

$$\text{Cycling economy} = \frac{\text{mean power}}{\text{mean } VO_2\text{(L/min)}}$$

Secondary outcomes were lactate concentrations measured at 10 min, 30 min, and 45 min. In the Mont Ventoux race, secondary outcomes were race time, average efficiency, and average power. Secondary safety outcomes were blood pressure, heart rate, adverse events, and coagulation and endothelial function markers. Additional measurements of skin blood flow and diagnostic aspects of detection of rHuEPO use were done and these will be published separately.

Statistical analysis

No studies have been published that allowed a formal power calculation based on enhancement of submaximal performance by rHuEPO in trained cyclists. Therefore, we based the power calculation on the increase in VO_{2max} observed in a previous study[9] with moderately trained participants (3.8 mL/min per kg). We assumed the effect in well trained participants would be smaller; therefore, the power calculation was done on a VO_{2max} increase of 1.7 mL/min per kg. To detect a difference of 1.7 mL/min per kg with a power of 80%, a sample size of 22 was needed, assuming that the common SD is 1.95, using a two-tailed t test with a 0.05 two-sided significance level. When taking into account a 10% attrition rate, 24 participants were required in both groups.

The mean power output per kg of 11 male professional cyclists during a 20 min constant-load test at 80% VO_{2max}, which was determined in a maximal exercise test with a 25 W/min ramp protocol, was 5.2 W/kg (SD 0.2).[10] With a sample size of 22 per treatment group, a difference of 0.172 W/kg could be detected in this population with a power of 80%. This difference would mean that a professional cyclist weighing 75 kg would go from an average of 390 W at 80% VO_{2max} to 402.9 W. With available calculators this power difference, for an athlete on a 9 kg racing bike, in racing position ('drops') at 25°C on a wind-still, flat terrain of 40 km, would produce a speed increase of approximately 0.5 km/h (from 43.80 km/h to 44.32 km/h), which is a relevant difference in cycling. On a mountain climb, that same increase in power would lead to an even larger relative increase in speed, expanding the effect on uphill race time - e.g., a decrease of approximately 2 min on a climb like Mont Ventoux.

To evaluate effects on performance for each variable we selected a suitable statistical model and undertook intention-to-treat analysis. Participants were included in all analyses of outcomes for which they had at least one measurement. Repeatedly measured data were analysed with a mixed model analysis of variance with treatment, time, and treatment by time as fixed factors, participants as a random factor and, if available, the (average) prevalue as covariate. These included parameters of maximal exercise test, haematology, coagulation, endothelial function, and vital signs measurements. Different timepoints for each variable are indicated in the tables.

We compared single measured data with an analysis of variance with factor treatment, and, if available, prevalue as covariate. We analysed parameters of the submaximal exercise test in this way.

We analysed the racing times with a parametric model for failure time (accelerated failure time regression model) with right censored values, to account for participants who did not reach the top of Mont Ventoux. The model for the response variable consists of a linear effect composed of the covariates and a random disturbance term. The covariates are treatment and prevalue. We log-transformed the time to arrival before analysis and the chosen distribution was normal. We chose the P_{max}/kg pretreatment as the prevalue to correct for possible differences in baseline performance.

The contrast that we calculated within the models was placebo versus rHuEPO. We report results of statistical models as estimated means at the different timepoints per group and estimates of the difference over the whole time period, including 95% CI (% for log-transformed parameters) and the p value of the contrasts.

We used a regression model to evaluate the association between haematological parameters and performance. We analysed the association between haemoglobin and haematocrit concentrations and several maximal and submaximal exercise variables with a mixed model regression, with treatment as covariate, a random participant intercept and slope, and an unstructured variance and covariance structure if feasible; a variance components variance and covariance structure otherwise. For the submaximal exercise test variables, there was only one after baseline measurement. We calculated the regression of time in race and haemoglobin and haematocrit concentrations at the time of the race with a regression model without random factors.

To evaluate the association between maximal, submaximal exercise tests and race performance, we calculated Spearman correlations for the maximal and submaximal exercise variables per kg as measured in the test closest to the race and power per kg and time in the race.

When 95% CIs are presented they reflect the estimated difference between the two treatment groups. Significance level was set at $p<0.050$. We did all calculations with SAS version 9.4.

RESULTS

Between March 7, 2016, and April 13, 2016, we enrolled 48 participants and had one reserve participant. The study took place for all participants simultaneously between April and June, 2016, with a follow-up before the end of August, 2016. One participant withdrew after the first dose administration and was replaced by the reserve participant, and another participant withdrew after the fourth dose administration. Both withdrawals were due to personal reasons and not related to the study treatment or medical concerns. In total, 48 participants were included in the analyses, with 24 in the rHuEPO group and 24 in the placebo group (Figure 2). Baseline characteristics were similar between treatment groups (Table 1). Table 2 has a more detailed breakdown of excluded participants before the screening visit. As no effect of the stratification variable age was observed on the results, all analyses were done disregarding this factor. All participants were living at sea level and did not spend any substantial amount of time at (simulated) high altitude. Furthermore, average cumulative cycle training duration per week recorded with the Pioneer equipment (4.9 h for rHuEPO versus 5.9 h for placebo, [estimated difference –16.5%, 95% CI of estimated difference –36.3 to 9.5]) and average training distance and power per week (186.2 km versus 202.0 km [–15.8, –63.5 to 31.8] and 202.1 W versus 205.3 W [–3.2, –25.5 to 19.2]) did not differ among treatment groups, nor did other training activities recorded in the diary (average 1.2 h [SD 0.9] per week for rHuEPO and 1.5 h [1.2] per week for placebo).

Participants in the rHuEPO group received eight doses during the study. Mean rHuEPO dose was 5000 IU per participant per week during the first 4 weeks of the study and 7000 IU in the subsequent 4 weeks. On five occasions a placebo injection was administered to participants exceeding 15% of haemoglobin increase compared with baseline or that had a haematocrit concentration that exceeded 52% (Table 3). The average administered rHuEPO dose was 48 000 IU (6000 IU/week), resulting in an average 12% increase in haemoglobin concentration up to a mean of 10.2 mmol/L and a 16% increase in haematocrit to 50%, whereas haemoglobin and haematocrit concentrations in the placebo group remained relatively stable during the study (Table 4). Diaries showed that participants had taken their supplements as instructed throughout the study period.

Haemoglobin concentrations and haematocrit were higher in the rHuEPO group compared with the placebo group over the treatment period (9.6 mmol/L versus 9.0 mmol/L [estimated difference 0.6, 95% CI 0.4–0.8] and 47.6% versus 44.3% [3.3, 2.5–4.1], respectively; Figure 3, Table 4). In the rHuEPO group, median haemoglobin concentration at baseline was 9.0 mmol/L (range 8.1–10.2) and median peak haemoglobin concentration was 10.1 mmol/L (range 9.0–11.5). The mean increase in haemoglobin concentration at the prerace measurement was 12%; in total, 14 (61%) of 23 participants completing treatment with rHuEPO achieved an increase of more than 10%. Participants who did not achieve the target haemoglobin range all received 5000 IU of rHuEPO in the first four doses of the study and a mean of 8000 IU in the latter four doses of the study.

Analysis of the effects on maximal exercise test variables showed that at baseline the mean maximal power output per kg was 4.37 W/kg (SD 0.365) in the

rHuEPO group compared with 4.36 W/kg (0.223) in the placebo group and mean absolute maximal power output was similar in both groups (335.14 W [34.46] in the rHuEPO group and 335.00 W [33.04] for the placebo group; Table 5). During the course of the study, maximal power output per kg did not differ between the rHuEPO group compared with the placebo group (4.61 W/kg versus 4.50 W/kg [estimated difference 0.11, 95% CI −0.00 to 0.22]; Table 5). However, absolute maximal power output did show a significant increase in the rHuEPO group compared with the placebo group over the entire treatment period (351.55 W versus 341.23 W [10.32, 3.47 to 17.17]; Table 5), which reached significance at the exercise test at 25 days (p=0.0073; data not shown). A similar significant increase in the rHuEPO group was seen in VO_{2max}, VO_2 at VT1 and VT2, and power at VT1 (Figure 4). rHuEPO treatment increased VO_{2max} by 10% compared with baseline, and the placebo group also improved by 4%. This results in a net improvement of about 5% over placebo. Similar effects were found on maximal power output, with an increase of about 4% for rHuEPO treatment compared with placebo. There was no indication that rHuEPO treatment had a stronger effect on maximal power output in the highest performing participants (Figure 5). Gross efficiency, lactate threshold, maximal heart rate, or any of the other respiratory parameters did not differ between groups (Table 5).

Analysis of the effects on the submaximal exercise test showed that mean power output during the study did not differ between groups either in absolute terms (283.18 W for rHuEPO versus 277.28 W for placebo [estimated difference 5.90, 95% CI −0.87 to 12.67]) or per kg (3.72 W/kg for rHuEPO versus 3.66 W/kg for placebo [0.06,−0.04 to 0.16]; Table 6). Mean VO_2 per kg (50.288 mL/min per kg for rHuEPO versus 49.642 mL/min per kg for placebo [0.646, −1.307 to 2.600]), cycling economy, mean heart rate, and lactate levels at 10 min, 30 min, and 45 min were similar between treatment groups (Table 6).

A total of 44 participants took part in the Mont Ventoux race, with 21 (48%) in the rHuEPO group. Two participants were unable to attend due to other engagements, and one participant was experiencing gastrointestinal complaints. Weather conditions on Mont Ventoux (afternoon of June 19, 2016) in Bedoin were around 20°C and 40 km/h northern wind, and at the top were around 5°C and 85 km/h northern wind, without precipitation. Out of the 44 participants, four (9%) did not complete the race due to exhaustion (n=2 placebo and n=2 rHuEPO). The mean time of the Mont Ventoux race did not differ among treatment groups (1 h 40 min 32 s for the rHuEPO group versus 1 h 40 min 15 s for the placebo group [estimated difference 0.3%, 95% CI −8.3 to 9.6]), nor did mean pedalling power during the race (3.03 W/kg for rHuEPO versus 3.09 W/kg for placebo [−1.7%, −11.0 to 8.6]; Table 7).

To evaluate whether participants could notice the effects of rHuEPO, all 47 participants that completed the study were asked whether they thought they had received rHuEPO or placebo during the study period. Overall, 27 (57%) of 47 participants correctly indicated their treatment. Out of the participants treated with rHuEPO, only nine (39%) of 23 thought they had received rHuEPO. Six (25%) of 24 participants treated with placebo thought this as well.

Evaluation of the association between haematological parameters and performance revealed a relation between haemoglobin and all maximal and submaximal exercise parameters, including maximal power output per kg, which showed a significant estimated slope in the rHuEPO group (slope 0.27 [95% CI 0.16–0.37]), but not the placebo group (Figure 6). Haemoglobin or haematocrit concentrations were not associated with Mont Ventoux race time.

The correlation between exercise test parameters and race performance was highest for mean power output per kg during the submaximal exercise test, but only partly predicted race time (correlation of −0.63), while power output per kg measured during the race appeared to be more predictive of race time (−0.79). The strongest correlation between a maximal exercise test parameter and average power output during the submaximal exercise test was with maximal power output (0.76). However, parameters from the maximal exercise test a week before the Mont Ventoux race and the uphill race times only showed a moderate correlation at best (spearman correlations range −0.36 to −0.58).

Safety evaluation of the rHuEPO treatment in the well trained cyclists revealed that weight and vital signs, such as heart rate and blood pressure, were similar between treatment groups (Table 8, Table 9). All observed adverse events were mild to moderate (grade 1–2), and the nature and incidence were similar in both groups (Table 10). No events of grade 3 or worse were observed. Of all coagulation and endothelial function markers measured, when compared with the placebo group, rHuEPO increased only E-selectin by 8.6% (95% CI 2.0–15.7) and P-selectin by 7.8% (1.5–14.5; Table 11).

DISCUSSION

The present study of rHuEPO treatment in well trained cyclists compared with those treated with a placebo showed that rHuEPO improved exercise performance during a maximal exercise test. By contrast, no improvement in exercise performance was observed in the submaximal test, nor in a real-life road cycle race. These outcomes indicate that the amount of performance increment in a maximal test might not be immediately translated to a real-life situation. Although no difference in adverse events was observed, the endothelial function markers E-selectin and P-selectin significantly increased in the rHuEPO group compared with the placebo group, potentially increasing the risk of thrombosis. The question remains whether these results can be generalised to the actual population of athletes doping with rHuEPO.

Average weekly training duration in this study was substantially less compared with professional cyclists (who train >20 h and >700 km per week), which was inevitable because participants were not allowed to be subject to anti-doping regulation and were therefore essentially amateur cyclists. Baseline values of maximal power output for our maximal exercise test were on average 335.07 W (SD 33.40) and of VO_{2max} were 55.63 mL/min per kg (SD 4.80) for both treatment groups. Elite cyclists have reported values of about 429 W and 73 mL/min per kg measured with a short exercise protocol, which would indicate that our participants were not comparable to this level of cyclist either.[6] However, because it takes 3–4 min for the body to acclimatise to a given workload and for lactate to be measured accurately,[11] and because longer protocols can be more sensitive to performance changes,[12] we selected a long exercise protocol with 5 min per power step rather than the 1–2 min protocols often used in exercise physiology. Shorter protocols overestimate maximal exercise parameters, including maximal power output up to 20% and VO_{2max} up to 7% compared with longer protocols.[8,13] Although our participants were not professional cyclists, based on the maximal exercise test they have similarity with the elite cyclists tested in a 3-min exercise protocol (similar values for maximal power output of 349 W and VO_{2max} of 60 mL/min per kg)[14] and elite triathletes in a 3-min protocol.[13] This result shows that our participants were well trained cyclists, and at least closely approached the level of elite cyclists. Based on VO_{2max} they are comparable to participants in previous studies[9,15,16] investigating the effects of rHuEPO on cycling performance with the highest exercise values (63–65 mL/min per kg) determined with short exercise protocols. Moreover, based on maximal power output our participants are better trained cyclists than participants reported in previous studies[9,17–21] with short protocols (from 311 W to 402 W).

The average administered rHuEPO dose (6000 IU/week) in our study is difficult to compare with previous studies[9,15–18,20–23] on the effects of rHuEPO in cyclists, which range from 4500 IU/week to 26 000 IU/week over a period of 4–12 weeks. Despite widely varying dose regimens these studies reported similar increases in haemoglobin concentration and haematocrit, without a clear dose-response association. This absence of a clear dose-response is likely to be due to a ceiling effect at the higher doses. The dose in our study is also consistent with known practices in professional cycling.[24]

The net improvement of about 5% for VO_{2max} when rHuEPO treatment was compared with the placebo was in line with previous studies.[9,15–23,25–28] Similar effects were found on maximal power output, with an increase compared with placebo of about 4%, which is also in line with the few previously reported effects,[9,17–21] although the effect there was seemingly slightly larger (6–13%).

Submaximal exercise parameters mimic the physical exertion of real-life cycling races more closely. Extracted from the maximal exercise test, VT1 and VT2 show an increase compared with placebo. However, these parameters do not directly reflect submaximal exercise as exerted during endurance performance. In a 45-min submaximal exercise test, the increases in mean power and VO_2 for the rHuEPO group were small and did not differ compared with placebo in this adequately powered study.

Results of other studies[17–19] found remarkable increases in reported submaximal tests, namely constant-load time-to-exhaustion tests, of 22–70%. These trials used short (between 3 min and 20 min) tests that, similar to the maximal exercise test, lead to exhaustion and therefore are less representative of real-life cycling. Our submaximal test was designed to closely mimic a road time trial of 45 min and in line with that was not intended to lead to exhaustion. Additionally, participants in the rHuEPO group in two of these trials were aware that they were treated with rHuEPO.[18,19] The third trial, a double- blind, placebo-controlled trial by Annaheim and colleagues,[17] did not find any effects of rHuEPO on VO_2 during this test similar to our finding. Our results could reflect a previous finding[29] that the best predictor of time trial performance in highly trained athletes is muscle oxidative capacity, in

contrast to the best predictor of maximal power output, which is oxygen delivery. The absence of a clear effect on average power in the submaximal test does not support an effect of rHuEPO on oxidative capacity.

The goal of using rHuEPO in professional sports is to improve performance during road races, not in maximal exercise tests. Participants therefore took part in a race designed to mimic a professional road race at Mont Ventoux about 12 days after the last dose of the treatment period, which also tested the validity of our laboratory exercise tests as biomarkers of real cycling performance. The two treatment groups did not differ in race time or mean power output, thereby raising doubt about the predictive value of the increase in maximal exercise test parameters by rHuEPO for performance in a road race. This outcome is further supported by the fact that rHuEPO treatment did not show an appreciable effect on a submaximal exercise test in the laboratory. Previous investigators[30,31] have suggested a strong predictive power of parameters obtained during laboratory maximal exercise tests such as maximal power output or power at VT for endurance capacity in a laboratory time trial (correlations of 0.80–0.91), which is similar to our findings (0.76). However, parameters from the maximal exercise test a week before the Mont Ventoux race and the uphill race times only showed a moderate correlation at best. This correlation could explain the discrepancy between the effect of rHuEPO treatment on maximal exercise test parameters and the absence of an effect on race time; many extra variables affect performance in real life. In line with this result, participants poorly predicted whether they had received rHuEPO or placebo during the study. Although the participants were not professional cyclists and might be slightly less sensitive to changes in performance, this result does seem to invalidate the claim of many professional cyclist that they could feel the effect of the rHuEPO treatment.

At the time of the first maximal exercise test, which was 11 days after the start of the treatment, the effects in the maximal exercise tests were already clearly visible. Only one other study[17] has investigated such early effects on exercise parameters, and their findings appear to confirm our findings. The magnitude of the increase at 11 days in the rHuEPO group was similar to the increase achieved in the placebo group during the entire study period. This early effect was unexpected because at this time the increase in haemoglobin concentration in the rHuEPO group was small, which followed the described time course of the effects on haemoglobin.[32]

Although our data show a correlation between haemoglobin concentration and performance in the maximal and submaximal exercise tests, this association is only true for the rHuEPO group; an association is absent in the placebo group, despite a similar range of haemoglobin concentrations. These findings indicate that there could be other mechanisms for the effect of rHuEPO on exercise tests, like via 2,3-bisphosphoglycerate or the monocarboxylate transporter.[33,34]

We did not observe a difference in incidence or severity of adverse events between the rHuEPO and placebo groups, nor did we detect any clinical signs of thrombosis, effects on blood pressure, coagulation factors, or most of the endothelial function markers. However, we did find a significant rise in E-selectin and P-selectin compared with the placebo group, which are cell adhesion molecules that play a crucial part in thrombogenesis and inflammation.[35,36] This rise might explain the observed increased thrombogenicity and increased risk of stroke after rHuEPO treatment in patients.[7] As the incidence of events such as stroke is relatively low in patients (2.6%) and will be even lower in healthy athletes,[7] our study did not have the power to detect such an increased risk. However, given the increase in endothelial function markers observed in this study, and with widespread and uncontrolled use of rHuEPO among athletes, it is not unlikely that in this population the risk of cardiovascular events also increased. Our study design has several inevitable limitations. Because of WADA regulations, it is currently impossible to do intervention studies with banned substances in professional cyclists. The question remains whether our data can be applied to professional cyclists. However, in our participants there was no indication that rHuEPO treatment had a stronger effect on maximal power output in the highest performing participants.

The size of our study might have generated insufficient statistical power to detect a difference on the road race. However, for an effect of doping to be relevant for cyclists, it should have a clear effect on time trial or race performance. The absence of an effect of rHuEPO treatment on both a 45-min submaximal exercise test and a road race indicates that the effect is at best very small, and disappears in all other variability that is present during such an event.

Finally, there have been suggestions in the literature that the effect of rHuEPO treatment is not, or only partly, mediated by the increase in red blood cell mass. Suggestions for such pleiotropic effects range from effects on the speed of recovery after exercise, direct effects on skeletal muscle, improved lipolysis, psychological

effects on fatigue or motivation, and effects on the immune system. In this case, the design of the tests or timing related to the administration of rHuEPO might have been suboptimal. For example, our race was not a multi-day event, such as the Tour de France, and therefore could not show potential effects of rHuEPO that only occur in such a setting. However, to be of clinical relevance, many of the other effects would have to show an effect on time trial or road race performance, which was not observed in our study.

In the clinical practice of nephrology and oncology in which rHuEPO is used therapeutically, a well recognised optimal dose exists, which results in lower than normal haemoglobin concentrations. When haemoglobin is restored to normal concentrations, mortality increases.[7] In these studies and studies with blood transfusion, such effects only become clear after a medicine is tested in circumstances that resemble the clinical situation. Our study shows that testing in a clinical situation is not different for drugs intended to enhance sport performance. Although we did not find clinical signs of adverse effects of rHuEPO treatment, the observed rise in endothelial function markers might indicate an increased thrombogenicity. Moreover, effects on relevant performance measures were small, largely disappeared in the submaximal test, and were undetectable in a real-life cycling race.

In summary, we showed that it is possible to test potential doping substances in well controlled clinical trials and that results are much less pronounced than claimed in popular literature and accounts. More clinical research like this study will provide the evidence base for the prohibited list and might lead to more focused attention and adequate information to athletes and their medical staff. Overall, the results of our study showed that rHuEPO treatment enhanced performance in well trained cyclists in a laboratory-based maximal exercise test leading to exhaustion, but did not improve submaximal exercise test or road race performance.

REFERENCES

1 The World Anti-doping agency. 2015 Anti-Doping Testing Figures Anti-Doping Testing Figures by Laboratory. Vol 22.; 2015. https://www.wada-ama.org/sites/default/files/resources/files/2015_wada_anti-doping_testing_figures_report_0.pdf (accessed May 24, 2017).

2 The World Anti-doping agency. Standard prohibited list 2017. January 2017. https://www.wada-ama.org/sites/default/files/resources/files/2016-09-29_-_wada_prohibited_list_2017_eng_final. pdf (accessed May 24, 2017).

3 The World Anti-doping agency. World anti-doping code 2015. 2015. https://www.wada-ama.org/en/resources/the-code/world-anti-doping-code. (accessed May 24, 2017).

4 Heuberger JAAC, Cohen Tervaert JM, Schepers FML, et al. Erythropoietin doping in cycling: Lack of evidence for efficacy and a negative risk-benefit. Br J Clin Pharmacol. 2013;75(6):1406-1421. doi:10.1111/bcp.12034.

5 K.Messmer, D.H.Lewis, L.Sunder-Plassmann, W.P.Kloverkorn, N.Mendler KH. Acute Normovolemic Hemodilution. Eur Surg Res. 1972;4(55-70):55-70.

6 Lucia A, Pardo J, Durantez A, Hoyos J, Chicharro JL. Physiological differences between professional and elite road cyclists. Int J Sport Med. 1998;19(0172-4622 (Print)):342-348.

7 Pfeffer MA, Burdmann EA, Chen C-Y, et al. A Trial of Darbepoetin Alfa in Type 2 Diabetes and Chronic Kidney Disease. N Engl J Med. 2009;361(21):2019-2032. doi:10.1056/NEMJoa0907845.

8 Mclelian TM. Ventilatory and Plasma Lactate Response with Different Exercise Protocols: A Comparison of Methods. Int J Sport Med. 1985;6:30-35.

9 Connes P, Perrey S, Varray A, Prefaut C, Caillaud C. Faster oxygen uptake kinetics at the onset of submaximal cycling exercise following 4 weeks recombinant human erythropoietin (r-HuEPO) treatment. PflugersArch. 2003;447(0031-6768 (Print)):231-238.

10 Lucía A, Hoyos J, Pérez M, Santalla A, Chicharro JL. Inverse relationship between VO2max and economy/efficiency in world-class cyclists. Med Sci Sports Exerc. 2002;34(12):2079-2084. doi:10.1249/01.MSS.0000039306.92778.DF.

11 Thoden JS. Testing aerobic power. In: MacDougall JD, Wenger HA, Green HJ, eds. Physiological Testing of the High-Performance Athlete. Champaign: Human Kinetics; 1991:107-174.

12 Bentley DJ, Newell J, Bishop D. Incremental exercise test design and analysis: Implications for performance diagnostics in endurance athletes. Sports Med. 2007;37(7):575-586. doi:10.2165/00007256-200737070-00002.

13 Bentley D, McNaughton L. Comparison of W (peak), VO2 (peak) and the ventilation threshold from two different incremental exercise tests: relationship to endurance performance. J Sci Med Sport. 2003;6(June):422-435. doi:10.1016/S1440-2440(03)80268-2.

14 Millán S, Bing K, Brill C, Hill JC, Miller LE. randomized controlled trial of Micro-Mobile compression® on lactate clearance and subsequent exercise performance in elite male cyclists. Open Access J Sport Med. 2013;4:221-227. doi:10.2147/OAJSM.S51956.

15 Birkeland KI, Stray-Gundersen J, Hemmersbach P, Hallen J, Haug E, Bahr R. Effect of rHEPO administration on serum levels of sTfR and cycling performance. Med SciSportsExerc. 2000;32(0195-9131 (Print)):1238-1243.

16 Ninot G, Connes P, Caillaud C. Effects of recombinant human erythropoietin injections on physical self in endurance athletes. J Sport. 2006;24:383-391.

17 Annaheim S, Jacob M, Krafft A, Breymann C, Rehm M, Boutellier U. RHEPO improves time to exhaustion by non-hematopoietic factors in humans. Eur J Appl Physiol. 2016;116(3):623-633. doi:10.1007/s00421-015-3322-6.

18 Wilkerson DP, Rittweger J, Berger NJA, Naish PF, Jones AM. Influence of recombinant human erythropoietin treatment on pulmonary O2 uptake kinetics during exercise in humans. J Physiol. 2005;568(Pt 2):639-652. doi:10.1113/jphysiol.2005.089920.

19 Thomsen JJ, Rentsch RL, Robach P, et al. Prolonged administration of recombinant human erythropoietin increases submaximal performance more than maximal aerobic capacity. Eur J Appl Physiol. 2007;101(4):481-486. doi:10.1007/s00421-007-0522-8.

20 Audran M, Gareau R, Matecki S durand, Chenard F, Marion CSM, Bressolle BF. Effects of erythropoietin administration in training athletes and possible indirect detection in doping control. Med Sci Sport Exerc. 1999;31(5):639-645.

21 Russell G, Gore CJ, Ashenden MJ, Parisotto R, Hahn AG. Effects of prolonged low doses of recombinant human erythropoietin during submaximal and maximal exercise. EurJ ApplPhysiol. 2002;86(1439-6319 (Print)):442-449.

22 Lundby C, Achman-Andersen NJ, Thomsen JJ, Norgaard AM, Robach P. Testing for recombinant human erythropoietin in urine: problems associated with current anti-doping testing. J Appl Physiol. 2008;105(2):417-419. doi:10.1152/japplphysiol.zdg-8311.2008.

23 Ekblom B, Berglund B. Effect of erythropoietin administration on maximal aerobic power. Scand J Med Sci Sport. April 1991:88-93.

24 Hamilton T, Coyle D. The Secret Race: Inside the Hidden World of the Tour de France. Random House Publishing Group; 2012.

25 Lundby C, Damsgaard R. Exercise performance in hypoxia after novel erythropoiesis stimulating protein treatment. ScandJ Med Sci Sport. 2006;16(0905-7188 (Print)):35-40.

26 Parisotto R, Gore CJ, Emslie KR, et al. A novel method utilising markers of altered erythropoiesis for the detection of recombinant human erythropoietin abuse in athletes. Haematologica. 2000;85(0390-6078 (Print)):564-572.

27 Rasmussen P, Foged EM, Krogh-Madsen R, et al. Effects of erythropoietin administration on cerebral metabolism and exercise capacity in men. J Appl Physiol. 2010;109(2):476-483. doi:10.1152/japplphysiol.00234.2010.

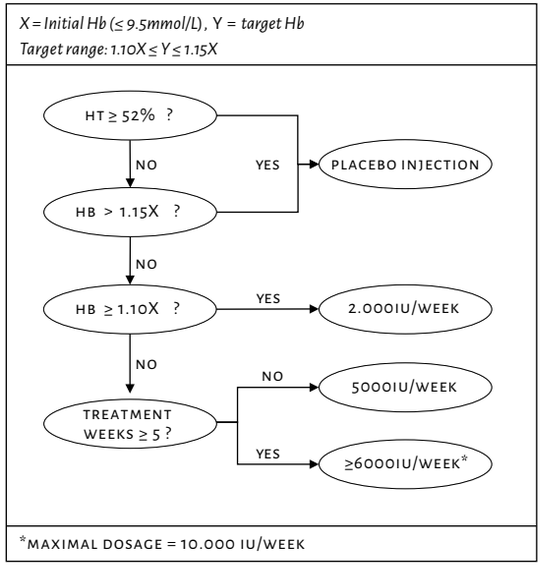
28 Souillard A, Audran M, Bressolle F, Gareau R, Duvallet A, Chanal JL. Pharmacokinetics and pharmacodynamics of recombinant human erythropoietin in athletes. Blood sampling and doping control. Br J Clin Pharmacol. 1996;42(0306-5251 (Print)):355-364.

29 Jacobs RA, Rasmussen P, Siebenmann C, et al. Determinants of time trial performance and maximal incremental exercise in highly trained endurance athletes. 2011:1422-1430. doi:10.1152/japplphysiol.00625.2011.

30 Amann M, Subudhi AW, Foster C. Predictive validity of ventilatory and lactate thresholds for cycling time trial performance. Scand J Med Sci Sport. 2006;16(1):27-34. doi:10.1111/j.1600-0838.2004.00424.x.

31 Bentley DJ, McNaughton LR, Thompson D, Vleck VE, Batterham AM. Peak power output, the lactate threshold, and time trial performance in cyclists. Med Sci Sport Exerc. 2001;33(0195-9131 (Print)):2077-2081.
32 Krzyzanski W, Jusko WJ, Wacholtz MC, Minton N, Cheung WK. Pharmacokinetic and pharmacodynamic modeling of recombinant human erythropoietin after multiple subcutaneous doses in healthy subjects. Eur J Pharm Sci. 2005;26(3-4):295-306. doi:10.1016/j.ejps.2005.06.010.
33 Crowley JP, Chazan JA, Metzger JB, Pono L, Valeri CR. Blood rheology and 2,3-diphosphoglycerate levels after erythropoietin treatment. Ann Clin Lab Sci. 1993;23(1):24-32.
34 Connes P, Caillaud C, Mercier J, Bouix D, Casties JF. Injections of recombinant human erythropoietin increases lactate influx into erythrocytes. J Appl Physiol. 2004;97(1):326-332. doi:10.1152/japplphysiol.00715.2003.
35 Myers D, Farris D, Hawley A, et al. Selectins Influence Thrombosis in a Mouse Model of Experimental Deep Venous Thrombosis. J Surg Res. 2002;108(2):212-221. doi:10.1006/jsre.2002.6552.
36 Merten M, Thiagarajan P. P-selectin in arterial thrombosis. Z Kardiol. 2004;93(11):855-863. doi:10.1007/s00392-004-0146-5.

FIGURE 1 DECISION TREE FOR DOSING Decision tree for dosing schedule that was applied before every administration of NeoRecormon/placebo during the 8 week treatment period.



Ht: haematocrit, Hb: haemoglobin, IU: International Units.

FIGURE 2 TRIAL PROFILE rHuEPO, recombinant human erythropoietin.

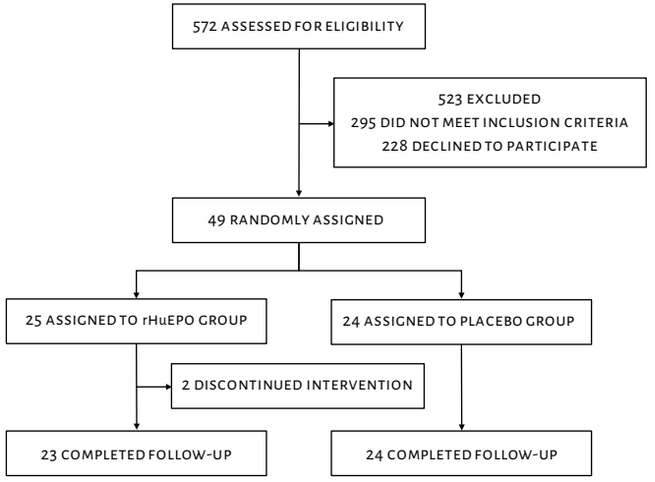


FIGURE 3 MEAN HAEMOGLOBIN CONCENTRATIONS DURING STUDY p<0.0001. rHuEPO, recombinant human erythropoietin

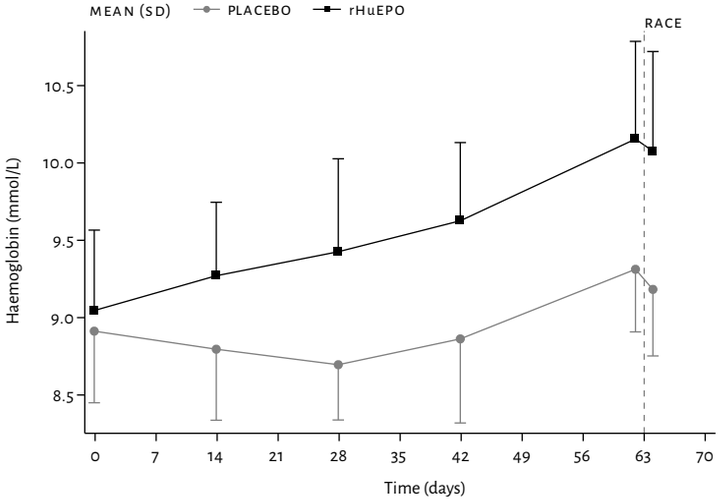


FIGURE 4 MAXIMAL POWER OUTPUT AND POWER AT VT1 DURING THE STUDY Mean maximal power output (p=0.055) and mean power output at VT1 (p=0.0100). P_{max}, maximal power output; rHuEPO, recombinant human erythropoietin; VT1, ventilatory threshold 1.

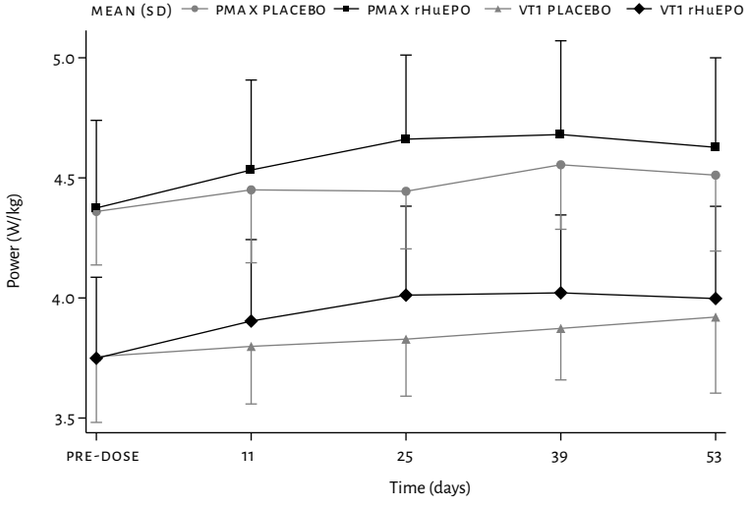


FIGURE 5 CORRELATION BASELINE AND DELTA MAXIMAL POWER AT WEEK 8 This figure displays the correlation between baseline maximal power output and the delta in maximal power output between baseline and week 8 for both treatment groups. Correlation rHuEPO: -0.21; Correlation Placebo: -0.12. rHuEPO: recombinant human erythropoietin; Regr: regression line.

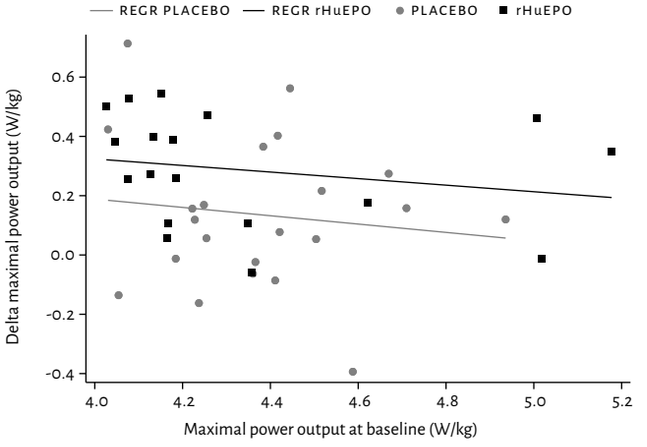
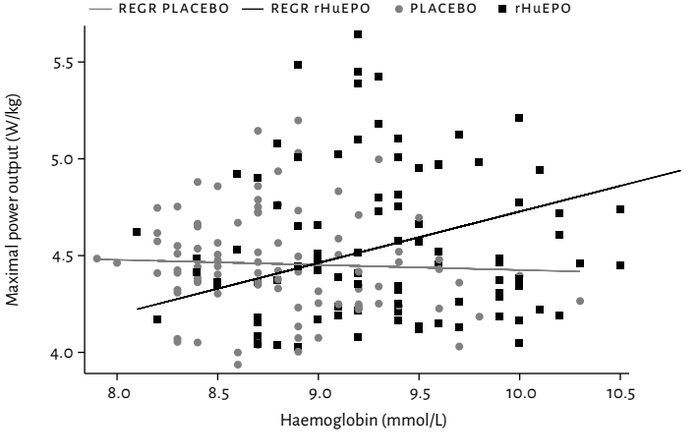


FIGURE 6 REGRESSION HAEMOGLOBIN LEVEL AND MAXIMAL POWER OUTPUT PER KG
This figure displays the regression line in a scatter graph between haemoglobin level and maximal power output per kg for both treatment groups.



*Regression rHuEPO: [P_{max} per kg] = 0.27 [CI 0.16; 0.37] * [Hb] + 2.1;*
*Regression Placebo: [P_{max} per kg] = -0.03 [CI -0.15; 0.10] * [Hb] + 4.7.*
rHuEPO: recombinant human erythropoietin; Regr: regression line; P_{max}: maximal power output.

TABLE 1 BASELINE CHARACTERISTICS Data are median (range) or mean (SD). rHuEPO, recombinant human erythropoietin; VO_{2max}, maximal oxygen consumption.

	Placebo	rHuEPO
n	24	24
Age (years)	33.8 (20.0 - 50.0)	33.5 (22.0 - 48.0)
Weight (kg)	76.9 (8.9)	77.0 (8.9)
Height (cm)	186 (6.7)	186 (7.9)
Hb (mmol/L)	8.9 (0.46)	9.0 (0.52)
Ht (L/L)	0.431 (0.0221)	0.433 (0.0222)
Maximal Power output per kg (W/kg)	4.36 (4.03 - 4.94)	4.37 (4.03 - 5.18)
VO_{2max} (mL/min/kg)	56.0 (4.111)	55.4 (5.132)

TABLE 2 EXCLUDED PARTICIPANTS Reasons for excluding participants from participation

Participant could not be contacted despite multiple attempts	27	
Participant was unwilling to discontinue membership with anti-doping regulation	32	
Study planning made it impossible for participant to participate	76	
Participant considered travel to study site (Leiden, the Netherlands) too far	8	
Participant was not interested in study participation anymore	52	
Participant did not agree with remuneration	3	
Withdrew before screening visit		198
Participant was not a cyclist	13	
Body mass index was too high	5	
Speed on solo cycle tour <30 km/h	76	
Weekly training effort <3 hours	34	
Participant could not communicate in Dutch language	82	
Participants' age exceeded 50	20	
Blood- or plasma donation within 3 months of study participation	4	
Medical history or prohibited medication use	10	
Preliminary screening by telephone		244
Withdrew consent before baseline visit	30	
Insufficient exercise performance (Pmax < 4W/kg)	45	
Hemoglobin > 9.8 mmol/L	5	
Medical history	1	
Failed in- and exclusion criteria		81
Total		**523**

TABLE 3 rHuEPO DOSAGES DURING THE STUDY COURSE Number of participants receiving the corresponding dose. *1 participant dropped out after week 4. IU: international units.

Dosage week	1	2	3	4	5*	6*	7*	8*
Dosage								
0	-				1	2	2	-
2000 IU	-		1	1	4	5	2	2
5000 IU	24	24	23	23	-	-	-	-
6000 IU	-				5	7	3	1
8000 IU	-				9	3	11	3
10000 IU	-				4	6	5	17

TABLE 4 HAEMATOCRIT AND HAEMOGLOBIN LEVELS Raw baseline (and SD) and EM (Estimated Mean) values of haematological parameters at the different time points for both treatment groups, including the estimated differences between the treatment groups. Data analysed with a mixed model analysis of variance with fixed factors treatment, time and treatment by time, random factor participant and the pre-value as covariate.

Parameter	Treatment	Raw baseline	EM Week 2	EM Week 4	EM Week 6	EM Pre Race	Difference between groups
Haemoglobin lab (mmol/L)	Placebo	8.9 (0.5)	8.8	8.7	8.9	9.4	0.60 (0.44, 0.77) p=<0.0001
	rHuEPO	9.0 (0.5)	9.2	9.4	9.6	10.1	
Haematocrit lab (L/L)	Placebo	0.431 (0.022)	0.437	0.436	0.438	0.460	0.0330 (0.0250, 0.0409) p=<0.0001
	rHuEPO	0.433 (0.022)	0.458	0.470	0.474	0.499	

TABLE 5 DIFFERENCE IN EXERCISE PERFORMANCE PARAMETERS AT MAXIMAL EXERCISE TEST BETWEEN EACH TREATMENT GROUP Data are raw baseline (SD) or estimated mean (EM) values. For log-transformed parameters, a back-transformed estimate of the difference in percentage is reported and geometric means for estimated mean. Data were analysed with a mixed model analysis of variance with three fixed factors (treatment, time, and treatment by time), one random factor (participant), and one covariate (pre-value). rHuEPO, recombinant human erythropoietin.

Parameter	Treatment	Raw baseline	EM Day 11	EM Day 25	EM Day 39	EM Day 53	Difference between groups
Maximal oxygen consumption (L/min)	Placebo	4.298 (0.486)	4.298	4.296	4.450	4.373	0.2237 (0.0824, 0.3650) p=0.0026
	rHuEPO	4.237 (0.466)	4.475	4.536	4.688	4.612	
Max. oxygen consumption per kg (mL/min/kg)	Placebo	55.946 (4.136)	56.392	56.528	58.616	58.122	2.7066 (0.9105, 4.5027) p=0.0041
	rHuEPO	55.322 (5.453)	58.564	59.472	61.611	60.838	
Maximal Power output (W)	Placebo	335.00 (33.039)	339.59	339.97	345.72	339.65	10.315 (3.465, 17.166) p=0.0040
	rHuEPO	335.14 (34.464)	346.19	351.60	354.95	353.46	
Maximal Power output per kg (W/kg)	Placebo	4.36 (0.223)	4.46	4.48	4.56	4.51	0.109 (-0.002, 0.220) p=0.055
	rHuEPO	4.37 (0.365)	4.53	4.60	4.66	4.65	
Lactate threshold Power (W)	Placebo	299.51 (7.671)	298.00	296.23	306.09	298.54	8.493 (-0.609, 17.595) p=0.067
	rHuEPO	290.87 (6.451)	305.65	310.66	311.05	305.47	
Lactate threshold Power per kg (W/kg)	Placebo	3.90 (0.215)	3.92	3.91	4.04	3.97	0.106 (-0.025, 0.238) p=0.11
	rHuEPO	3.84 (0.304)	4.01	4.09	4.12	4.05	
Lactate threshold VO_2 (L/min)	Placebo	4.005 (0.461)	3.948	3.934	4.079	3.963	0.1160 (-0.0165, 0.2485) p=0.084
	rHuEPO	3.855 (0.459)	4.002	4.160	4.209	4.017	
Lactate threshold VO_2 per kg (mL/min/kg)	Placebo	52.159 (3.239)	51.918	51.921	53.885	52.749	1.4375 (-0.4529, 3.3279) p=0.13
	rHuEPO	50.864 (4.476)	52.502	54.863	55.674	53.185	
Ventilatory Threshold 1 VO_2 (L/min)	Placebo	3.863 (0.431)	3.862	3.857	3.916	3.953	0.1781 (0.0481, 0.3081) p=0.0083
	rHuEPO	3.747 (0.434)	3.989	4.057	4.156	4.099	
Ventilatory Threshold 1 VO_2 per kg (mL/min/kg)	Placebo	50.291 (4.056)	50.671	50.772	51.538	52.657	2.0875 (0.3802, 3.7949) p=0.018
	rHuEPO	48.912 (4.974)	52.092	53.261	54.560	54.076	
Ventilatory Threshold 1 Power (W)	Placebo	288.69 (32.883)	290.92	292.28	294.66	294.68	9.701 (2.437, 16.964) p=0.010
	rHuEPO	287.08 (30.562)	299.60	301.72	303.73	306.30	
Ventilatory Threshold 1 Power per kg (W/kg)	Placebo	3.76 (0.274)	3.81	3.84	3.87	3.92	0.119 (0.018, 0.219) p=0.022
	rHuEPO	3.75 (0.337)	3.91	3.96	4.00	4.04	
Ventilatory Threshold 2 VO_2 (L/min)	Placebo	4.077 (0.452)	4.089	4.058	4.145	4.125	0.1621 (0.0021, 0.3221) p=0.047
	rHuEPO	3.985 (0.462)	4.190	4.212	4.343	4.319	
Ventilatory Threshold 2 VO_2 per kg (mL/min/kg)	Placebo	53.097 (4.295)	53.658	53.422	54.577	54.873	1.9471 (-0.1110, 4.0053) p=0.063
	rHuEPO	52.005 (5.114)	54.816	55.279	57.149	57.074	
Ventilatory Threshold 2 Power (W)	Placebo	306.77 (32.450)	311.34	310.80	312.35	311.43	8.423 (-0.855, 17.700) p=0.074
	rHuEPO	306.15 (35.301)	316.76	316.96	320.88	325.01	
Ventilatory Threshold 2 Power per kg (W/kg)	Placebo	3.99 (0.285)	4.08	4.09	4.11	4.14	0.100 (-0.031, 0.230) p=0.13
	rHuEPO	3.99 (0.354)	4.15	4.15	4.23	4.29	

(continuation on next page)

TABLE 5 *(continuation of previous page)*

Gross Efficiency (%)	Placebo	21.5 (0.792)	21.7	21.8	21.4	21.4	0.4% (-1.6%, 2.5%) p=0.69
	rHuEPO	21.9 (1.096)	21.8	22.0	21.4	21.7	
Heart Rate (BPM)	Placebo	182 (8.78)	182	182	184	181	0.7 (-1.7, 3.1) p=0.56
	rHuEPO	184 (8.17)	183	182	184	182	
Tidal Volume (L)	Placebo	3.029 (0.458)	3.115	3.139	3.162	3.173	0.0288 (-0.1060, 0.1637) p=0.67
	rHuEPO	3.160 (0.340)	3.168	3.177	3.214	3.146	
Respiratory Frequency (1/min)	Placebo	54.3 (8.566)	53.2	52.2	53.8	51.6	0.54 (-2.00, 3.08) p=0.67
	rHuEPO	51.5 (5.887)	52.4	52.7	54.5	53.4	
Respiratory minute ventilation (L/min)	Placebo	161.8 (18.383)	163.2	161.5	167.4	161.7	2.84 (-4.01, 9.69) p=0.41
	rHuEPO	161.3 (21.627)	163.1	164.0	172.9	165.2	
Respiratory Quotient	Placebo	1.07 (0.044)	1.07	1.06	1.06	1.06	-0.003 (-0.022, 0.016) p=0.77
	rHuEPO	1.08 (0.050)	1.07	1.05	1.06	1.06	

TABLE 6 DIFFERENCE IN EXERCISE PERFORMANCE PARAMETERS AT SUBMAXIMAL EXERCISE TEST BETWEEN EACH TREATMENT GROUP Data are raw baseline (SD) or estimated mean (EM) values. For log-transformed parameters, a back-transformed estimate of the difference in percentage is reported, and geometric means for estimated mean. Data were analysed with an analysis of variance with factor treatment, and, if available, pre-value as covariate. rHuEPO, recombinant human erythropoietin.

Parameter	Treatment	Raw baseline	EM Day 46	EM Change from baseline	Difference between groups
Power output submaximal (W)	Placebo	268.00 (27.751)	277.28	7.656	5.898 (-0.872, 12.668) p=0.086
	rHuEPO	270.83 (30.503)	283.18	13.554	
Power output submaximal per kg (W/kg)	Placebo	3.50 (0.199)	3.66	0.135	0.062 (-0.035, 0.159) p=0.20
	rHuEPO	3.53 (0.298)	3.72	0.197	
Average VO_2 submaximal (L/min)	Placebo	3.631 (0.345)	3.758	0.0866	0.0624 (-0.0862, 0.2110) p=0.40
	rHuEPO	3.701 (0.510)	3.821	0.1490	
Average VO_2 submaximal per kg (mL/min/kg)	Placebo	47.594 (3.930)	49.642	1.6189	0.6463 (-1.3074, 2.6000) p=0.51
	rHuEPO	48.180 (4.657)	50.288	2.2652	
Average Heart Rate submaximal (BPM)	Placebo	160 (9.47)	160	-1.0	-0.9 (-4.5, 2.7) p=0.62
	rHuEPO	162 (11.53)	159	-1.9	
Submaximal Lactate 10 min (mmol/L)	Placebo	2.51 (1.090)	2.45	-0.148	-0.301 (-0.864, 0.262) p=0.29
	rHuEPO	2.68 (1.456)	2.15	-0.449	
Submaximal Lactate 30 min (mmol/L)	Placebo	2.78 (1.232)	3.23	0.477	-0.050 (-0.698, 0.598) p=0.88
	rHuEPO	2.67 (1.200)	3.18	0.427	
Submaximal Lactate 45 min (mmol/L)	Placebo	3.33 (1.458)	5.09	1.448	-0.065 (-1.401, 1.271) p=0.92
	rHuEPO	3.80 (1.882)	5.03	1.383	
Cycling economy (W/L/min)	Placebo	73.87 (4.117)	73.86	0.5%	0.2% (-3.0%, 3.5%) p=0.89
	rHuEPO	73.52 (4.429)	74.03	0.7%	

TABLE 7 RESULTS RACE ON MONT VENTOUX EM (Estimated Mean) values of Mont Ventoux race results for both treatment groups, including the estimated differences between the treatment groups. For log transformed parameters a back transformed estimate of the difference in percentage is reported, and geometric means for EM. The racing times are analysed with a parametric model for failure time (accelerated failure time regression model) with right censored values. The model for the response variable consists of a linear effect composed of the covariates and a random disturbance term. The covariates are treatment and pre-value. The time to arrival is log-transformed before analysis and the chosen distribution is normal. The P_{max}/kg pre-treatment value is chosen as pre-value to correct for possible differences in baseline performance.

Parameter	Treatment	EM	Difference between groups
Average time in race (hr:m:s)	Placebo	1:40:15	0.3% (-8.3%, 9.6%) p=0.95
	rHuEPO	1:40:32	
Average Efficiency (%)	Placebo	49.8	0.40 (-3.95, 4.76) p=0.85
	rHuEPO	50.2	
Average Power (W)	Placebo	266.37	-3.5% (-13.0%, 7.0%) p=0.48
	rHuEPO	256.93	
Average Power per kg (W/kg)	Placebo	3.09	-1.7% (-11.0%, 8.6%) p=0.73
	rHuEPO	3.03	

TABLE 8 WEIGHT Raw baseline (and SD) and EM (Estimated Mean) values of weight at the different time points for both treatment groups, including the estimated differences between the treatment groups. Data analysed with a mixed model analysis of variance with fixed factors treatment, time and treatment by time, random factor participant and the pre-value as covariate.

Parameter	Treatment	Raw baseline	EM Week 2	EM Week 4	EM Week 6	EM Week 7	EM Week 8	EM Pre race	EM Post race	Difference between groups
Weight (kg)	Placebo	76.94 (8.94)	76.57	76.25	76.12	75.83	75.80	77.43	75.80	0.303 (-0.507, 1.112) p=0.46
	rHuEPO	77.03 (8.99)	76.65	76.61	76.50	76.41	76.27	77.50	75.98	

TABLE 9 VITAL SIGNS Raw baseline (and SD) and EM (Estimated Mean) values of vital signs at the different time points for both treatment groups, including the estimated differences between the treatment groups. Data analysed with a mixed model analysis of variance with fixed factors treatment, time and treatment by time, random factor participant and the pre-value as covariate.

Parameter	Treatment	Raw baseline	EM Week 1	EM Week 2	EM Week 3	EM Week 4	EM Week 5	EM Week 6	EM Week 7	Difference between groups
Systolic blood pressure (mmHg)	Placebo	119 (10)	119	116	117	117	118	115	116	2.3 (-1.1, 5.7) p=0.18
	rHuEPO	120 (11)	122	121	119	118	118	119	118	
Diastolic blood pressure (mmHg)	Placebo	71 (7)	73	74	72	71	72	72	71	0.4 (-2.0, 2.8) p=0.73
	rHuEPO	74 (9)	73	72	74	73	72	71	73	
Heart rate supine (BPM)	Placebo	53 (6)	54	53	52	52	51	53	52	-1.3 (-3.9, 1.4) p=0.34
	rHuEPO	55 (11)	52	50	53	50	50	52	51	

TABLE 10 GRADE 1–2 TREATMENT-EMERGENT ADVERSE EVENTS IN DIFFERENT SYSTEM ORGAN CLASSES Data are number of participants with TEAE (and total number of TEAEs) by System Organ Class and Preferred Term, grouped by intensity and treatment group. TEAEs of grade 1–2 were only reported if they occurred in more than 10% of the subjects. rHuEPO, recombinant human erythropoietin; TEAE, treatment-emergent adverse event.

System organ class	Adverse events	N subjects rHuEPO group grade 1-2	N subjects Placebo group grade 1-2
General disorders and administration site conditions	Fatigue	4 (4)	3 (3)
Immune system disorders	Seasonal allergy	3 (3)	1 (1)
Infections and infestations	Nasopharyngitis	3 (3)	1 (1)
Musculoskeletal and connective tissue disorders	Arthralgia	4 (5)	3 (3)
Musculoskeletal and connective tissue disorders	Myalgia	2 (2)	3 (3)
Musculoskeletal and connective tissue disorders	Pain in extremity	0 (0)	4 (4)
Nervous system disorders	Headache	6 (8)	3 (3)
Skin and subcutaneous tissue disorders	Rash	0 (0)	3 (3)

TABLE 11 COAGULATION AND ENDOTHELIAL FUNCTION MARKERS Raw baseline (and SD) and EM (Estimated Mean) values of coagulation and endothelial function markers at the different time points for both treatment groups, including the estimated differences between the treatment groups. Data analysed with a mixed model analysis of variance with fixed factors treatment, time and treatment by time, random factor participant and the pre-value as covariate.

Parameter	Treatment	Raw baseline	EM Day 11	EM Day 14	EM Day 25	EM Day 28	EM Day 39	EM Day 42	EM Day 53	EM Pre race	Difference between groups
Activated partial Thromboplastin time (s)	Placebo	30.7 (1.4)	30.1	30.1	29.8	30.0	30.4	29.6	30.4	30.0	1.5% (-0.3%, 3.4%) p=0.097
	rHuEPO	29.9 (1.9)	29.8	30.4	30.4	30.4	31.0	30.4	31.1	30.6	
Prothrombin time (s)	Placebo	14.5 (0.6)	14.5	14.3	14.5	14.5	14.6	14.3	14.3	14.2	-0.0% (-1.6%, 1.6%) p=1.00
	rHuEPO	14.4 (0.9)	14.5	14.4	14.7	14.3	14.4	14.4	14.4	14.0	
Fibrinogen (g/L)	Placebo	2.6 (0.3)	2.4	2.5	2.5	2.5	2.4	2.4	2.5	2.5	1.3% (-3.9%, 6.8%) p=0.62
	rHuEPO	2.7 (0.4)	2.5	2.5	2.4	2.6	2.4	2.5	2.5	2.5	
D-dimer (ng/mL)	Placebo	231.6 (145.1)	253.4	222.3	213.1	217.6	186.8	217.1	205.5	251.8	-1.3% (-17.0%, 17.4%) p=0.88
	rHuEPO	258.5 (157.3)	217.1	216.4	242.8	204.9	190.8	218.6	221.4	228.9	
Creatinine phosphokinase (U/L)	Placebo	142 (84)	123	140	126	142	127	137	152	173	-9.4% (-26.9%, 12.2%) p=0.36
	rHuEPO	213 (179)	135	137	119	129	111	139	132	117	
Beta Thromboglobulin (pg/mL)	Placebo	16113 (8283)	26836	11777	22008	14520	15088	16944	15554	23881	11.2% (-10.0%, 37.6%) p=0.32
	rHuEPO	31194 (85294)	26350	13585	27421	14135	19539	16419	19785	25535	
E-selectin (pg/mL)	Placebo	5100 (1978)	4986	4964	5097	5166	4731	4740	4942	4953	8.6% (2.0%, 15.7%) p=0.011
	rHuEPO	5543 (1912)	5446	5117	5475	5270	5568	5251	5477	5387	
Prothrombin fragment1+2 (pmol/L)	Placebo	118.6 (53.4)	169.4	102.9	167.2	108.1	116.4	134.5	122.6	137.1	-0.5% (-15.9%, 17.8%) p=0.95
	rHuEPO	368.2 (1262.8)	160.5	106.1	160.8	109.9	127.5	125.5	126.5	131.1	
Factor VIII (%)	Placebo	129 (37)	119	120	129	129	125	119	117	130	3.7% (-3.1%, 10.9%) p=0.29
	rHuEPO	137 (43)	129	123	136	132	120	125	128	131	
P-selectin (pg/mL)	Placebo	8789 (2112)	9037	8714	8524	8930	8810	8721	8261	8599	7.8% (1.5%, 14.5%) p=0.016
	rHuEPO	9710 (2285)	9010	9120	9248	9590	9846	9267	9173	9789	
PF4 (pg/mL)	Placebo	43019 (46073)	68339	23340	51235	28821	33209	35929	33026	63845	13.1% (-9.3%, 41.1%) p=0.27
	rHuEPO	43988 (85138)	60195	23125	60555	31296	52786	36164	43549	72593	
Thrombine: Antithrombine (ng/mL)	Placebo	2.850 (3.457)	4.304	1.495	3.663	1.575	1.860	2.029	1.795	1.869	4.1% (-20.0%, 35.5%) p=0.76
	rHuEPO	14.737 (63.963)	3.063	1.495	4.358	1.604	2.126	1.932	2.175	2.261	
Thrombomodulin (pg/mL)	Placebo	1383 (1358)	1323	1347	1320	1368	1322	1328	1364	1385	-2.8% (-10.2%, 5.2%) p=0.48
	rHuEPO	1936 (2756)	1252	1283	1218	1317	1319	1315	1387	1377	
Von Willebrand Factor (%)	Placebo	88.434 (27.669)	81.953	86.168	85.932	91.320	86.046	88.634	83.627	90.113	-0.9% (-8.0%, 6.8%) p=0.81
	rHuEPO	100.096 (25.426)	87.897	86.114	82.513	94.138	80.810	86.512	81.131	89.014	

5 REPEATABILITY AND PREDICTIVE VALUE OF LACTATE THRESHOLD CONCEPTS IN ENDURANCE SPORTS

PLOS ONE 2018, 13(11):E0206846

Jules AAC Heuberger, Pim Gal, Frederik E Stuurman, Wouter AS de Muinck Keizer, Yuri Mejia Miranda and Adam F Cohen

ABSTRACT

Blood lactate concentration rises exponentially during graded exercise when muscles produce more lactate than the body can remove, and the blood lactate-related thresholds are parameters based on this curve used to evaluate performance level and help athletes optimise training. Many different concepts of describing such a threshold have been published. This study aims to compare concepts for their repeatability and predictive properties of endurance performance. Forty-eight well-trained male cyclists aged 18-50 performed 5 maximal graded exercise tests each separated by two weeks. Blood lactate-related thresholds were calculated using eight different representative concepts. Repeatability of each concept was assessed using Cronbach's alpha and intra-subject CV and predictive value with 45 minute time trial tests and a road race to the top of Mont Ventoux was evaluated using Pearson correlations. Repeatability of all concepts was good to excellent (Cronbach's alpha of 0.89-0.96), intra-subject CVs were low with 3.4-8.1%. Predictive value for performance in the time trial tests and road race showed significant correlations ranging from 0.65-0.94 and 0.53-0.76, respectively. All evaluated concepts performed adequate, but there were differences between concepts. One concept had both the highest repeatability and the highest predictability of cycling performance, and is therefore recommended to be used: the Dmax modified method. As an easier to apply alternative, the lactate threshold with a fixed value of 4 mmol/L could be used as it performed almost as well.

INTRODUCTION

The measurement of blood lactate is extensively used in sports medicine, although there is debate on how lactate affects fatigue in endurance athletes.[1] Nevertheless, the concentration of lactate in the blood relative to the exercise intensity is a relevant marker of endurance performance.[2–5] This can be visualised in a blood lactate curve (BLC) using a maximal graded exercise test (GXT): as the workload on the athlete increases over time, blood lactate concentrations (bLa) are measured at defined intervals. During high intensity contractions lactate is formed along with H^+ in the muscles,[6] followed by an increased elimination of lactate from plasma.[7,8] When elimination becomes saturated, bLa will start to rise when production exceeds clearance. This (exponential) rise in bLa in the BLC is of importance, as the corresponding exercise intensity is associated with endurance performance since it correlates with the transition from aerobic to anaerobic workout.[9] Since the 1960's BLCs have been analysed trying to accurately determine a point in this curve that predicts endurance performance. Although many terms have been used for this point, in this work they will be termed lactate threshold (LT) concepts. BLCs and LT concepts can be used to assess 'endurance fitness' in athletes,[10] and to evaluate the effects of and to prescribe training exercises for individual athletes.[4,5] Therefore these measures are relevant in sports medicine, both in amateur and professional sports. But as LT is based on a maximal exercise test protocol that does not directly mimic endurance exercise, finding a single point in the resulting BLC that has a strong relation to endurance performance is challenging. Moreover, determining where this single point lies in the relatively smooth curve, that is the result of a complex system of factors, can prove difficult as well. On the other hand, the more accurate method of determining maximum lactate steady state (MLSS), using several sessions with different workloads takes more time, which is the reason why an approximation of MLSS using lactate threshold concepts was developed.[11]

A previous literature review showed that there are many methods used to analyse the BLCs, with approximately 25 different concepts identified in literature to describe some form of LT.[9] These different concepts are used interchangeably throughout scientific studies and in sports and show variable repeatability and predictive value. Moreover, populations that were included in different studies often differed in training status, age and category of sport. For these reasons

there is debate about these LT concepts.[9] The aim of this study is to evaluate the repeatability and predictive value of representative concepts using a large dataset of BLCs from a group of well-trained cyclists who performed multiple GXTs, time trials (TT) and an uphill road race (RR) in the setting of a clinical study.

MATERIALS AND METHODS

Study design and participants

Blood lactate curve data in this paper were generated in a previously published study.[12] Briefly, the study was a double-blind, randomised, placebo controlled, parallel, single centre trial to evaluate the effects of recombinant human erythropoietin (rHuEPO) in forty-eight healthy male cyclists aged 18 to 50. Informed consent was obtained from all individual participants included in the study. The study was approved by the Independent Ethics Committee of the Foundation Evaluation of Ethics in Biomedical Research (Stichting Beoordeling Ethiek Biomedisch Onderzoek, Assen, Netherlands). The study is registered in the Dutch Trial Registry (Nederlands Trial Register), number NTR5643. For inclusion, participants had to be well-trained, as evaluated by a maximum power-to-weight ratio during the GXT at screening that should exceed 4 W/kg. During the eleven week study duration, twenty-four participants received weekly rHuEPO injections and twenty-four received placebo injections for eight weeks. Participants had to maintain their regular training schedule during the study.

Procedures

Maximal exercise tests

Five GXTs were performed on a Monark LC4r ergometer (COSMED, Rome, Italy) with approximately 2-week intervals between each test, see Figure 1. After a two-minute warm-up at 75 Watts, the GXT dictated an increase in pedalling resistance to 175 Watts, which increased by an additional 25 Watts every five minutes. Between 4:15 and 4:45 into each step and immediately after termination of the exercise test, blood was drawn to measure bLa. Gas exchange was measured using a Quark CPET system (COSMED, Rome, Italy) and breath-by-breath sampling technology. During the test cadence had to be maintained between 70 and 90 rpm. The test terminated when cadence could not be maintained above 70 rpm or when a participant stopped the test.

Lactate determination

During the GXTs blood for lactate determination was drawn via an IV cannula (Venflon 7 Pro Safety, BD, Switzerland) with a 30 cm extension set between the cannula and a three way stopcock for blood sampling in the antecubital vein. Before the first and after every sampling the stopcock and extension set were flushed with 2 mL saline. Before blood sampling 0.5 mL was withdrawn from the stopcock to remove any remaining saline. Next, 1 mL of blood was taken from the stopcock. Within ten seconds from withdrawal the blood was placed on the Lactate Pro 2 (Arkray, Kyoto, Japan) strip which was then inserted in the Lactate Pro 2 device. The same device was used throughout the whole study and was given at least 20 minutes to adjust to the room temperature before sampling.

Time trial tests

The time trial tests were performed twice on the same ergometer used for the GXT, with the first (TT1) 3-8 days after the first GXT and the second (TT2) one week after GXT four. Participants were instructed to produce the highest mean power output during a 45-minute period at a cadence of 70-90 rpm, attempting to mimic competitive cycling time trials. At the start of the test pedalling resistance was set at 80% of the maximal power reached during GXT1. Participants could adjust the power by indicating to increase or decrease in power by steps of 10 Watts. They were informed of the remaining time on a regular basis during the test.

Mont Ventoux race

Approximately one week after the last GXT participants competitively climbed the Mont Ventoux (Vaucluse département, France) via Bédoin, a climb of approximately 21.5 km with an average gradient of 7.5%. The race was preceded by a stage of 110 km in the French Provence (total elevation gain 1524 m) that was completed collectively. Racing bikes of participants were equipped with a Single Leg Power Meter SGY-PM910H2 (Pioneer Europe, Antwerp, Belgium) with Shimano Ultegra 6800 crank (Shimano, Osaka, Japan) to log power data on the bicycle during the race. Data were uploaded to the dedicated database Cyclo-Sphere.

Lactate threshold concepts

The BLCs from the GXTs were then used to calculate several representative LT concepts. Concepts were selected as follows: First, published concepts were retrieved from a review by Faude *et al*[9] and by a literature research within the PubMed database. The database was searched for the search terms 'lactate threshold', 'aerobic threshold', 'anaerobic threshold', 'endurance performance' or 'maximal lactate steady state' or similar terms in different combinations. The references of the selected articles were searched for further relevant articles. Secondly, retrieved concepts were divided into seven different categories: *Fixed bLa value*, *Visually determined rise from baseline*, *Predetermined increase*, *Intersection of tangents and other derivatives*, *Minimum lactate equivalent*, *Using the recovery bLa curve*, *Tangent to bLa curve*. A few retrieved concepts could not be implemented, reasons being lacking lactate concentrations in the recovery phase after exercise and no availability of the full text article describing the method of the concept despite various efforts obtaining it. (*Using the recovery bLa curve*; *Tangent to bLa curve*). From each remaining category, concepts that were representative and were used frequently in other research were selected. If there were multiple concepts in one category that were commonly used and fundamentally different in methodology, more than one concept of that category was included in the analysis. Selecting multiple commonly used, but very similar concepts from one category was not deemed useful for the purpose of this study. This resulted in a final selection of eight concepts from the five implementable categories for analysis in our study.

Implementation of lactate threshold concepts

All selected concepts were implemented according to the articles that described the concept. When exact reproduction of the method was not feasible due to the use of different parameters (e.g. running velocity was used), we approximated the description as close as possible (e.g. we used power output). For concepts that required data fitting of the blood lactate curve a third-order polynomial was chosen, based on the shape of the blood lactate curve data and given that it is a proven method, although there is no generally accepted method for data fitting.[9] An example of a blood lactate curve with a depiction of all lactate threshold concepts is shown in Figure 2.

LT1 Similar to what Tanaka described[13] we plotted bLa (mmol/L) *versus* power (W). Three authors (JH, WdMK and PG) were asked to independently select the first point in the BLC that marks a substantial increase above resting level. LT1 was defined as the power value corresponding to the point selected by at least two researchers, or in cases without consensus, the three researchers discussed until consensus was reached.

LT2 Coyle *et al*[14] determined LT as 1 mmol/L above a visually determined baseline in the BLC. We took the lactate measurement chosen as LT1 and calculated the mean of the measurements preceding this point to create an average baseline value. The power value belonging to the first measured lactate value after baseline that supersedes the baseline value plus 1 mmol/L was considered LT2.

LT3 As Dickhuth *et al*,[15] we determined the minimum lactate equivalent (the lowest value when bLa is divided by work intensity) using third-order polynomial fitting and added 1.5 mmol/L to the corresponding bLa, termed individual anaerobic threshold in the paper, to find the power value on the fitted polynomial of the BLC and termed it LT3.

LT4 As described by Amann *et al*,[16] we calculated the first rise of 1 mmol/L or more between two bLa measurements where the next rise was similar or larger than 1 mmol/L. The measurement that preceded this first increase was considered LT4.

LT5 Based on the method described by Dickhuth *et al*,[17] we divided bLa (mmol/L) by the 30 second average VO_2 (mL/min/kg) and plotted it against power. These values were interpolated with a third-order polynomial and the power value at the lowest point in this curve was considered LT5.

LT-4MMOL A widely used concept is the LT-4mmol method, as described for example by Sjodin *et al*[18] The power in the interpolated third-order polynomial BLC that corresponds to a bLa of 4 mmol/L was considered LT-4mmol.

D_{MAX} AND D_{MAX} MODIFIED Similar to the method proposed by Cheng *et al*,[19] we plotted bLa versus power, interpolated with a third-order polynomial and

plotted a line from the first measurement to the last measurement. The point in the interpolated BLC that has the maximum perpendicular distance with that line was considered Dmax. A modified version as described by Bishop *et al*,[20] uses the measurement that precedes an increase of at least 0.4 mmol/L instead of the first bLa measurement to draw the line to the last measurement, which is termed Dmax modified (Dmax-mod).

Data management

Data was stored in a validated database system (Promasys, Omnicomm Inc., Fort Lauderdale, USA) and checked for accuracy and completeness. Blinded data review before code-breaking and analysis was performed according to a standard procedure at our unit. This included evaluating whether the GXT was performed to maximal ability, which was based on power, VO_2 and bLa values and report by the subject.

Statistical analysis

We used statistical software R version 3.4.0[21] to plot measurements, calculate the third-order polynomial that best fits the data using polynomial regression with the R-function lm(y~poly(3)), implement the LT concepts and perform the statistical testing. R was used with the following packages: dplyr 0.5.0,[22] psych 1.7.5,[23] tidyr 0.6.3.[24] Data of all subjects enrolled in the study were used in the analysis.

Repeatability

To measure repeatability we determined the weighted intra-subject coefficient of variation (CV) and the Cronbach's alpha based on the five GXT results for each LT concept. Weighted intra-subject CV was calculated correcting for missing values (CV based on the sum of the variance per subject multiplied by the amount of measurements, divided by the total amount of measurements). Both the weighted intra-subject CV and Cronbach's alpha were calculated only using data from participants receiving placebo, as there might have been longitudinal effects of rHuEPO treatment on the GXTs.

Predictive properties

For the predictive properties we calculated the Pearson correlation between each LT concept and the mean power of the corresponding relevant endurance parameter. The LT concept from the GXT closest in time to the endurance tests TT1 and TT2 and road race (see Figure 1), namely GXT 1, 4 and 5 respectively, were used for correlations between the LT concept and corresponding average power output. In addition, the difference between each measurement pair was calculated and averaged to create the mean difference between the LT concept and endurance test power. This value indicates how the power at the LT concept translates to average endurance power in a time trial or race. For these Pearson correlation and mean difference analyses both subjects receiving rHuEPO and placebo were included. This was done as LT concepts are designed to be a predictive parameter for endurance exercise, which should be irrespective of a subject being treated with rHuEPO or not. In addition, given that the measurements of each pair are at most a week apart, no changes in the LT concept or endurance performance are expected due to rHuEPO. Moreover, GXT1 and TT1 were performed before starting the treatment period, and no rHuEPO administrations took place between GXT5 and the race. For these analyses therefore no treatment effect was expected and pooling was considered appropriate.

RESULTS

In total 49 subjects entered the study, of which 47 were completers (Figure 3); one subject dropped out after having performed the first GXT and time trial test and was replaced. One other subject dropped out after completing two GXTs and one time trial test and was not replaced. Subject characteristics can be found in Table 1. Of the remaining 238 planned GXTs, five were not performed due to illness or injury. An additional 22 were excluded from analysis, five due to having less than four bLa samples for the GXT, most others due to the fact that subjects indicated having physical problems (e.g. illness/injury, sore legs from recent exercise) potentially affecting test results, leaving 211 GXTs (of which 109 from placebo subjects) with analysable lactate threshold data. A total of 96 time trial tests were performed and used in the analysis, and power data of 37 subjects was available for the road race. Out of the 47 subjects that completed the study, three could not participate in the road race, four did not reach the finish line due to exhaustion, and three did not have a power meter on their bike, therefore lacking power data for the road race.

Lactate threshold concepts and endurance

All eight LT concepts were successfully implemented on the GXT data; for LT1 which was determined visually by three researchers, a unanimous decision about the lactate threshold was reached in 56.8% of the tests, in 40.0% of the cases two out of three researchers agreed and there was originally no consensus in 3.2% of the tests. Several concepts were based on the third-order polynomial data fitting, mean r-squared values of all individual curves were 0.978 (SD = 0.032, range 0.716-1.000). Mean values for each LT concept of the placebo group can be found in Table 2. Mean (SD) power output for TT1 was 268 W (28 W) in the placebo group and 271 W (29 W) in the rHuEPO group, and estimated mean for TT2 was 277 W and 283 W for the placebo and rHuEPO groups respectively. Estimated mean power during RR were 266 W and 257 W for the two groups, during a mean race time of 1 h 37 min 45 s (SD = 12 min 40 s) and 1 h 38 min 23 s (SD = 14 min 9 s), respectively.

Repeatability

The overall intra-subject CV of each LT concept is indicated in Table 2, and shows some minor differences between concepts, with LT3, LT-4mmol, Dmax and Dmax-mod having CVs < 5% and LT5 having the highest intra-subject CV with 8.1%. The Cronbach's alpha values for all LT concepts in the placebo group are between 0.89 and 0.97 and although 95% CIs largely overlap, the same four concepts as observed for intra-subject CVs perform best with Cronbach's alpha values >0.95 (Table 3).

Predictive properties

Pearson correlation coefficients and the mean difference between each correlation pair are listed in Table 4. All correlations are highly significant ($p<0.0002$), indicating the null hypothesis that the correlation is equal to zero can be rejected. The strength of the relationship differs for different concepts. Correlation with TT1 was very strong for Dmax-mod and strong for all other concepts except LT5, which showed a moderate correlation. Correlation with TT2 was strong for all concepts except LT5, which showed a moderate correlation. Correlation with RR was strong for Dmax and Dmax-mod, and moderate for all other concepts. Dmax-mod has the highest correlation with time trial test 1 (r=0.94), LT-4mmol with time trial test 2 (r=0.85) and Dmax-mod with road race power (r=0.76). The mean difference with the endurance parameters differs substantially between concepts, ranging from the lactate threshold on average being up to 45.3 W lower than the related endurance parameter for LT5 to 36.6 W higher for LT-4mmol. Linear regression between each LT concept and average race power, including accompanying r^2 values, is plotted in Figure 4.

DISCUSSION

All LT concepts that were included in this analysis performed good on repeatability and reasonable to good on predicting a lab-based time trial and a real-life road race. Nevertheless, this study identified several LT concepts that outperformed the others in the setting of this trial. The best method being Dmax-mod, but Dmax, LT-4mmol and LT3 performed well too.

Methodology

The design of the exercise protocol, for example stage duration, is known to impact blood lactate curves.[25,26] We selected an exercise protocol with five minute stages and 25 W increments because it takes 3-4 minutes for the body to reach steady state and lactate accompanying that effort level can be measured accurately.[27] In addition, longer protocols may be more sensitive to performance changes.[25] As described in more detail elsewhere,[12] GXT results show our subjects were well-trained with maximal power output and VO_{2max} values comparable to elite cyclists and triathletes when using longer exercise protocols.[28,29] All evaluated concepts were applied to data from the same exercise tests, with the same sampling and assay method, and the same fitting procedure was used for those applicable concepts. As a result such factors could not have affected the comparison between concepts within this study. The current study was designed in that way to give the most accurate estimate of performance parameters and its controlled set-up seems to be the most robust and valuable way to determine differences between concepts. Nevertheless, when any of these factors are changed (e.g. using a different exercise test protocol) it is possible the outcomes might not translate perfectly. With regards

to data fitting, the third-order polynomial in the applicable concepts performed well given the high mean r-squared values observed.

Selection of concepts

After inspection of all identified lactate concepts, it became clear that there were similarities between quite some of them. For this reason, the concepts were grouped into categories, and a selection was made of concepts to be analysed to have at least one representative per category and thereby ensuring that results from this study would be informative for all regularly used lactate concepts. The selection includes concepts such as a fixed lactate value (usually at 4 mmol/L) and the visually determined LT concepts that have been used since the conception of the LT, and more recent concepts such as LT3, LT4 and Dmax and Dmax-mod.[9]

Mean threshold

The mean power output (Table 2) is relatively constant over time for each concept. These results confirm there was no placebo-effect on any of the LT concepts, although such an impact would already theoretically be improbable. What can also be seen is that not all concepts seem to identify the same point in the blood lactate curve: LT5 gives the lowest estimate of LT (228.4 W), much lower than other concepts (274.1-302.7 W). LT-4mmol and Dmax-mod have the highest estimates (302.7 and 301.5 W), indicating these concepts identify different intensities of performance and have different physiological meanings. Applying the terminology as described in Faude *et al*,[9] based on mean threshold and mean difference with TT and RR (Table 4), some concepts seem to be more related to the aerobic threshold (LT5), others to the aerobic/anaerobic transition (e.g. LT1, LT4, Dmax) or the anaerobic threshold (LT-4mmol, Dmax-mod).

Repeatability

Intra-subject CVs over all five measurements were low (3.4-8.1%) and Cronbach's alphas high (0.89-0.97), indicating repeatability of all concepts over the study period of approximately 8 weeks was good. This corresponds well to previous findings of repeatability for power or speed at different lactate concepts, both in terms of CV, determined at 1.3-5.9% in a meta-analytic review,[30] and in terms of Pearson correlations 0.88-0.96[31,32] or ICC 0.98 - 0.99.[33] One study applied different LT concepts to the same dataset from two exercise tests and showed that intra-subject CVs and correlation was good for LT2, LT-4mmol and a concept similar to LT4 (CV 3-4% and $r \geq 0.85$), but not for Dmax (10.3% and 0.57).[34] Our data, based on more subjects (24 versus 14) and more measurements per subject (5 versus 2), disputes this relatively poor repeatability for Dmax. However, our study does show differences between concepts, with LT3, LT-4mmol, Dmax and Dmax-mod having the lowest intra-subject CV (<5%) and the highest Cronbach's alpha (>0.95).

Correlation with performance

As we have established that CV and repeatability for all LT concepts was good, the most relevant question is whether these concepts correlate to actual endurance performance. As previously indicated, it is highly unlikely that the rHuEPO treatment impacted this particular correlation analysis. When analysing the groups separately, some differences in correlation coefficients could be observed between the two groups (data not shown), but these differences were already present for the correlation between GXT1 and TT1 when treatment had not yet started, indicating that this was not due to rHuEPO treatment. Because combining all subjects generates more informative and robust results being based on a bigger population, pooling the groups was considered justified.

Data in Table 4 show that for all concepts correlations with time trial tests were higher compared to the road race (based on all subjects median of all concepts r=0.875 for TT1 and 0.82 for TT2, *versus* 0.61 for RR). This is most likely partly due to additional variability in the road race due to the circumstances (e.g. weather, uphill racing with changes in steepness over the course, and differences in race duration (range 72-126 minutes)). Possibly there was also a minor impact of using different equipment for power measurement during the RR, as it was not measured on the ergometer but on the subjects' bike. What can also be seen is that correlation of the LT concepts with TT1 in general is slightly higher than with TT2. More importantly however, correlation for both time trials show that the ranking among different concepts is very similar, confirming the results are robust. It seems that in general,

using a technique of interpolation for the BLC has superior performance, as LT concepts that were based on the third-order polynomial derived from the individual lactate concentration measurements (LT3, LT5, LT-4mmol, Dmax, Dmax-mod) performed better than the ones that used actual measured bLa values without interpolation (LT1, LT2 and LT4), with the exception of LT5. This poor performance of LT5 is most likely due to the fact that it is conceptually different from the other concepts; it is the power at the minimum lactate equivalent, in this case the lowest value for the lactate-VO_2 ratio. In contrast, LT3 also uses a form of the minimum lactate equivalent, but it adds 1.5 mmol/L to this value. As can be seen in Table 2 and Figure 2, this leads to LT5 on average determining a point even before the first rise in lactate concentration as determined by LT1. This concept therefore relates to much lower (aerobic) work intensities than the other concepts. Additionally it is less repeatable (see Table 3). From all tested concepts LT5 correlates least with 45 min TT performance, but for the longer RR performance relative to the other concepts it performs somewhat better than for TT. This could mean that is this concept is more related to long-term exercise efforts.

Many studies previously evaluated correlations of LT concepts with endurance performance, of which most used running performance. An overview of these studies by Faude *et al* shows a median r = 0.84-0.92 for several different LT concepts for endurance distances (>5km),[9] comparable to our results. There are fewer studies that have compared LT concepts and their correlation with different types of cycling endurance performance,[16,20,26,35–38] but correlation with endurance performances (30-90 minutes) for each concept seem to vary between these studies, see Table 5. In addition, the comparison between concepts within these studies shows varying conclusions about which is the best concept. This could partially be due to differences between studies, for example study populations differ (mean VO_{2max} ranges from 48 to 68 mL/min/kg, and some studying female, others male cyclists and/or triathletes). However, they are more or less as heterogeneous as our population with an SD of 4-8 mL/min/kg on VO_{2max}. The applied exercise protocols all used long stages similar to ours (3-5 minutes), although the increases in workload differ (20-50 W). Finally, correlation to endurance exercise was based on time trials that lasted between approximately 30 to 90 minutes (our TT of 45 min at the lower end and RR of on average 98 min at the higher end), a difference that might impact the correlation to different LT concepts. Nevertheless, taking these differences into account, comparison is possible, albeit with some caution. Moreover, a robust and valid LT concept should perform well in any of these datasets. What can be observed is that all these concepts except LT1, Dmax and Dmax-mod have shown correlations below 0.75, and that in all four direct comparisons that evaluated both Dmax and LT-4mmol, Dmax showed a higher correlation. This latter finding could be due to the fact that LT-4mmol is less robust to changes in settings such as exercise protocol duration, sampling site and lactate analyser because of its fixed nature. Our study expands on this information, and compared to previous studies as reviewed in Table 5, is based on approximately 2-4 times more subjects, therefore allowing for more robust conclusions. This is especially true since our population is a heterogeneous well-trained, and therefore relevant, group (range maximal power output at baseline 256 – 425 W). Similar to what can be extracted from the literature, our study too shows that Dmax and Dmax-mod have highest correlations with time trial performance, although LT-4mmol and LT3 show a similarly high correlation in our study. For the correlation with RR, there are slightly larger differences between concepts. Correlation is highest for Dmax and Dmax-mod, mainly because for the other concepts correlation for a few subjects is very poor, as visualised in Figure 4 (e.g. for LT-4mmol). These findings combined, we conclude that Dmax, and even more so Dmax-mod, have the best correlation with endurance performance. One recent study evaluated correlation between MLSS, which could be considered to be the gold standard for the physiological endurance threshold, and different LT concepts generated from GXTs with different protocol durations.[26] This study concluded that for a GXT with 4-minute steps (most similar to our GXT), correlation was high for many of the concepts, but validity was highest for LT-2.5mmol, Dmax-mod, and two modified versions of Dmax-mod. In contrast, LT2, LT-4mmol and Dmax showed much higher mean differences with MLSS and therefore were designated as invalid estimates of MLSS. Combining these findings with our own results, Dmax-mod determined in a GXT with approximately 5-minute stages is both a valid estimate of MLSS and has a high correlation with actual endurance performance.

Absolute power difference

The mean difference of each concept with the endurance parameter gives an indication of how the absolute power of the LT concept corresponds to the average

power produced during TT and RR. On average, power is higher compared to the endurance test for each concept (except the poorest performing concept LT5). This difference in power between LT concepts and endurance test is possibly due to having to sustain the power for a much longer time during the endurance tests, needing a systematic lower power in order to cope with the effort. Interestingly, Dmax-mod and LT-4mmol, concepts that show among the highest correlations, have the largest difference in absolute power (approximately 30 W). Given the high correlation with performance this should not disqualify these concepts, but one should take into account that there is a systematic difference with endurance performance of approximately 30 W.

CONCLUSIONS

LT concepts are correlated with endurance performance, but a review showed that many different concepts are used in literature, which is undesirable.[9] Also for cycling performance, there is no consensus on which LT concept should be applied and results vary highly.[16,20,35–38] In this study we compared eight different representative LT concepts on the same large cycling performance dataset to evaluate repeatability and predictive properties. All concepts showed high repeatability, and correlated with endurance performance. However, LT3, LT-4mmol, Dmax and Dmax-mod showed the best repeatability, and had the highest correlation with time trial performance. As correlation with performance was consistently high for Dmax and Dmax-mod, also with the uphill road race, the latter performing slightly better on each criterion, and because Dmax-mod was previously shown to be a valid estimate of MLSS, we would recommend using Dmax-mod when analysing the blood lactate curve.

REFERENCES

1 Cairns SP. Lactic acid and exercise performance : culprit or friend? *Sports Med*. 2006;36(4):279-291.
2 Atkinson G, Davison R, Jeukendrup A, Passfield L. Science and cycling: current knowledge and future directions for research. *J Sports Sci*. 2003;21(9):767-787.
3 Kindermann W, Simon G, Keul J. The significance of the aerobic-anaerobic transition for the determination of work load intensities during endurance training. *Eur J Appl Physiol Occup Physiol*. 1979;42(1):25-34.
4 Londeree BR. Effect of training on lactate/ventilatory thresholds: a meta-analysis. *Med Sci Sports Exerc*. 1997;29(6):837-843.
5 Antonutto G, Di Prampero PE. The concept of lactate threshold. A short review. *J Sports Med Phys Fitness*. 1995;35(1):6-12.
6 Robergs RA, Ghiasvand F, Parker D. Biochemistry of exercise-induced metabolic acidosis. *Am J Physiol Regul Integr Comp Physiol*. 2004;287(3):R502-516.
7 MacRae HS, Dennis SC, Bosch AN, Noakes TD. Effects of training on lactate production and removal during progressive exercise in humans. *J Appl Physiol (1985)*. 1992;72(5):1649-1656.
8 Stanley WC, Gertz EW, Wisneski JA, Neese RA, Morris DL, Brooks GA. Lactate extraction during net lactate release in legs of humans during exercise. *J Appl Physiol (1985)*. 1986;60(4):1116-1120.
9 Faude O, Kindermann W, Meyer T. Lactate threshold concepts: how valid are they? *Sports Med*. 2009;39(6):469-490.
10 Coyle EF, Coggan AR, Hopper MK, Walters TJ. Determinants of endurance in well-trained cyclists. *J Appl Physiol (1985)*. 1988;64(6):2622-2630.
11 Billat VL, Sirvent P, Py G, Koralsztein JP, Mercier J. The concept of maximal lactate steady state: a bridge between biochemistry, physiology and sport science. *Sports Med*. 2003;33(6):407-426.
12 Heuberger J, Rotmans JI, Gal P, et al. Effects of erythropoietin on cycling performance of well trained cyclists: a double-blind, randomised, placebo-controlled trial. *Lancet Haematol*. 2017;4(8):e374-e386.
13 Tanaka H. Predicting running velocity at blood lactate threshold from running performance tests in adolescent boys. *Eur J Appl Physiol Occup Physiol*. 1986;55(4):344-348.
14 Coyle EF, Martin WH, Ehsani AA, et al. Blood lactate threshold in some well-trained ischemic heart disease patients. *J Appl Physiol Respir Environ Exerc Physiol*. 1983;54(1):18-23.
15 Dickhuth H-H, Yin L, Niess A, et al. Ventilatory, lactate-derived and catecholamine thresholds during incremental treadmill running: relationship and reproducibility. *Int J Sports Med*. 1999;20(2):122-127.
16 Amann M, Subudhi AW, Foster C. Predictive validity of ventilatory and lactate thresholds for cycling time trial performance. *Scand J Med Sci Sports*. 2006;16(1):27-34.
17 Dickhuth H.-H.; Huonker M. MT, Drexler H., Berg A., Keul J. Individual anaerobic threshold for evaluation of competitive athletes and patients with left ventricular dysfunctions. *Advances in ergometry*. 1991.
18 Sjodin B, Jacobs I. Onset of blood lactate accumulation and marathon running performance. *Int J Sports Med*. 1981;2(1):23-26.
19 Cheng B, Kuipers H, Snyder AC, Keizer HA, Jeukendrup A, Hesselink M. A new approach for the determination of ventilatory and lactate thresholds. *Int J Sports Med*. 1992;13(7):518-522.
20 Bishop D, Jenkins DG, Mackinnon LT. The relationship between plasma lactate parameters, Wpeak and 1-h cycling performance in women. *Med Sci Sports Exerc*. 1998;30(8):1270-1275.
21 Chambers J. Project R. 3.4.0:https://www.r-project.org/.
22 Hadley Wickham RF, Lionel Henry, Kirill Müller. dplyr: A Grammar of Data Manipulation. 2017; https://cran.r-project.org/web/packages/dplyr/.
23 Revelle W. psych: Procedures for Psychological, Psychometric, and Personality Research. 2017; https://cran.r-project.org/web/packages/psych/.
24 Hadley Wickham LH. tidyr: Easily Tidy Data with 'spread()' and 'gather()' Functions. 2017; https://cran.r-project.org/web/packages/tidyr/.
25 Bentley DJ, Newell J, Bishop D. Incremental exercise test design and analysis: implications for performance diagnostics in endurance athletes. *Sports Med*. 2007;37(7):575-586.
26 Jamnick NA, Botella J, Pyne DB, Bishop DJ. Manipulating graded exercise test variables affects the validity of the lactate threshold and [Formula: see text]. *PLoS One*. 2018;13(7):e0199794.
27 Thoden JS. Testing aerobic power. In: MacDougall JD, Wenger HA, Green HJ, eds. *Physiological testing of the high-performance athlete*. Champaign: Human Kinetics; 1991:107-174.
28 San Millan I, Bing K, Brill C, Hill JC, Miller LE. Randomized controlled trial of Micro-Mobile Compression(R) on lactate clearance and subsequent exercise performance in elite male cyclists. *OpenAccess J Sports Med*. 2013;4:221-227.
29 Bentley DJ, McNaughton LR. Comparison of W(peak), VO2(peak) and the ventilation threshold from two different incremental exercise tests: relationship to endurance performance. *J Sci Med Sport*. 2003;6(4):422-435.
30 Hopkins WG, Schabort EJ, Hawley JA. Reliability of power in physical performance tests. *Sports Med*. 2001;31(3):211-234.
31 Weltman A, Snead D, Stein P, et al. Reliability and validity of a continuous incremental treadmill protocol for the determination of lactate threshold, fixed blood lactate concentrations, and VO2max. *Int J Sports Med*. 1990;11(1):26-32.
32 Grant S, McMillan K, Newell J, et al. Reproducibility of the blood lactate threshold, 4 mmol.l(-1) marker, heart rate and ratings of perceived exertion during incremental treadmill exercise in humans. *Eur J Appl Physiol*. 2002;87(2):159-166.
33 Pfitzinger P, Freedson PS. The reliability of lactate measurements during exercise. *Int J Sports Med*. 1998;19(5):349-357.
34 Pallares JG, Moran-Navarro R, Ortega JF, Fernandez-Elias VE, Mora-Rodriguez R. Validity and Reliability of Ventilatory and Blood Lactate Thresholds in Well-Trained Cyclists. *PLoS One*. 2016;11(9):e0163389.
35 Borszcz FK, Tramontin AF, de Souza KM, Carminatti LJ, Costa VP. Physiological Correlations With Short, Medium, and Long Cycling Time-Trial Performance. *Research quarterly for exercise and sport*. 2018;89(1):120-125.
36 Nichols JF, Phares SL, Buono MJ. Relationship between blood lactate response to exercise and endurance performance in competitive female master cyclists. *Int J Sports Med*. 1997;18(6):458-463.

37 Bentley DJ, McNaughton LR, Thompson D, Vleck VE, Batterham AM. Peak power output, the lactate threshold, and time trial performance in cyclists. *Med Sci Sports Exerc.* 2001;33(12):2077-2081.
38 McNaughton LR, Roberts S, Bentley DJ. The relationship among peak power output, lactate threshold, and short-distance cycling performance: effects of incremental exercise test design. *J Strength Cond Res.* 2006;20(1):157-161.

FIGURE 1 STUDY DESIGN Study design showing timing of different tests. Time point 0 weeks indicates start of treatment (rHuEPO or placebo) for all participants. GXT, graded exercise test; TT, time trial test; RR, road race.

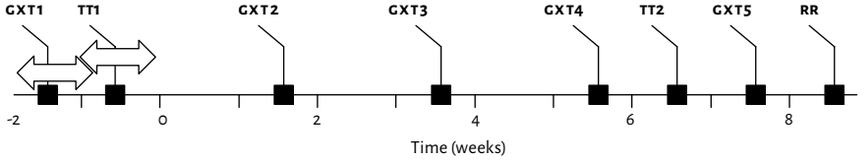


FIGURE 2 GRAPHICAL REPRESENTATION OF LACTATE THRESHOLD CONCEPTS Example of a blood lactate curve with the location of the different lactate threshold concepts for this particular curve. Open circles: observed blood lactate values at each exercise intensity; Black curve: third-order polynomial; Black dashed line: baseline; Light grey circle and arrow: LT1, observer-determined first rise in blood lactate; Medium light grey circle and arrow: LT2, first observed blood lactate value more than 1 mmol/L above baseline; Ligtest grey circle and arrow: LT3, minimum lactate equivalent (blood lactate divided by power) plus 1.5 mmol/L; Medium dark grey circle and arrow: LT4, first blood lactate value that shows an increase of at least 1 mmol/L; Grey circle and arrow: LT5, minimum lactate equivalent (blood lactate divided by VO_2); Light grey circle and arrow and dashed line: LT-4 mmol, value at 4 mmol/L; Darkest grey circle and arrow and dashed line: Dmax, value with the maximum perpendicular distance to the polynomial from the dashed line; Dark grey circle and arrow and dashed line: Dmax-mod, value with the maximum perpendicular distance to the polynomial from the dashed line.

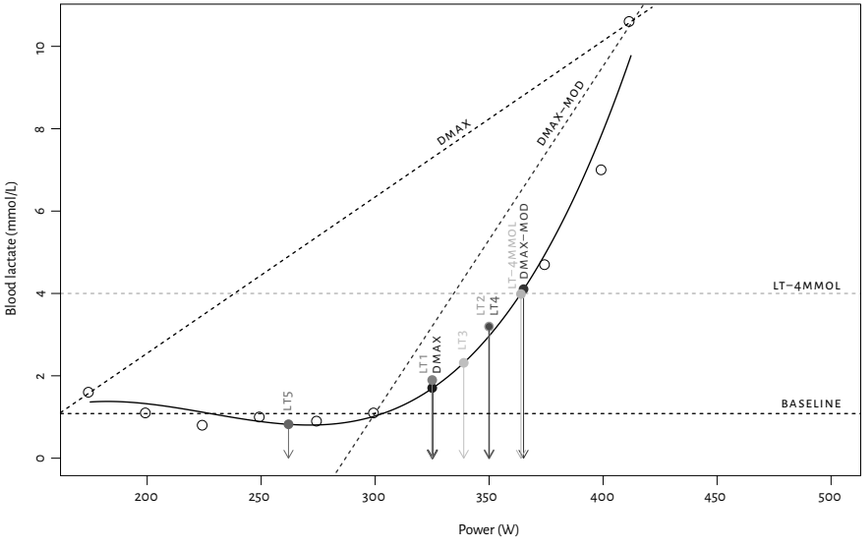


FIGURE 3 CONSORT FLOWCHART

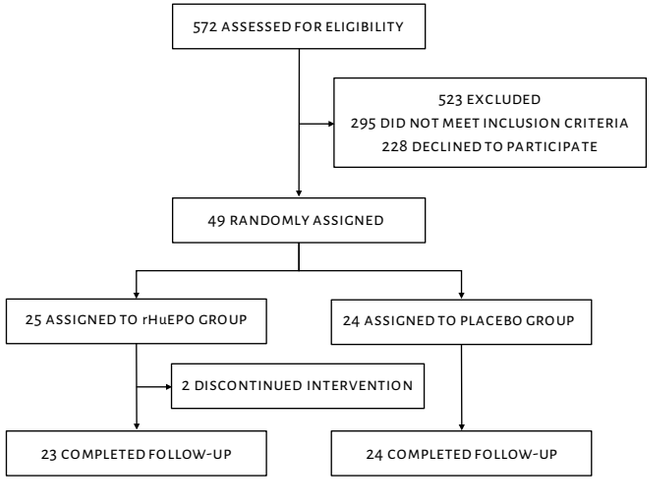


FIGURE 4 LINEAR REGRESSION LACTATE THRESHOLD CONCEPT POWER AND AVERAGE RACE POWER Linear regression of lactate threshold power and average race power per LT concept for all subjects depicting linear regression line (solid line) and 95% confidence interval (dotted lines). r^2: R-squared or coefficient of determination is the proportion of the variance in the dependent variable that is predictable from the independent variable.

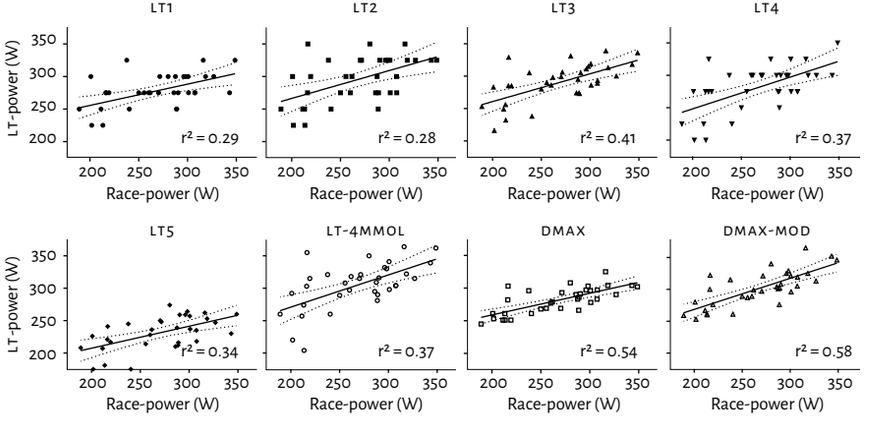


TABLE 1 SUBJECT DISPOSITION Values are presented as mean (standard deviation (SD) where appropriate; range). VO_{2max}: maximal oxygen consumption.

	All subjects	Placebo subjects
N	48	24
Age (years)	33.6 (20.0-50.0)	33.8 (20.0-50.0)
Weight (kg)	76.9 (9.0; 59.2-95.6)	76.9 (8.9; 59.2-95.6)
Height (cm)	186 (7.3; 172-203)	186 (6.7; 174-203)
Maximal Power output per kg (W/kg)	4.36 (4.03-5.18)	4.36 (4.03-4.94)
VO_{2max} (mL/min/kg)	55.7 (4.6; 45.3-67.5)	56.0 (4.1; 47.0-62.8)

TABLE 2 MEAN LACTATE THRESHOLD CONCEPT POWER OUTPUT Weighted mean power output (SD; range) for the placebo group at every exercise test. Overall combined power output (based on 109 GXTS) for each lactate threshold concept (CV). CV is weighted intra-subject CV.

GXT number	LT1 (W)	LT2 (W)	LT3 (W)	LT4 (W)	LT5 (W)	LT-4MMOL (W)	Dmax (W)	Dmax-mod (W)
1	283.3 (29.9; 225-350)	292.9 (37.2; 250-375)	286.1 (32.9; 219-352)	275.0 (41.1; 200-350)	225.0 (31.2; 175-283)	301.8 (41.0; 222-381)	275.7 (24.6; 222-323)	299.5 (35.3; 225-367)
2	283.0 (22.3; 225-375)	293.2 (31.0; 250-375)	288.7 (29.2; 231-373)	276.1 (34.0; 175-375)	231.8 (25.7; 175-300)	305.0 (33.0; 234-389)	280.0 (23.6; 233-339)	301.2 (28.7; 237-369)
3	281.0 (29.5; 225-400)	290.5 (27.9; 250-400)	285.7 (26.3; 240-390)	272.6 (33.5; 225-375)	224.5 (29.5; 175-318)	300.8 (28.9; 253-411)	278.7 (20.9; 250-343)	297.5 (29.6; 257-413)
4	283.7 (35.8; 225-400)	292.4 (38.0; 225-425)	291.6 (29.1; 240-392)	272.8 (36.9; 200-400)	229.8 (29.5; 175-323)	307.0 (34.2; 249-415)	284.0 (22.7; 232-338)	308.7 (35.6; 251-396)
5	278.3 (28.5; 225-325)	290.2 (37.5; 225-350)	285.0 (31.6; 216-339)	271.7 (37.9; 200-350)	230.4 (30.5; 175-274)	297.2 (38.9; 204-364)	280.3 (20.3; 245-325)	298.9 (29.2; 253-365)
Overall	282.1 (5.7%)	292.2 (5.0%)	287.7 (3.6%)	274.1 (5.6%)	228.4 (8.1%)	302.7 (3.8%)	280.0 (3.4%)	301.5 (4.3%)

TABLE 3 CRONBACH'S ALPHA FOR EACH LACTATE THRESHOLD CONCEPT Cronbach's alpha for the placebo group for each lactate threshold concept with 95% confidence interval (CI).

Lactate threshold concept	Cronbach's alpha	Lower 95% CI	Upper 95% CI
LT1	0.91	0.85	0.96
LT2	0.95	0.92	0.98
LT3	0.97	0.94	0.99
LT4	0.94	0.91	0.98
LT5	0.89	0.82	0.96
LT 4-mmol	0.96	0.94	0.99
Dmax	0.96	0.93	0.98
Dmax-mod	0.96	0.94	0.98

TABLE 4 PREDICTIVE VALUE OF LACTATE THRESHOLD CONCEPTS Pearson correlation between each lactate threshold concept in GXT1 and time trial test 1 (TT1), GXT4 and time trial test 2 (TT2) and GXT5 and average road race (RR) power for all subjects combined. All correlations are significant ($p<0.0002$). To determine potential differences in power output between the LT concept and time trial power or race power, mean difference (SD) between each measurement pair is calculated. Negative values indicate lactate threshold power is higher than exercise test average power.

Lactate threshold concept	Pearson correlation			Mean difference (SD)		
	TT1	TT2	RR	TT1	TT2	RR
	n=42	*n*=46[a]	*n*=34[b]	*n*=42	*n*=46[a]	*n*=34[b]
LT1	0.78	0.74	0.54	-11.3 (18.3)	-8.7 (22.1)	-9.8 (37.2)
LT2	0.87	0.80	0.53	-23.2 (16.0)	-18.5 (19.3)	-27.4 (39.3)
LT3	0.88	0.84	0.64	-16.2 (14.3)	-16.1 (18.1)	-21.9 (33.8)
LT4	0.78	0.82	0.61	-3.5 (23.7)	-0.6 (26.2)	-13.5 (36.1)
LT5	0.67	0.65	0.58	43.7 (21.7)	45.3 (23.5)	39.1 (35.9)
LT-4MMOL	0.88	0.85	0.61	-31.7 (19.4)	-32.3 (23.0)	-36.6 (36.1)
Dmax	0.89	0.82	0.73	-4.4 (12.1)	-3.8 (15.8)	-13.4 (32.4)
Dmax-mod	0.94	0.84	0.76	-27.3 (11.8)	-29.9 (16.6)	-33.7 (29.1)

a for LT5 *n* = 44; *b for* LT5 *n* = 32

TABLE 5 REPORTED CORRELATIONS BETWEEN LT CONCEPTS AND ENDURANCE PERFORMANCE Literature data for LT concepts and correlation with 30–90 minute-during performances. LTlog: the power output at which bLa starts to increase when log(bLa) is plotted against log(power output).

Correlation reported in publication	Lactate threshold concept						
	LT1	LT2	LT4	LT-4MMOL	Dmax	Dmax-mod	LTlog
Amann[16]	-	0.72	0.59	0.60	-	-	-
Bentley[37]	-	-	-	0.54	0.77	-	0.91
Bishop[20]	0.81	-	0.61	0.81	0.84	0.83	0.69
Borszcz[35]	-	0.31	-	0.56	0.75	-	-
McNaughton[38]	-	-	-	0.90	0.91	-	0.86
Nichols[36]	-	-	0.88	0.67	-	-	-

6
ADDITIVE EFFECT OF ERYTHROPOIETIN USE ON EXERCISE-INDUCED ENDOTHELIAL ACTIVATION AND HYPERCOAGULABILITY IN ATHLETES

SUBMITTED IN REVISED FORM: EUROPEAN JOURNAL OF APPLIED PHYSIOLOGY

Jules AAC Heuberger, Jelle J Posthuma, Dimitrios Ziagkos, Joris I Rotmans, Johannes MA Daniels, Pim Gal, Frederik E Stuurman, Henri MH Spronk, Hugo ten Cate, Jacobus Burggraaf, Matthijs Moerland and Adam F Cohen

7
SENSITIVITY AND SPECIFICITY OF DETECTION METHODS FOR ERYTHROPOIETIN DOPING IN CYCLISTS

SUBMITTED IN REVISED FORM: DRUG TESTING AND ANALYSIS

Jules AAC Heuberger, Peter van Eenoo, Joris I Rotmans, Pim Gal, Frederik E Stuurman, Titiaan E Post, Johannes MA Daniels, Herman Ram, Olivier de Hon, Jacobus Burggraaf and Adam F Cohen

ABSTRACT

Recombinant human erythropoietin (rHuEPO) is used as doping a substance. Anti-doping efforts include urine and blood testing and monitoring the Athlete Biological Passport (ABP). Data on the performance of these methods are incomplete, therefore this study aimed to evaluate the performance of two common urine assays and the ABP. In a randomised, double-blinded, placebo-controlled trial, 48 trained cyclists received a mean dose of 6000 IU rHuEPO (epoetin-beta) or placebo by weekly injection for 8 weeks. Seven timed urine and blood samples were collected per subject. Urine samples were analysed by sarcosyl-PAGE and isoelectric focusing methods in the accredited DoCoLab in Ghent. A selection of samples, including any with false presumptive findings, underwent a second sarcosyl-PAGE confirmation analysis. Haematological parameters were used to construct a module similar to the ABP and analysed by two evaluators from an Athlete Passport Management Unit. Sensitivity of the sarcosyl-PAGE and isoelectric focusing assays for the detection of erythropoietin abuse were 63.8% and 58.6%, respectively, with a false presumptive finding rate of 4.3% and 6%. None of the false presumptive findings tested positive in the confirmation analysis. Sensitivity was highest between 2-6 days after dosing, and dropped rapidly outside this window. Sensitivity of the ABP was 91.3%. Specificity of the urine assays was high, however, the detection window of rHuEPO was narrow, leading to questionable sensitivity. The ABP, integrating longitudinal data, is more sensitive, but there are still subjects that evade detection. Combining these methods might improve performance, but will not resolve all observed shortcomings.

INTRODUCTION

Athletes have used drugs to enhance sports performance throughout history, despite all efforts to ban these doping substances. To detect cheating athletes, the World Anti-Doping Agency (WADA) oversees a global doping control system, using blood and urine samples that are collected both in competition and out of competition. If an athlete is tested positive, this will usually lead to a suspension of two or four years, depending on the substance that is found and the circumstances of the case, up to lifetime ineligibility for recidivists.[1] One substance that is widely abused, especially in endurance sports, is recombinant human erythropoietin (rHuEPO). rHuEPO has shown to positively affect maximal performance (e.g. [2,3]) and although there is a lack of clear scientific proof whether it also improves actual endurance performance,[4] the belief in erythropoietin's effects on performance among athletes, their staff and the general public is overwhelming. Recent confessions to rHuEPO misuse by many cyclists from the 1990-2010 era, including former champions like Lance Armstrong, unfolded a rHuEPO epidemic, at least within cycling. Testing for rHuEPO use was introduced in 2000, and results from the testing figures published by WADA seem to indicate that rHuEPO is still used by athletes.[5] In the period 2012-2017,[5] a total of 200.451 urine samples were tested for presence of erythropoiesis stimulating agents (ESAs) and 305 samples (0.15%) produced adverse analytical findings (AAFs). In addition, 12.966 blood samples were tested, of which 61 samples (0.44%) produced AAFs. These AAFs are interpreted as proof an athlete used an ESA. In addition to these specific tests for rHuEPO, the WADA introduced the Athlete Biological Passport (ABP) in 2009, in which several biomarkers of an individual are measured longitudinally. The haematological module containing haematological markers is applied to monitor use of erythropoiesis stimulating agents such as rHuEPO. Abnormalities in these markers can be used to start targeted testing in that individual, or even as a doping violation by itself.

Although doping tests are intended to discourage athletes from using doping, detect the ones that are violating doping rules, and protect the athletes that play fair, their ability to do so completely depends on the performance of the tests and an intelligent testing program collecting samples within the window of detection. While technical documents describing the urine assays for detecting ESA use are available through WADA[6] and articles on the methods are published,[7–11] the critical

assay characteristics of sensitivity and specificity and the detection window, based on well-designed blinded studies, are not sufficiently (publicly) available.[12] We believe that scientific evidence underpinning internationally and widely applied doping tests should be available in order to understand the value of the tests, protect fair playing athletes and collaborate on improving doping detection. For that reason, we aimed to evaluate the performance of three main methods being used in rHuEPO detection by determining sensitivity and specificity of the sarcosyl-PAGE and Isoelectric Focusing (IEF) assays on urine samples, and of the ABP and its haematological module.

METHODS

Study design and participants

For the purpose of this study, blood and urine samples were used that were obtained in a double-blind, randomised, placebo controlled, parallel, single centre study on rHuEPO, in which forty-eight healthy male cyclists between 18 and 50 years of age were included. The design of this trial was described elsewhere.[4] In short, the main inclusion criteria were maximum power-to-weight ratio during the maximal exercise test at screening exceeding 4 W/kg, Hb between 8.0 and 9.8 mmol/L (equivalent to 12.8 – 15.7 g/dL) and Ht below 48% at screening and not being subject to anti-doping regulation or using medication that could potentially interact with the study drugs or study assessments. All participants gave written informed consent prior to any study-related activity. The study was approved by the Independent Ethics Committee of the Foundation 'Evaluation of Ethics in Biomedical Research' (Stichting Beoordeling Ethiek Biomedisch Onderzoek), Assen, NL. The study is registered in the Dutch Trial Registry (Nederlands Trial Register, NTR) under study number NTR5643.

Randomisation and masking

Randomisation to the rHuEPO or the placebo group (1:1) was done by a randomisation code generated by an un-blinded statistician who was otherwise not involved in the execution of the study. Enrolment of subjects was performed by a blinded study physician.

Procedures

Treatment

Epoetin-beta (NeoRecormon, Roche, Basel, Switzerland), prepared from multidose vials containing a lyophilisate of 50 000 IU epoetin-beta and 10 mL solvent for solution for injection, or saline (0.9% NaCl) was administered subcutaneously with weekly abdominal injections for 8 weeks. Dosing aimed to reach a target Hb of 10-15% increase compared to the baseline Hb value, similar to previous studies investigating effects of rHuEPO on performance.[13] The first four injections contained 5000 IU per injection and the dose was modulated by an un-blinded, non-study related physician to 6000 IU, 8000 IU or 10 000 IU in the final four weeks in case the Hb level was below the target range, doses similar to known practices in professional cycling.[14] When the Hb was in the target range 2000 IU was administered. If the Ht was equal to or exceeded 52% or if the Hb exceeded the upper limit of the Hb range (increase of 15% compared to baseline), a placebo injection was administered.

All participants also received open-label daily oral doses of 200 mg ferrous fumarate and 50 mg ascorbic acid (both Pharmachemie BV, Haarlem, NL).

Urine and blood samples for doping detection

Urine samples were collected for each subject and 50 mL was stored in Falcon tubes (Greiner Bio-One International GmbH, Austria) at -70 °C. Samples were collected at approximately 1 hour, 2 days, 4 days and 7 days after the second dose and at approximately 12 days after the last dose. These samples were taken while subjects had not exercised in the preceding hours. To evaluate potential effects of exercise on the outcome of the test, a sample was taken as soon as the subject could urinate after the maximal exercise at approximately 4 days after the second dose and after a race to the top of Mont Ventoux at approximately 12 days after the last dose. Samples were shipped to the WADA accredited anti-doping laboratory DoCoLab UGent by courier in two separate batches. The lab was aware these samples were research samples, but was blinded to the administered treatment.

All haematology samples were drawn after participants were seated with their feet on the floor for at least ten minutes, at the following time points: 1-6 weeks before the first dose, directly before the first, third, fifth and seventh dose, approximately 12 days and 4 weeks after the last dose. In general this means there

were 2 weeks between two samples. Samples were analysed within 2 hours at the Leiden University Medical Centre (LUMC) on a Sysmex XN9000 analyser (Sysmex Corporation, Kobe, Japan). Samples taken at the Mont Ventoux race were split into two aliquots, of which one was kept at room temperature for determination of leucocytes and leucocytes differentiation and one was kept between 2-8 °C for determination of haemoglobin, haematocrit, erythrocytes, reticulocytes, MCV, MCH, MCHC, thrombocytes, immature reticulocyte fraction (IRF) and red blood cell distribution width (RDW-SD). Samples were driven by courier to the LUMC and analysis took place within 36 hours of collection.

Sarcosyl-PAGE urine assay

The sarcosyl-PAGE urine assay was used as described in the WADA technical document.[6] In short, immunopurification of the sample was performed with antibodies other than the one used for immunoblotting before loading 15 µL of sample on pre-cast polyacrylamide gels (NuPAGE BisTris gels, 10% T, 1.5 mm, as described previously[15]), with 15 slots for test samples per gel. For purification, Stemcell ELISA was used as described previously.[11] In addition, a negative control sample, positive control sample containing rHuEPO, and a reference to enable to define apparent molecular mass were run in separate slots. Running buffer used was sodium N-lauroyl sarcosinate, and after the electrophoretic separation double-blotting was performed with the monoclonal mouse anti-human EPO antibody clone AE7A5 and a secondary antibody. The secondary antibody was a goat anti-mouse polyclonal antibody IgG (H+L), cross-adsorbed, HRP (Thermo scientific product number 31432) and used in combination with a streptavidin horseradish peroxidase complex (Biospa, Milano, Italy) and a substrate (SuperSignal West Femto Maximum Sensitivity; Thermo Scientific, Waltham, MA, USA). The electrophoretic patterns of ESAs were revealed by the use of an amplified chemiluminescent system. Detection of exogenous EPO, in this case epoetin-beta, was done on the basis of the characteristic band shape ("broad band") and different (higher) apparent molecular mass than endogenous EPO. A sample was termed positive if a single band was visible above the negative control EPO, or if a mixed band both at the endogenous and exogenous location or a diffuse or faint area of the band above the corresponding endogenous band were visible, as per WADA technical document.[6]

Isoelectric focusing urine assay

The IEF urine assay was used as described in the WADA technical document.[6] In short, immunopurification of the sample was performed with antibodies other than the one used for immunoblotting before loading 20 µL of sample on the gel, with 30 slots for test samples per gel. For purification, Stemcell ELISA was used as described previously.[11] IEF was performed in a pH range compatible with the isoelectric points (pI) of both the natural EPO and its recombinant analogues. IEF was performed under denaturing conditions (approximately 7M urea). After the electrophoretic separation double-blotting was performed with the monoclonal mouse anti-human EPO antibody clone AE7A5 and a secondary antibody. The secondary antibody was a goat anti-mouse polyclonal antibody IgG (H+L), cross-adsorbed, HRP (Thermo scientific product number 31432) and used in combination with a streptavidin horseradish peroxidase complex (Biospa, Milano, Italy) and a substrate (SuperSignal West Femto Maximum Sensitivity; Thermo Scientific, Waltham, MA, USA). The electrophoretic patterns of ESAs were revealed by the use of an amplified chemiluminescent system. Detection of exogenous EPO, in this case epoetin-beta, was done on the basis of the relative location of the bands compared to the position of the bands corresponding to the rHuEPO Biological Reference Preparation (BRP) of the European Pharmacopeia (equimolar mixture of epoetin-alfa and -beta), which define the basic and acidic areas. A sample was termed positive if there were at least 3 acceptable, consecutive bands in the basic area, and if the 2 most intense bands measured by densitometry were in the basic area, as per WADA technical document.[6] It should be noted that the isoelectric focusing assay is currently not used in routine practice by the DoCoLab and therefore the associated analyses are exploratory.

Urine sample evaluation (screening)

Samples were designated as screening negative, screening positive or non-detectable by the DoCoLab in Ghent. Samples were non-detectable if no band was detected in the EPO (endogenous or exogenous) designated region.

Confirmation analysis

After results of the screening analysis were available, the outcome was analysed by unblinded Centre for Human Drug Research (CHDR) researchers. All samples

from placebo treated subjects that generated a positive result in the screening analysis (here termed false presumptive findings) with the sarcosyl-PAGE assay were re-analysed with the sarcosyl-PAGE assay. Both the initial and confirmation methodology were the same. In order to keep this confirmation analysis blinded to the DoCoLab, true positive, true negative and false negative samples were included as well, and this selection of samples was relabelled by CHDR staff for analysis. Samples were only deemed positive, similar as for routine-analysis, when they would be forwarded for a mandatory second opinion. The WADA Technical Document requires that all presumptive adverse analytical findings are evaluated by a recognised expert. After evaluation by the laboratory performing the analysis, samples are therefore forwarded to a WADA recognised expert. Currently, these experts are co-authors of the WADA Technical Document and have extensive scientific knowledge on ESAs, their detection methodology and potential issues in the detection of ESAs. In the framework of this study, samples were not subjected to a second opinion and the interpretation was solely based upon the laboratory performing the analysis. No WADA recognised expert as per the Technical Document was involved. In case of doubt, samples were deemed negative.

Osmolality

Osmolality was determined by taking an aliquot from the homogenised, centrifuged samples, using the freezing point method on an Osmometer Auto & Stat, OM-6050 (Arkray, Kyoto, Japan). Osmolality determination was initiated after all analyses had been performed in order to provide insights into potential effects of osmolality on assay results.

Athlete Biological Passport (ABP)

A method similar to the adaptive model of the ABP was constructed internally using Microsoft Excel 2013 (Redmond, United States), as the official WADA software was not made available for this study. Individual curves of each separate haematological parameter from blood samples were produced by plotting the observed value, and the 99.7% confidence interval (CI) around that value. For the first measurement, the 99.7% CI was based on the population mean (Pop. mean) ± 3xSD of all available data points from the screening and baseline haematological samples of all subjects that completed these visits (Pop. SD); for IRF and RDW-SD only the value at baseline was used as they were not determined at screening. For following data points, with each additional available measurement the 99.7% CI was based more on individual data, and calculated as follows: second data point:

$$\frac{\text{Pop. mean + individual value 1}}{2} \pm \text{SD}[\text{Pop. mean + individual value 1}] + 2 * \text{Pop. SD}$$

third data point:

$$\frac{\text{Pop. mean + individual value 1 and 2}}{3} \pm 2 * \text{SD}[\text{Pop. mean + individual value 1 and 2}] + \text{Pop. SD}$$

fourth data point:

$$\frac{\text{Pop. mean + individual value 1,2 and 3}}{4} \pm 3 * \text{SD}[\text{Pop. mean + individual value 1, 2 and 3}]$$

fifth data point:

$$\frac{\text{Pop. mean + individual value 1, 2, 3 and 4}}{5} \pm 3 * \text{SD}[\text{Pop. mean + individual value 1, 2, 3 and 4}]$$

etc.

The OFF-score (OFFS) was calculated using the following formula: ([haemoglobin concentration in g/dL]) * 10 - (60 * √([Reticulocyte percentage])).[16] Graphs of each separate variable belonging to one subject were combined into one document and sent with a coded (blinded) identifier to the DoCoLab in Ghent for evaluation of a subject being negative, suspicious or positive. Two independent researchers at the DoCoLab, which is designated as an Athlete Passport Management Unit (APMU) by WADA, independently evaluated the data and two sets of results were generated and analysed, so that reproducibility could also be evaluated.

Exercise

Two urine samples per subject were collected after exercise: one after a maximal exercise test to exhaustion at 4 days (±1 day) after the second dose administration and one approximately 12 days after the last dose after participants competitively completed an uphill road race on the Mont Ventoux. This race was directly preceded by 110 km cycling.

Data management

All data were stored in a clinical trial database (Promasys, Omnicomm Inc., Fort Lauderdale, USA) and checked for accuracy and completeness. A blinded data review was performed before code-breaking and analysis according to a standard procedure at our unit.

Statistical analysis

Standard assay characteristics were calculated based on the reported results. For determining overall urine assay characteristics, only samples that were designated 'negative' or 'positive' were included in the analysis. When determining sensitivity over time, two analyses were done, one including and one excluding the samples designated 'non-detectable'. For this analysis, samples were binned per day, and only bins with more than one sample were included in the analysis. When determining ABP assay characteristics, suspicious and positive were combined as being an indication of rHuEPO use.

Role of the funding source

This was an investigator initiated study by the foundation CHDR in Leiden in collaboration with DoCoLab (Ghent, B), Leiden University Medical Centre, the Anti-Doping Authority of The Netherlands and the department of Pulmonary Diseases, VU University Medical Centre, Amsterdam. There was no external funding source.

RESULTS

Forty-nine participants were recruited. One participant withdrew after the first dose administration and was replaced by a reserve participant, whereas another participant withdrew after the fourth dose administration. Both withdrawals were due to personal reasons and not related to the study treatment or medical concerns. All 48 subjects were included in the urine detection analysis. Due to the incomplete ABP profile of the withdrawn subject, 47 subjects were included in the analyses for the ABP, of which 23 in the rHuEPO group and 24 in the placebo group. All participants were living at sea level and did not spend any substantial amount of time at (simulated) high altitude.

Participants in the rHuEPO group received eight dosages of epoetin-beta during the study. Mean rHuEPO dose was 5000 IU per week during the first four weeks of the study and 7000 IU in the subsequent four weeks. On five occasions a placebo injection was intentionally administered to subjects in the rHuEPO group, because of the measured haematocrit or haemoglobin levels. Of all 336 urine samples, 6 were not collected due to withdrawal from the study or subjects not being able to perform the particular visit. One additional sample from a rHuEPO treated subject was not analysed with the IEF assay as it was erroneously not included in the sample shipment at the time of the IEF analysis.

Urine analysis

Sarcosyl-PAGE urine assay

A total of 330 urine samples were analysed for rHuEPO using sarcosyl-PAGE, of which 17 samples (5%) were termed non-detectable. Of these, 14 (82%) belonged to subjects in the rHuEPO group and 3 (17%) to subjects in the placebo group. Based on the remaining 313 samples, the sensitivity of sarcosyl-PAGE after the screening analysis was 63.5% and specificity was 95.7% (with 7 false presumptive findings), see Table 1.

A total of 14 samples were analysed for a second time in the confirmation analysis, including all seven samples producing a false presumptive finding in the screening analysis. All 14 samples were correctly identified as negative in placebo subjects and positive for rHuEPO subjects. The assay characteristics when taking both analyses into account show specificity increased to 100%, see Table 1.

When comparing samples that were collected before and directly after a maximal exercise test on the same day, sensitivity was 100% at 4 days, 92% and 100% at 5 days, 57% and 0% at 11 days, 33% and 0% at 12 days and 0% at 13 days since last dosing.

Sarcosyl-PAGE sensitivity over time

When tested approximately 1-2 hours after the second rHuEPO dosing (and 7 days after the first), 50% of rHuEPO treated subjects tested positive, see Figure 1. The sensitivity increased to approximately 81% when urine was collected 2 days after the last dose, and was 95-100% between day 3 and 6 after the last dose. After 7 days, sensitivity decreased again to approximately 50% and was 29% at 11 days, 18% at 12 days and 0% at 13 and 16 days after the last dose. When samples designated as non-detectable are excluded, sensitivity slightly increases, but is very similar.

Isoelectric Focusing urine assay

A total of 329 urine samples were analysed for rHuEPO using isoelectric focusing, of which 68 samples (21%) could not be declared either positive or negative. Of these, 51 (75%) belonged to subjects in the rHuEPO group and 17 (25%) to subject in the placebo group. Based on the remaining 261 samples, the sensitivity of isoelectric focusing after a single analysis was 58.6% and specificity was 94.0% (with 9 false presumptive findings), see Table 2. For isoelectric focusing no confirmation analysis was performed. The focus of the urine analysis was on the sarcosyl-PAGE assay as that is the standard assay at the DoCoLab in Ghent, whereas the isoelectric focusing assay was only implemented for this study, and for this reason we decided not to perform a confirmation analysis for this method.

When comparing samples that were collected before and directly after a maximal exercise test on the same day, sensitivity was 50% and 89% at 4 days, 33% and 83% at 5 days, 42% and 0% at 11 days, and 0% at for all samples at 12 and 13 days since last dosing.

Isoelectric Focusing sensitivity over time

When tested approximately 1-2 hours after the second rHuEPO dosing (and 7 days after the first), 42% of rHuEPO treated subjects tested positive, see Figure 2. The sensitivity increased to approximately 86% when urine was collected 2 days after the last dose, and was 50-68% between day 3 and 5 after the last dose. After 6 and 7 days, sensitivity decreased again to approximately 25 and 22% and was only 21% at 11 days and 0% at 12, 13 and 16 days after the last dose. When samples designated as non-detectable are excluded, sensitivity increases significantly (Figure 2).

Athlete Biological Passport

A total of 47 ABPs based on the collected blood samples were rated for rHuEPO use by the two independent APMU staff members. Results of the ABP performance by evaluator 1 and evaluator 2 are depicted in Table 3. The sensitivity of the ABP by both evaluators was 91.3% and specificity was 100% for evaluator 1 and 95.8% for evaluator 2. Evaluator 1 classified 14 ABPs as positive, 7 as suspicious and 26 as negative. Evaluator 2 classified 15 ABPs as positive, 7 as suspicious and 25 as negative. Of all 47 ABPs, 39 were scored identically by the evaluators. When the results for suspicious and positive were combined, 44 ABPs were scored identically. One discrepancy between evaluators was that evaluator 2 designated an ABP of a placebo subject as suspicious, which was correctly designated as negative by evaluator 1. Additionally, both evaluators designated a separate ABP as negative while it belonged to a subject from the rHuEPO group. Interestingly, the ABP from one rHuEPO treated subject was designated as negative by both evaluators. This ABP, together with a true positive and a true negative subject's ABP can be seen in Figure 3.

DISCUSSION

The present study evaluated the performance of current doping detection methods for epoetin-beta (Neorecormon) use by athletes. The sarcosyl-PAGE urine assay based on 330 samples was shown to be highly sensitive for the applied dose of 5000 IU within a small time range of two to six days after dosing. Of all non-rHuEPO samples 4.3% was labelled screening positive, but after the standard confirmation analysis no false positives were found. Findings for isoelectric focusing showed a slightly lower sensitivity and slightly higher false presumptive finding rate of 6%. Evaluation of the ABP based on haematology samples had high sensitivity of 91.3% for rHuEPO use. Specificity was also high with 100% and 95.8% for the two evaluators. However, this research brought forward some concerns of all three methods.

Sarcosyl-PAGE urine assay performance

The sarcosyl-PAGE assay performed well two to six days after dosing, but outside this window sensitivity falls rapidly. Given the long-lasting haematological effect of rHuEPO (red blood cell lifespan is around 120 days[17]), high frequency testing and intelligent testing strategies (e.g. targeted testing in combination with the ABP) would be necessary to attain a high probability to catch athletes abusing rHuEPO. A second concern with this assay is that a small part of the samples (5%) were non-detectable, especially because this result occurred more often in rHuEPO treated subjects (83% of these samples). Non-detectable results in general have been associated with a variety of reasons. Degradation of EPO via neuraminidase activity, with consecutive loss of sialic acid moieties will result in a shift in molecular weight of EPO to lower molecular weight sections.[18] Protease activity might lead to disappearance

of EPO. Of course suppression of endogenous production due to a negative feedback mechanism will also lower endogenous EPO levels, potentially below the limit of detection of the method. Finally, urine dilution could disturb the assay.[19] With regards to dilution however, only four samples with a non-detectable result (24% of all non-detectable samples) were strongly diluted, having an osmolality of <180 mOsm/kg (approximately corresponding to a specific gravity of <1.004,[20] the applied cut-off for not accepting a sample in anti-doping procedures). In contrast, there were 30 samples (10%) with similarly low osmolality that did have a positive or negative assay result. Therefore osmolality (alone) does not seem to explain a non-detectable result, and the cause for such a result remains unclear. This finding also questions the validity of the current process to reject samples based on specific gravity. Of the 17 samples designated as non-detectable for the sarcosyl-PAGE assay 13 had the same result in IEF, indicating the problem is probably sample-related. Given that only three subjects had more than one sample (namely two, and for one subject three) designated as non-detectable, this effect does not seem to be subject-related. On top of this unknown effect leading to non-detectable results, athletes might engage actively in manipulative behaviour to prevent detection of rHuEPO use. Implementing an additional immunoassay in the routine doping procedure for non-detectable samples evaluating if overall erythropoietin levels are sufficiently present in such samples, would therefore help interpreting the result.

The aim of the screening procedure is to forward all potential positives and not to exclude any false negatives and in the confirmation procedure, this 'philosophy is reversed'; i.e. not to report any false positives, as described previously.[21] The sarcosyl-PAGE assay partly does allow for this at reasonable economically viable conditions with less than 5% false presumptive findings forwarded to confirmation. What is remarkable however, is that when the results from the same sample in the screening and confirmation analyses are compared, at least in some cases these are clearly different, see Figure 4. This indicates that the difference in outcome might not solely be due to implementing the screening or confirmation procedure, but possibly also due to variability in assay performance.

Finally, sensitivity of the assay did not seem to be affected by the maximal exercise test. The race however did seem to impact sensitivity, with six subjects that tested positive at days 11 and 12 before the race testing negative after the race. Of these six negative post-race samples, two were non-detectable, potentially due to effects of the exercise and hydration. The other four samples however, were designated negative, indicating the race potentially affected the ability to detect rHuEPO: these four subjects went from positive to negative in approximately 8 hours, which is remarkable (see for two examples Figure 5). One explanation could be that the concentration of erythropoietin in the post-race samples was lower (possibly due to changes in urine concentration during the race), and that the faint signal that was detected pre-race was therefore not visible post-race.

Isoelectric focusing urine assay performance

Findings for the isoelectric focusing assay were similar to the sarcosyl-PAGE assay, although overall sensitivity was slightly worse for IEF, which is similar as in a previous study with micro-doses of rHuEPO.[22] The most striking difference between the two assays was a four-fold higher incidence of samples not being detectable for the isoelectric focusing (21%). This could be due to the fact that the isoelectric focusing is not currently being used by the DoCoLab in Ghent and was only implemented for this study, and less routine led to more samples being designated as non-detectable. Additionally, it could be due to the effect of previously described neuraminidase activity on IEF results.[18] The limit of detection for sarcosyl-PAGE as determined with reference standards was the same as for IEF. Samples were stored frozen immediately following collection until analysis. Samples were defrosted overnight in a fridge and aliquoted immediately before analysis. While for sarcosyl-PAGE mostly protease activity will lead to undetectable results and neuraminidase activity will lead to a shift (lower molecular weight) of the rHuEPO band, we believe that neuraminidase activity might lead to complete 'disappearance' of the EPO bands in IEF and not in sarcosyl-PAGE. This may contribute to the observed difference. Similar to the worrying finding for sarcosyl-PAGE, the majority of the non-detectable samples (75%) in IEF belonged to rHuEPO treated subjects. This higher incidence in samples from doped athletes in both IEF and sarcosyl-PAGE indicates that, while the fact that no EPO can be detected at all is not a proof of doping, this information might be useful to plan target and/or follow up tests as rHuEPO use might be associated with a non-detectable result. To rule out problems with the individual assay when such a result occurs, it could be useful to add a positive control to each sample, for example using EPO from a different animal species, as is common for most other doping analyses.

There was no overlap in false presumptive findings between the sarcosyl-PAGE and IEF assays, indicating that such a result is not related to the sample.

In contrast to the sarcosyl-PAGE, the maximal exercise test potentially did have an effect on the assay result of the isoelectric focusing: sensitivity was higher directly after the test compared to before. For the race however, we observed a similar outcome as for the sarcosyl-PAGE: three subjects tested positive before the race, after the race two tested negative, and one non-detectable. The impairment in signal for these post-race samples as observed in the sarcosyl-PAGE assay (Figure 5) was not observed in the IEF for these two samples. However, this is solely based upon observed signal intensities and not a quantitative observation. As no internal standard is used, factors such as recovery differences in sample preparation between different batches/samples might also play a role. If there indeed is an effect of exercise on the IEF test results, this effect was opposite for the race, compared to the maximal exercise test.

Athlete Biological Passport performance

In addition to evaluating the urine assays, we constructed an ABP so we could evaluate all tools used to detect rHuEPO use. This method had a high sensitivity of 91.3%; given that this data is collected longitudinally and the method is less dependent on the time of sampling in relation to dosing compared to the urine assay, this approach not surprisingly has a better chance to detect doping athletes. Nevertheless, almost 10% of subjects were not identified as suspicious or positive based on their ABP. See as an example the false negative subject in Figure 3, whose variation in haematological parameters is actually very similar to that of the placebo subject depicted in the left column of panels. This means that some athletes might not be at risk for detection based on the ABP despite using rHuEPO. Moreover, in practice sensitivity of the ABP might fall due to potential lower doses being used (micro-dosing) and less optimal timing of blood samples as the anti-doping organisation does not have the information on dosing times. Other studies evaluating the performance of the ABP showed somewhat ambiguous findings. One study used approximately half the weekly dose of our study after a high dose period of 250 IU/kg three times a week and found 100% subjects to have at least one suspicious or abnormal measurement.[23] This difference with our findings could be due to the very high starting dose used. Additionally, this study did not investigate evaluator determination based on the ABP, but only whether individual values were outside the ABP reference ranges. Also remarkable is that a second, unblinded, study using doses building up to a similar dose to our weekly dose, found no measurements during the study being flagged as abnormal.[24] This study used 99.9% likelihood ranges and did not investigate evaluator determination based on the ABP. Our own model used similar limits (99.7%) indicating the fact that an evaluator interpreted the ABP in our study might give more insight into the observed pattern.

Limitations

There are several limitations to our study. Firstly, although unlikely, it cannot be ruled out that placebo treated subjects administered rHuEPO outside the proceedings of the study.

Secondly, for reasons of restrictions in time and personnel, confirmation analysis was not performed for isoelectric focusing, and only for a selection of 14 samples, including all false presumptive findings, for sarcosyl-PAGE. In this respect, our study design did not follow regular anti-doping procedures to forward all screening positive samples to a confirmation analysis, nor did it perform a third (B-sample) analysis or a second opinion by an expert. As the confirmation analysis only included a limited number of samples, the characteristics of this confirmation analysis by itself, including the false positive rate, is somewhat uncertain.

Furthermore, most subjects received higher doses (8000 or 10 000 IU) at the last dose and the accompanying urine samples taken more than 10 days after dosing, than at the second dose (5000 IU), which could have impacted the urine assay results over time. But even with a higher dose sensitivity was well below 40% for these samples.

For the ABP, a limitation was that samples were analysed on a different type of analyser, although from the same manufacturer, than the Sysmex XT 2000i that is used by anti-doping laboratories. In addition, the ABP algorithm used by the WADA was not available to us, and so the current method was an approximation of the official procedure. Nevertheless, these discrepancies should not have a major impact on the ABP profiles and so the outcome of the review of the ABPs, we feel, can be considered indicative of the first stage of ABP performance at the APMU.

Information about factors that might influence markers in the ABP (e.g. altitude training) would be available for the evaluator's assessment. This information was not systematically recorded in the present study, and although no such events were reported by subjects, occurrence could have impacted ABP review. It should also be noted that the relatively high sampling frequency applied in this study, with intervals of as short as 2 weeks, could lead to the individual reference ranges moving along with the observed values more than with longer intervals and thereby reducing the chance values to fall outside the range. The selected window is however the smallest acceptable window according to the ABP and so these data should still be correctly interpretable. Finally, in the WADA setting, after assessment of an ABP by an APMU, suspicious ABPs will be sent for review to a WADA haematology expert. If this expert confirms the suspicious finding, this and two other experts will determine, based on all available information and possibly additional requested information, whether the profile is indeed indicative of blood doping; if consensus is reached. For this study, several WADA haematology experts were approached to participate in assessment of the ABPs of this study, but all of them declined. Therefore, the results reported here can only be considered to resemble the first part of the official system. But as only suspicious results are forwarded to the expert, ABP sensitivity would at best stay at the same level.

CONCLUSION

The sarcosyl-PAGE urine assay for rHuEPO (specifically epoetin-beta) did not show false positive results after confirmation analysis, but it does have shortcomings such as a limited detection window. In addition, it is of some concern that in this study we observed a higher rate of non-detectable samples in subjects treated with rHuEPO. Adding a positive control to each sample and measuring total erythropoietin levels using an immunoassay could address the latter issue. The IEF assay in our study was inferior to the sarcosyl-PAGE assay. The ABP had a much higher sensitivity than the assays using a single urine sample, although it too did not classify all passports correctly. In summary, we showed that all three methods evaluated in this study have their shortcomings and challenges and that it is critical to continue research to improve existing and develop new doping detection methods.

REFERENCES

1 World_Anti-Doping_Agency. World Anti-Doping Agency Code. 2015; https://www.wada-ama.org/en/resources/the-code/world-anti-doping-code. Accessed August 8, 2018.
2 Ekblom B. BB. Effect of erythropoietin administration on maximal aerobic power. *Scand J Med Sci Sports*. 1991;1(88-93).
3 Annaheim S, Jacob M, Krafft A, Breymann C, Rehm M, Boutellier U. RHEPO improves time to exhaustion by non-hematopoietic factors in humans. *Eur J Appl Physiol*. 2016;116(3):623-633.
4 Heuberger J, Rotmans JI, Gal P, et al. Effects of erythropoietin on cycling performance of well trained cyclists: a double-blind, randomised, placebo-controlled trial. *Lancet Haematol*. 2017;4(8):e374-e386.
5 World_Anti-Doping_Agency. Anti-doping testing figures. 2012-2017; https://www.wada-ama.org/en/resources/laboratories/anti-doping-testing-figures-report. Accessed August 8, 2018.
6 World_Anti-Doping_Agency. Technical Document for the harmonization of analysis and reporting of erythropoiesis stimulating agents (ESAs) by electrophoretic techniques. 2014; https://www.wada-ama.org/en/resources/science-medicine/td2014epo. Accessed August 8, 2018.
7 Reichel C. SARCOSYL-PAGE: a new electrophoretic method for the separation and immunological detection of PEGylated proteins. *Methods Mol Biol*. 2012;869:65-79.
8 Reichel C. Recent developments in doping testing for erythropoietin. *Anal Bioanal Chem*. 2011;401(2):463-481.
9 Bowers LD. The analytical chemistry of drug monitoring in athletes. *Annu Rev Anal Chem (Palo Alto Calif)*. 2009;2:485-507.
10 Lasne F, Thioulouse J, Martin L, de Ceaurriz J. Detection of recombinant human erythropoietin in urine for doping analysis: interpretation of isoelectric profiles by discriminant analysis. *Electrophoresis*. 2007;28(12):1875-1881.
11 Reihlen P, Blobel M, Kempkes R, et al. Optimizing SAR-PAGE. *Drug Test Anal*. 2015;7(11-12):1014-1016.
12 Berry DA. The science of doping. *Nature*. 2008;454(7205):692-693.
13 Heuberger JA, Cohen Tervaert JM, Schepers FM, et al. Erythropoietin doping in cycling: lack of evidence for efficacy and a negative risk-benefit. *Br J Clin Pharmacol*. 2013;75(6):1406-1421.
14 Hamilton TC, D. *The secret race: inside the hidden world of the Tour de France.* New York: NY: Random House Publishing Group; 2012.
15 Reichel C, Abzieher F, Geisendorfer T. SARCOSYL-PAGE: a new method for the detection of MIRCERA- and EPO-doping in blood. *Drug Test Anal*. 2009;1(11-12):494-504.
16 Gore CJ, Parisotto R, Ashenden MJ, et al. Second-generation blood tests to detect erythropoietin abuse by athletes. *Haematologica*. 2003;88(3):333-344.
17 Shrestha RP, Horowitz J, Hollot CV, et al. Models for the red blood cell lifespan. *J Pharmacokinet Pharmacodyn*. 2016;43(3):259-274.
18 Reichel C, Kulovics R, Jordan V, Watzinger M, Geisendorfer T. SDS-PAGE of recombinant and endogenous erythropoietins: benefits and limitations of the method for application in doping control. *Drug Test Anal*. 2009;1(1):43-50.
19 Lamon S, Robinson N, Sottas PE, et al. Possible origins of undetectable EPO in urine samples. *Clin Chim Acta*. 2007;385(1-2):61-66.
20 Imran S, Eva G, Christopher S, Flynn E, Henner D. Is specific gravity a good estimate of urine osmolality? *J Clin Lab Anal*. 2010;24(6):426-430.
21 Badoud F, Guillarme D, Boccard J, et al. Analytical aspects in doping control: challenges and perspectives. *Forensic Sci Int*. 2011;213(1-3):49-61.
22 Dehnes Y, Shalina A, Myrvold L. Detection of recombinant EPO in blood and urine samples with EPO WGA MAIIA, IEF and SAR-PAGE after microdose injections. *Drug Test Anal*. 2013;5(11-12):861-869.
23 Clark B, Woolford SM, Eastwood A, Sharpe K, Barnes PG, Gore CJ. Temporal changes in physiology and haematology in response to high- and micro-doses of recombinant human erythropoietin. *Drug Test Anal*. 2017;9(10):1561-1571.
24 Ashenden M, Gough CE, Garnham A, Gore CJ, Sharpe K. Current markers of the Athlete Blood Passport do not flag microdose EPO doping. *Eur J Appl Physiol*. 2011;111(9):2307-2314.

FIGURE 1 SARCOSYL-PAGE SENSITIVITY VERSUS TIME SINCE DOSE Sensitivity for the sarcosyl-PAGE assay over time in the rHuEPO treated subjects. The numbers at each data point represent the number of samples that are in a bin. Left panel includes samples that were designated 'Non-detectable', right panel excludes these samples.

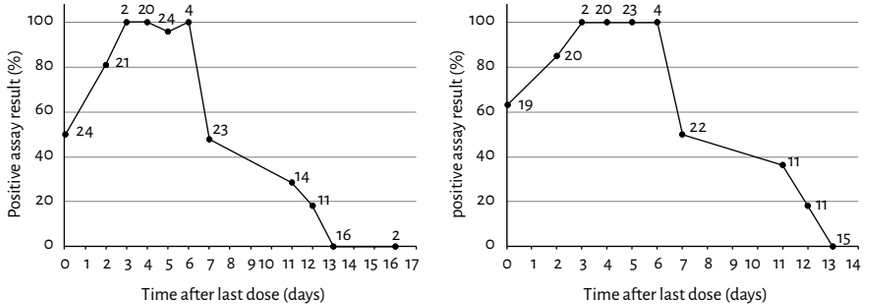


FIGURE 2 ISOELECTRIC FOCUSING SENSITIVITY VERSUS TIME SINCE DOSE Sensitivity for the isoelectric focusing assay over time in the rHuEPO treated subjects. The numbers at each data point represent the number of samples that are in a bin. Left panel includes samples that were designated "Non-detectable", right panel excludes these samples.

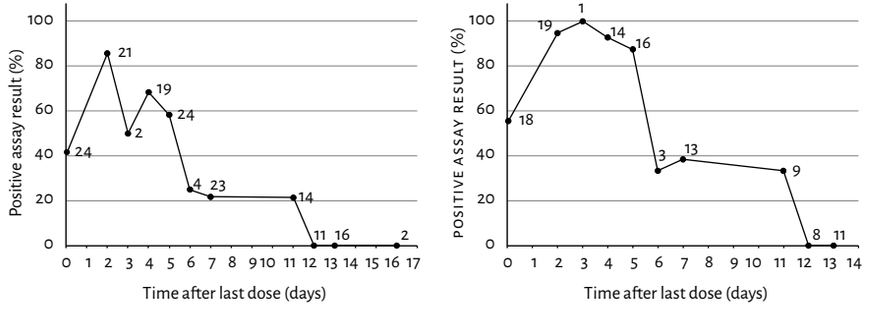


FIGURE 3 ATHLETE BIOLOGICAL PASSPORT GRAPHS EXEMPLARY SUBJECTS Six main panels of the Athlete Biological Passport of three exemplary subjects. Black data points and the accompanying grey numbers are the values observed for this parameter in the particular subject. Grey points are the calculated 99.7% CI of the (individual) normal values calculated as described in the text. Left column of panels shows a subject that received placebo that was correctly identified as not using rHuEPO. Middle column of panels shows a subject that received rHuEPO weekly for 8 weeks starting at time point 0 that was correctly identified as using rHuEPO. Right column of panels shows a subject that received rHuEPO weekly for 8 weeks starting at time point 0 that was incorrectly identified as not using rHuEPO. OFFS: OFF-score; IRF: Immature reticulocyte fraction.

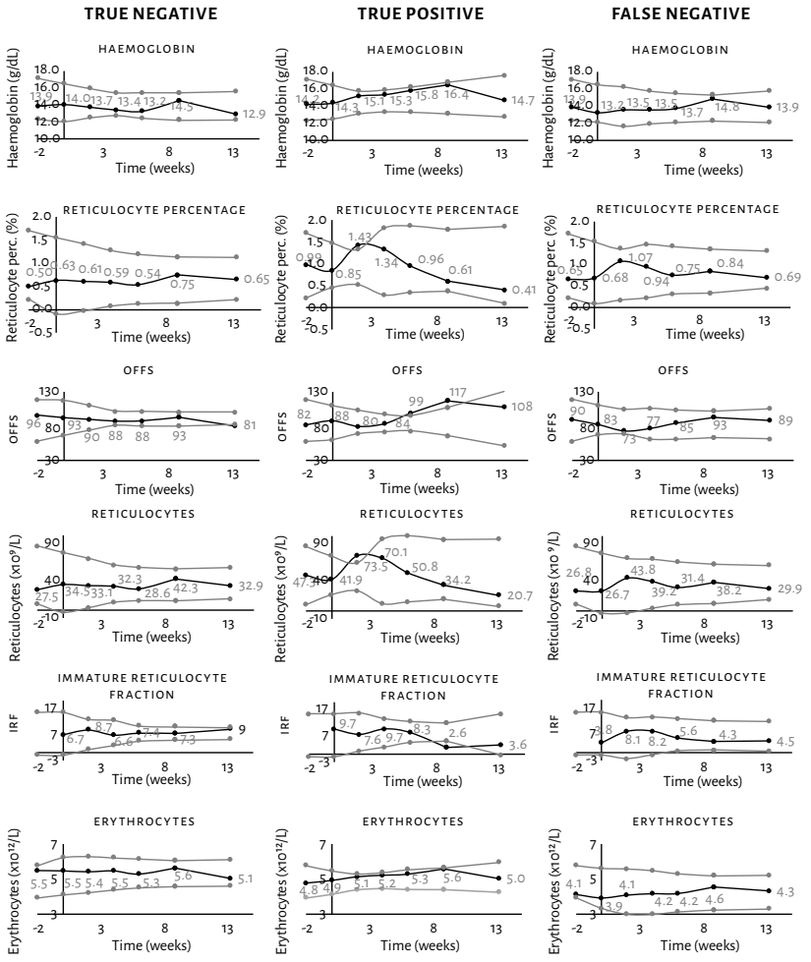


FIGURE 4 SARCOSYL-PAGE URINE ASSAY RESULTS OF SCREENING AND CONFIRMATION ANALYSIS OF FALSE PRESUMPTIVE FINDINGS Images of the sarcosyl-PAGE assay results taken from different gels with the nearest references (labelled ESA-MIX) and the negative quality control (labelled NU) for each relevant lane. Letters indicate the relevant lanes, with the matching upper-case and lower-case letters indicating the same sample. The white line indicates the level above which rHuEPO will stain. A) In lanes 10.39, 11.38, 12.37 and 17.33 urine samples taken from subjects receiving placebo were analysed and determined as suspect for rHuEPO use. In lanes 10.1009, 11.1008, 12.1007 and 17.1003 the same samples are analysed in the confirmation analysis. B) In lanes 12.193, 6.57 and 16.49 urine samples taken from subjects receiving placebo were analysed and determined as suspect for rHuEPO use. In lanes 3.1015, 6.1012 and 9.1010 the same samples are analysed in the confirmation analysis. All indicated samples in the screening analysis show staining above the white line which is what made them suspicious, with lanes 10.39, 11.38 and 12.37 even staining a large portion above the white line. This is not the case in the same samples in the secondary analysis (compare with lanes 10.1009, 11.1008, 12.1007).

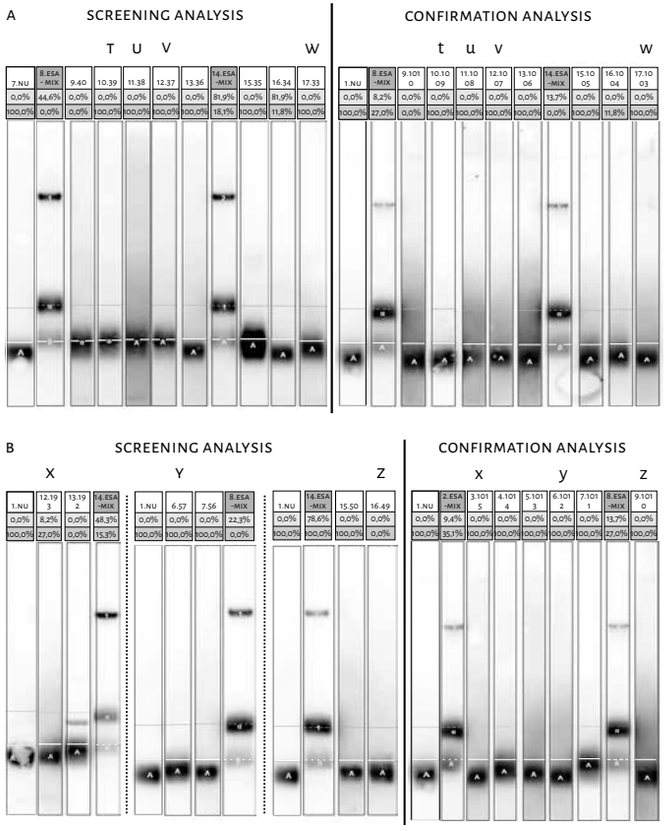


FIGURE 5 EXAMPLE OF CHANGES IN SARCOSYL-PAGE RESULTS DURING THE RACE Images of the sarcosyl-PAGE assay results taken from two gels. In lanes 16.268 and 17.267 urine samples taken pre-race from two subjects receiving rHuEPO were analysed. In lanes 5.315 and 6.316 samples taken post-race from the same subjects were analysed. Letters indicate the relevant lanes, with the A and a indicating lanes from one subject, and B and b from the other. Lanes labelled ESA-MIX are reference samples, lanes labelled NU are negative quality controls, the white line indicates the level where rHuEPO will stain. In lanes 16.268 and 17.267 staining is clearly present above the white line, while this is not the case in the samples from the same subjects that were taken only approximately 8 hours later, after the race, in lanes 5.315 and 6.316.

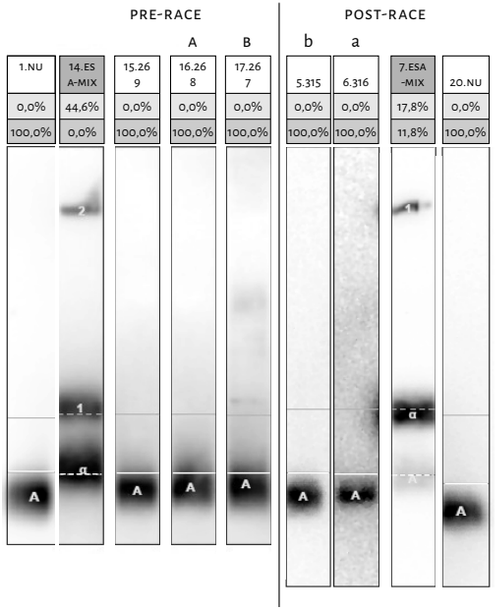


TABLE 1 SARCOSYL-PAGE URINE ASSAY PERFORMANCE CHARACTERISTICS (TOTAL N = 330)

Measure	Screening analysis	Screening + Confirmation analysis
	Value (%)	Value (%)
Sensitivity	63.8	63.8
Specificity	95.7	100.0
False negative rate	36.2	36.2
False positive rate	4.3	0.0
Precision	93.1	100.0
False discovery rate	6.9	0.0
False omission rate	25.6	24.8
Negative predictive value	74.4	75.2

TABLE 2 ISOELECTRIC FOCUSING URINE ASSAY PERFORMANCE CHARACTERISTICS (TOTAL N = 329)

Measure	Value (%)
Sensitivity	58.6
Specificity	94.0
False negative rate	41.4
False positive rate	6.0
Precision	87.8
False discovery rate	12.2
False omission rate	24.6
Negative predictive value	75.4

TABLE 3 ATHLETE BIOLOGICAL PASSPORT PERFORMANCE CHARACTERISTICS (TOTAL N = 47)

Measure	Evaluator 1	Evaluator 2
	Value (%)	Value (%)
Sensitivity	91.3	91.3
Specificity	100.0	95.8
False negative rate	8.7	8.7
False positive rate	0.0	4.2
Precision	100.0	95.5
False discovery rate	0.0	4.5
False omission rate	7.7	8.0
Negative predictive value	92.3	92.0

BRITISH JOURNAL OF CLINICAL PHARMACOLOGY 2018, 84(8):1830-1838

Jules AAC Heuberger*, Sven C van Dijkman* and Adam F Cohen

*These authors contributed equally to this work

ABSTRACT

Salbutamol is used in the management of obstructive bronchospasm, including that of some elite athletes. It is claimed that high salbutamol (oral) doses may also have an anabolic effect. Therefore, inhalation of salbutamol is restricted by the World Anti-Doping Agency (WADA) to a maximal daily dose. Urine is tested for violations, but recent cases have resulted in a debate regarding the validity of this approach. It was our aim to determine whether current approaches are sufficiently able to differentiate approved usage from violations. We extracted pharmacokinetic parameters from literature for salbutamol and its sulphated metabolite. From these parameters, a semi-physiological pharmacokinetic model of inhaled and orally administered salbutamol was synthesised, validated against literature data, and used to perform clinical trial simulations (N=1000) of possible urine concentrations over time resulting from WADA-allowed and oral unacceptable dosages. The synthesised model was able to predict the literature data well. Simulations showed a very large range of salbutamol concentrations, with a significant portion of virtual subjects (15.4%) exceeding the WADA threshold limit of 1000 ng/mL at 1-hour post-dose. The observed large variability in urine concentrations indicates that determining the administered dose from a single untimed urine sample is not feasible. The current threshold inadvertently leads to incorrect assumptions of violation, whereas many violations will go unnoticed, especially when samples are taken long after drug administration. These issues, combined with the dubious assertion of its anabolic effect, leads us to conclude that the large effort involved in testing should be reconsidered.

INTRODUCTION

Recently one of the most successful male cyclists of the last decade, Chris Froome, came into disrepute due to news of a potential doping violation. Doping control revealed a salbutamol (albuterol in the US) concentration exceeding the allowed limit of 1000 ng/mL in a urine sample provided by the British rider during the Vuelta a España of 2017.[1] The WADA prohibited list indicates that salbutamol use is allowed in inhaled dosages up to "1600 micrograms over 24 hours in divided doses not to exceed 800 micrograms over 12 hours starting from any dose", which is considered the maximum therapeutic dose for athletes with a so-called Therapeutic Use Exemption (TUE). Froome was in possession of such a TUE, but "the presence in urine of salbutamol in excess of 1000 ng/mL [...] is presumed not to be an intended therapeutic use of the substance and will be considered as an Adverse Analytical Finding unless the athlete proves, through a controlled pharmacokinetic study, that the abnormal result was the consequence of the use of the therapeutic dose (by inhalation) up to the maximum dose indicated above." Currently, over six months since the concerned urine sample, there is still no news of such a controlled study, which is perhaps not surprising: the burden is with the accused and setting up a robust study requires expertise. Additionally, inter-occasion variability will influence the results of each attempt.

Salbutamol acts on the beta-2-adrenoceptor as a sympathomimetic, commonly prescribed to counteract bronchoconstriction due to allergic and exercise-induced asthma, as well as chronic obstructive pulmonary disease.[2] This mechanism of action has led athletes to believe that there might be a performance enhancing effect of beta-2 agonists both through relaxing smooth muscle cell in the lung, and anabolic effects on skeletal muscles. Two reviews extensively investigated the evidence for these effects and concluded that inhaled beta-2 agonists have no positive effects on muscle strength, sprint or endurance performance, and that only high, systemic dosages can improve muscle strength and sprint performance, but not endurance performance.[3,4] In summary, beta-2 agonists might give an advantage in sports, but only at very high concentrations and for very short (power) disciplines. It is therefore doubtful that multi-stage (endurance) cyclists like Chris Froome would benefit even from high doses of beta-2 agonists apart from when treating asthmatic symptoms. Nevertheless, the WADA currently is of the opinion that all beta-2 agonists are banned substances, necessitating doping control measures. For this purpose, urine analysis is performed, which leads to a problem for the substances that are allowed

with a TUE up to a maximum dose: salbutamol, formoterol and salmeterol. For these substances, urine concentration is used to determine whether the maximum dose was exceeded. It is, however, impossible to determine dose from a single urine concentration due to several pharmacological factors that are at play. Salbutamol plasma pharmacokinetics in particular, are highly variable, mainly due to variability in lung and gut absorption, volume(s) of distribution, metabolism (including first-pass effect) and renal clearance.[5–7] Subsequently, urine pharmacokinetics are even more variable due to the additional factor of urine concentrating abilities of the kidney (urine osmolality between 50 and 1200 mOsm/kg),[8] volumetric production of urine[9] and micturition. On top of all this, the time since last dose in a doping control setting is unknown. Despite these facts, the WADA rules seem to assume that a urine concentration above the set threshold indicates a high chance of having detected use of more than the allowed dose. This chance is apparently considered sufficiently high to warrant sanctioning the athlete but supporting evidence in literature is lacking. The aim of this study is therefore to evaluate the current WADA approach to this problem and determine whether this approach is able to differentiate approved use from violations. As a part of that aim, we evaluate whether WADA approved doses of salbutamol may lead to unacceptable urine concentrations, considering the multiple sources of variability, using a PBPK approach.

METHODS

Several literature data sources were used to synthesise a semi-physiological pharmacokinetic model of plasma and urine salbutamol concentrations.[5–7] In short, a PK model of salbutamol in dogs was used as the basis and extrapolated to humans using allometric scaling.[7] Literature data on clearance of salbutamol and its main metabolite sulphated salbutamol was added to the base model.[5] A separate compartment was constructed to take into account the production of urine based on factors such as cardiac output and concentrating efforts of the kidneys.[9,10] Further adjustment of parameters was performed to align a visual predictive check against literature data of salbutamol plasma and urine concentration from Haase *et al*.[6] All modelling was performed in an environment consisting of Piraña v2.8.1,[11] psn v4.2.0[12] and NONMEM v7.3.[13] The statistical software R® v3.3.0[14] was used for pre and post-processing of data, and statistical and graphical analysis.

Literature model synthesis

Several pharmacokinetic models exist for salbutamol in literature; however, none of these include the urinary pharmacokinetics of salbutamol. A number of articles describe the concentrations of salbutamol in both plasma and urine without describing the parameters involved in the transfer of salbutamol between the two fluids. The only available model that approximates the absorption of salbutamol through both lungs and gut was built on data from dogs.[7] The model was implemented in NONMEM and the parameter values extrapolated to humans using allometric theory.[15] Variability on the fraction of inhalation absorbed through the gut and through the lungs was incorporated to account for inter-patient variability in inhalation efficiency, and differences in inhaler types and formulations.

The allometrically extrapolated model was further expanded by introducing a first-pass effect of gut absorbed salbutamol to its main metabolite sulphated salbutamol (S-SAL). Furthermore, the renal clearances of salbutamol and S-SAL were incorporated from pharmacokinetic parameters reported by Morgan *et al*.[5] Additional compartments for both salbutamol and S-SAL in urine were constructed, assuming that when the compounds are eliminated from plasma at the clearance rates provided by Morgan *et al*,[5] they are directly introduced to urine without delay.

A separate compartment was built to describe urine volume production, to allow voiding this compartment, similar to micturition, thereby better approximating physiology and allowing for the investigation of several micturition scenarios. Urine formation was assumed to occur at a constant rate, calculated using several physiological parameters, such as cardiac output (CO) scaled allometrically to weight ($CO = 0.166 * weight^{0.79}$).[10] The CO for a typical person of 70 kg then is 4.76 L, with a coefficient of variance (CV) of 20%. Of this CO, 21% flows through the kidneys. Typical haematocrit is 0.409 (CV: 7%) in trained cyclists,[16] thus 59.1% of blood is plasma, with 19% of plasma entering the renal capsule. Of the resulting glomerular filtrate, 99.2% is reabsorbed, leaving 0.8% to leave the kidneys as urine. Combining these numbers, we calculate the urine production in litre per hour for a typical adult weighing 70 kg as follows:[9]

$$4.76\ L/min * 60\ min * 21\% * 59.1\% * 19\% * 0.8\% = 0.054\ L/h \qquad \text{EQ. 1}$$

This amounts to 1.2 litres of urine per day, with a 95% prediction interval (PI) of 0.66 – 1.92 L/day, corresponding well with the typical volumes of 1-2 litres per day. This also includes variability due to concentrating by the kidney. Cardiac stroke volume is increased in elite athletes, but due to reduced heart rate at rest the cardiac output is not significantly different from normal healthy controls.[17] And although cardiac output is significantly increased (roughly 3 times) during exercise, renal blood flow is restricted, leading to only a mild to moderate reduction in the urine production rate.[18,19]

Model validation and calibration

We extracted data points of plasma and urine salbutamol concentrations from Haase *et al*[6] using ImageJ v1.50[20] and the Figure_Calibration plugin.[21] Haase *et al* administered 13 subjects with a single salbutamol inhalation of 1600 mcg, followed by regular plasma and urine sample collections. Validation of the synthesised literature model was performed through simulation of salbutamol plasma and urine concentrations in 1000 subjects (weight mean: 84 kg, SD: 17 kg), after a single 1600 mcg dose, and bladder voiding at the indicated time points where a urine sample was collected. The resulting simulated concentrations were compared to the plasma and urine concentrations over time in exercised, dehydrated subject data extracted from Haase *et al* and graphically presented in a visual predictive check plot.[22] Calibration of the parameters derived from the dog model was required, to better correspond to the Haase *et al* data.

Model simulation

Subsequently, we simulated a twice-daily 800 mcg inhaled salbutamol administration at steady-state (maximal allowed dose) in 1000 virtual subjects and determined the percentage of subjects attaining urine salbutamol concentrations above the 1000 ng/mL threshold to determine whether approved doses of salbutamol may lead to unacceptable urine concentrations. Finally, we also investigated a two week regimen of 8 mg oral salbutamol tablets (a dose shown to increase sprint power in elite athletes[23]) in 1000 simulated subjects, to determine the length of washout period associated with non-adverse findings in doping control. Both simulations (inhaled and oral dosing regimens) were used to evaluate whether the current WADA approach is able to differentiate approved use from violations.

RESULTS

Model synthesis

The synthesised model consisted of eight compartments; one each for absorption through the gut and lung, one central and one peripheral distribution compartment for parent salbutamol, one central metabolite compartment, one salbutamol and one metabolite urine compartment, and a single urine volume compartment (Figure 1). When administered through inhalation, 20% of the dose reaches central circulation through the lung, with 80% of the dose ingested and absorbed through the gut. Half of the gut-absorbed amount experiences a first-pass effect and is absorbed as sulphate metabolite, with the other 50% reaching the circulation as unchanged salbutamol. Total apparent clearance of salbutamol parent drug is 13.1 L/h (of which 18.6%, or 2.4 L/h is the rate of conversion to S-SAL), and clearance of S-SAL is 5.9 L/h. The higher apparent S-SAL clearance compared to conversion of salbutamol to metabolite in circulation explains why S-SAL only accumulates after oral administration of salbutamol, and not in the case of IV administration.[5]

Model validation

After several adjustments of pharmacokinetic parameters (Table 1), plasma and urine salbutamol concentrations were adequately predicted, with only a minor bias, and reasonably similar variation bandwidth (Figure 2). The dual absorption peaks, distribution and elimination phases were well-described.

Urine concentrations after appropriate salbutamol use

Simulations of steady-state urine concentrations over time resulting from a bi-daily administration of the approved 800 mcg show a large spread, with a significant portion (15.4%) of the simulated population achieving urine concentrations above the threshold of 1000 ng/mL at the peak concentration at 1 hour post-dose (Figure

3). At 12 hours after the dose administration, or right before the next inhalation, 0.7% of the population still showed urine concentrations above the threshold. It should be pointed out that these numbers do not take into account voiding the bladder before urine testing. In other words, Figure 3 shows the concentrations that would be measured at a certain time after administration when the bladder is voided for the first time since dose administration.

Urine concentrations after ergogenic salbutamol use

Simulations of urine concentrations over time resulting from a daily 8 mg oral dose (unacceptable usage), show that concentrations decline rapidly below the threshold after ceasing the regimen (Figure 4). Within the first 24 hours, the vast majority of subjects is already below the threshold, and after 2-3 days, none of the athletes will produce urine concentrations above the threshold.

DISCUSSION

We synthesised a model based on literature data alone that was able to adequately describe and predict the complex pharmacology of salbutamol in plasma and urine. The developed model was used to simulate possible outcomes of the maximum allowed dose of salbutamol in elite athletes, to show how current (WADA) standards for urine collection and analysis do not adequately take into account the large number of factors contributing to variability (dose amount and timing and physiological variability). This large variability leads to large uncertainty in determining the dose that was used, showing the implemented approach is not fit for purpose. It will lead to incorrect accusations of violation, whereas many violations will go unnoticed.

Applied correction factors by WADA

The general WADA rules for salbutamol use and control depicted in the introduction apply in all cases, but WADA describes two additional relevant specifics for the procedure. First, when collecting the urine sample from the athlete, the doping control officer will check the specific gravity of the urine produced by the athlete. If this is smaller than 1.005 when measured with a refractometer or smaller than 1.010 when measured with lab sticks, the athlete is required to provide further samples, until a suitable sample is collected.[24] The testing authority together with the laboratory decide which samples shall be analysed, although it is not clear on which criteria this decision is based. This procedure is presumably designed to prevent false negative findings due to diluted urine. However, it should be noted that due to bladder voiding for the first sample, the urine concentration of any subsequent sample will be driven by the plasma concentration at that moment and therefore potentially be a substantial underestimation, as plasma concentrations have fallen since dose administration. Correcting urine concentrations for urine osmolality (more accurate than specific gravity) of a sample would therefore be a much more rational way to normalise urine concentrations, and in addition avoid the need for multiple samples and their drawbacks.

Secondly, WADA corrects a measured urine concentration for assay variability (measurement uncertainty) by adding to the threshold concentration a guard band. The guard band corresponds to the expanded measurement uncertainty of the assay giving > 95% coverage interval for a result at the threshold concentration based on a 1-tailed normal distribution. It is calculated as 1.645 * ucMax (the maximum acceptable combined standard uncertainty of the assay, being 100 ng/mL for salbutamol), rounded up to 2 significant figures. A sample is determined to contain an adverse analytical finding only if the concentration is above the threshold plus the guard band, which is called the decision limit (i.e. 1200 ng/mL for salbutamol).[25] When this decision limit is applied, 9.95% of subjects were above the limit at 1-hour post-dose in our simulated scenario. Throughout the rest of this article the limit of 1000 ng/mL is used for clarity, as this is termed the threshold level by the WADA, also in the Prohibited List.

WADA deems it unnecessary to account for instability of salbutamol or its metabolite in urine, which potentially could impact measured salbutamol urine concentrations, as previous research showed that both salbutamol and its conjugate metabolites seem to be stable in urine.[26]

Involving pharmacology

So, the procedure involving specific gravity is rather dubious, and although the correction for assay variability seems appropriate, variability due to pharmacological processes are not discussed by WADA. Our developed pharmacokinetic model visualises the underlying pharmacokinetic theory linking urine concentration and dose, namely, the administered dose being absorbed into the circulation and appearing in urine as a fixed proportion of the amount in blood (renal clearance). This concentration in urine varies with dose amount and time after dose, bioavailability and absorption rate from the lung and gut (for inhalation), distribution over the body, renal clearance, urine volume and voiding. From a clinical perspective, this is very similar to creatinine clearance, and as clinicians might know: it is impossible to calculate the 'dose of creatinine' using only a single urine concentration. Therefore, to be able to make an informed estimate of the dose, one would need to know factors such as the time of dose (which in the doping control setting they do not), the physiological variability of the athlete (not known), volume of urine over a timed period (not known) and the plasma concentration (not known). And because these factors are unknown in the doping control setting, dose cannot be determined from the urine concentration. This is exemplified by simulations from the model in Figure 3, right panel, showing that even from a single dosing scenario, an extremely wide range of urine concentrations can be found. This would perhaps not be as problematic if a certain dose (the maximum allowed dose) would never lead to urine concentrations above the threshold for any subject, but our simulations show that this is not the case either. When collecting urine at 1 hour (close to t_{max}), 15.4% of our simulated subjects are above the threshold if they did not void since dosing. Conversely, due to the large variability and the unknown frequency of voiding and time since dosing, there is also a high chance of finding urine concentrations below the threshold with doses above the maximum allowed dose. Given the relatively short half-life of salbutamol, anabolic use (i.e. large oral doses) could even be halted shortly before a race, with no urine concentrations above threshold to be found by urine doping tests. Using our model, simulations of daily oral administration of 8 mg salbutamol over two weeks, with a washout period of 2 days resulted in 0% urine concentrations above the WADA threshold (Figure 4) when applying a normal micturition pattern (3 times a day). Moreover, within 24 hours, the majority of subjects would already produce a urine concentration below the threshold. This indicates that even high doses that have been shown to improve peak sprint power in elite endurance athletes,[23] will not lead to adverse analytical findings in the majority of cases if the athlete ceases dosing at least the day before an event.

Burden of proof

In the current situation, the WADA does seem to acknowledge the problem of variability to some extent as an athlete that produced a urine sample with an unacceptably high salbutamol concentration, is given the possibility to prove this was a result of a dosing scheme within WADA limits by means of a controlled pharmacokinetic study. Hereby the WADA transfers the responsibility of resolving the flaws in the rules designed by WADA itself to the athlete. Setting up such a study and getting the desired result will take months at least. And even if an athlete does prove his innocence, this could already do major damage to a reputation (see the Froome case). This is to say, if showing innocence will be successful at all, as this might not prove simple. Although intra-subject variability will be smaller than the previously described inter-subject variability, substantial variability will still be present within a subject. It is therefore not unlikely that many trials will be needed to produce another urine sample that exceeds the threshold with allowed dosages, even more so because it will be difficult to reproduce the circumstances leading to the original finding. Aside from this being a very expensive and time-consuming venture for the athlete, the fact that the foundation on which these WADA rules are based in the first place are flawed as we have shown, makes placing the burden of proof with the athlete completely unacceptable.

Alternative solutions

All these arguments make one wonder why the current procedure is being used by WADA. Speculation about a used dose of salbutamol from a urinary concentration (even when it would be adjusted for osmolality) is open to serious criticism and cannot be used to affect the career of an athlete. A better approach would be to collect timed urine samples over a specified period and possibly take a midpoint blood sample. In addition, S-SAL concentrations in urine could be incorporated, as

this metabolite would accumulate more, especially after oral dosing, due to slower clearance and the first-pass effect. One may in fact use an approach using Bayesian hierarchical modelling, as proposed by Mu & Ludden, to determine the most probable dose and dose-administration time.[27] Such an approach would require the development of an accurate population pharmacokinetic model of salbutamol (and S-SAL) in cyclists, including salbutamol in urine, with samples corrected for osmolality. This approach could help approximate the dose but would require large amounts of samples to be taken to achieve sufficient accuracy and precision. Without such dense data, uncertainty in estimating the dose would remain due to individual pharmacokinetic variability and the unknown factor of time since dosing. Seen how the performance enhancing activity of beta-2 agonists is dubious, especially in endurance sports, the question arises whether it is worth the effort of screening for these compounds.

Limitations

In this paper we reviewed the procedure implemented by the WADA in the control for salbutamol doping and indicated that these are fundamentally flawed. Many of these problems are no different from those encountered in clinical practice and clinical pharmacology, and so the theory and knowledge from these disciplines were applied in this study to the issue of doping control. We developed a population pharmacokinetic model based on literature data to substantiate and quantify that theory, but there are some limitations to the model. No individual data on both plasma and urine concentrations was available for model development, such as that of Haase *et al*.[6] Such data would allow proper estimation of inter-individual variability, leading to better prediction of concentration bandwidths. However, given the good performance of the visual predictive check in Figure 2, there does not seem to be over-prediction of the variability in urine concentration. The results of this visual predictive check also make it unlikely that inter-laboratory or inter-assay differences would impact extrapolation of our model to the WADA doping control laboratory setting: the pharmacokinetic analysis in the study used for the plot was performed by the WADA-accredited doping control laboratory in Norway.[6] Furthermore, variability on physiological parameters could only partially be based on actual data. For example, data on additional variability due to the extreme circumstances during professional cycling is not available and thus could not be accounted for. Two variabilities were fixed to 23% (omega=0.05), which is typically referred to when discussing physiology, and deemed reasonable. In addition, several parameters originally derived from dogs required manual adjustment to properly align with the Haase *et al*[6] data. Furthermore, the salbutamol concentration in the doping control assay used by the WADA is based on the sum of the glucuronide conjugate (expressed as the free drug) and free salbutamol concentrations. For the purpose of this study, we did not explicitly take into account the glucuronide conjugate of salbutamol, as its contribution to the concentration measured in urine is only very limited, with concentrations below the LLOD of 2 ng/mL in 100% of subjects after inhalation of 800 mcg of salbutamol and 70% of subjects after oral administration of 8 mg (and <3% compared to the unconjugated salbutamol for the remaining subjects).[28] However, if there would be an impact, this would only add to the variability in observed urine concentrations. Similarly, this variability might increase if we would incorporate variable urine production (due to increased fluid intake or dehydration) rather than a constant urine production rate that was applied in the model.

Finally, our model is supported by data from a clinical study,[28] reporting that out of 28 subjects (including 8 asthmatic elite athletes) inhaling a single dose of 800 mcg of salbutamol, there was one subject with a urine concentration above the 1000 ng/mL threshold when analysing the urine sample taken 4 hours post-dose (urine collection between 0-4 hours). This means in this study, 3.6% of subjects exceeded the urine concentration threshold with allowed use, which is similar to the observed 3.0% when simulating 1000 subjects with our model in this scenario, supporting the validity of our model.

Our approach therefore shows the feasibility of modelling exercises, including integration of pre-clinical and clinical data, with the possibility of clinical trial simulations / not-in-trial simulations and optimal design. Above all, we show that this approach, originating from traditional clinical pharmacology, can also be applied in this setting of doping control.

CONCLUSION

Pharmacokinetic theory dictates that it is not possible to derive the administered dose with any certainty from a single random urine concentration when there is no information about timing of the dose, urine volume or osmolality and individual physiological variability. Using a pharmacokinetic model based on literature data we substantiated this notion. We demonstrate that the current approach to detect excessive salbutamol use is fundamentally flawed and cannot differentiate between illegal and allowed use and inadvertently leads to incorrect assumptions of violation. If the community is determined to control for excessive salbutamol use, these procedures should be changed. The expertise present in the field of clinical pharmacology is clearly relevant in doping control, and we therefore advocate a closer collaboration between the two disciplines to work towards a sport that is as clean and fair as possible.

REFERENCES

1 Internationale UC. UCI statement on Christopher Froome 2017; http://www.uci.ch/pressreleases/uci-statement-christopher-froome/.
2 Johnson M. Molecular mechanisms of β2-adrenergic receptor function, response, and regulation. *Journal of Allergy and Clinical Immunology.* 2006;117(1):18-24.
3 Cairns SP, Borrani F. β2Adrenergic modulation of skeletal muscle contraction: key role of excitation-contraction coupling. *The Journal of Physiology.* 2015;593(21):4713-4727.
4 Pluim BM, de Hon O, Staal JB, et al. β2-Agonists and Physical Performance. *Sports Medicine.* 2011;41(1):39-57.
5 Morgan DJ, Paull JD, Richmond BH, Wilson JEvered E, Ziccone SP. Pharmacokinetics of intravenous and oral salbutamol and its sulphate conjugate. *British Journal of Clinical Pharmacology.* 1986;22(5):587--593.
6 Haase CB, Backer V, Kalsen A, Rzeppa S, Hemmers-bach P, Hostrup M. The influence of exercise and dehydration on the urine concentrations of salbutamol after inhaled administration of 1600 µg salbutamol as a single dose in relation to doping analysis. *Drug Testing and Analysis.* 2016;8(7):613-620.
7 Auclair B, Wainer IW, Fried K, Koch P, Jerussi TP, Ducharme MP. A population analysis of nebulized (R)-albuterol in dogs using a novel mixed gut-lung absorption PK-PD model. *Pharmaceutical Research.* 2000;17(10):1228--1235.
8 Sands JM, Layton HE. The Physiology of Urinary Concentration: an Update. *Seminars in nephrology.* 2009;29(3):178-195.
9 OpenStax_College. *Anatomy & Physiology.* http://cnx.org/content/col11496/latest/: Rice University Press; 2013 Accessed 10 Apr 2018.
10 Holt JP, Rhode EA, Kines H. Ventricular volumes and body weight in mammals. *American Journal of Physiology.* 1968;215(3):704-715.
11 Keizer RJ, Karlsson MO, Hooker a. Modeling and Simulation Workbench for NONMEM: Tutorial on Pirana, PsN, and Xpose. *CPT: pharmacometrics \& systems pharmacology.* 2013;2(November 2012):e50.
12 Lindbom L, Pihlgren P, Jonsson N. PsN-Toolkit - A collection of computer intensive statistical methods for non-linear mixed effect modeling using NONMEM. *Computer Methods and Programs in Biomedicine.* 2005;79:241--257.
13 Bauer RJ. NONMEM user guide introduction to NONMEM 7.3.0. 2015 Available at https://nonmem.iconplc.com/nonmem730/nm730.pdf (last accessed 20 Mar 2018).
14 R: A language and environment for statistical computing. 2014.
15 Anderson BJ, Holford NHG. Mechanism-based concepts of size and maturity in pharmacokinetics. *Annual review of pharmacology and toxicology.* 2008;48:303--332.
16 Mørkeberg JS, Belhage B, Damsgaard R. Changes in Blood Values in Elite Cyclist. *Int J Sports Med.* 2009;30(02):130-138.
17 Pluim BM, Chin JC, De Roos A, et al. Cardiac anatomy, function and metabolism in elite cyclists assessed by magnetic resonance imaging and spectroscopy. *European Heart Journal.* 1996;17(8):1271-1278.
18 Ho CW, Beard JL, Farrell PA, Minson CT, Kenney WL. Age, fitness, and regional blood flow during exercise in the heat. *Journal of Applied Physiology.* 1997;82(4):1126-1135.
19 Chapman CB, Henschel A, Minckler J, Forsgren A, Keys A. The effect of exercise on renal plasma flow in normal male subjects. *The Journal of Clinical Investigation.* 1948;27(5):639-644.
20 Eliceiri K, Schneider CA, Rasband WS, Eliceiri KW. NIH Image to ImageJ : 25 years of image analysis historical commentary NIH Image to ImageJ : 25 years of image analysis. *Nature Methods.* 2012;9(7):671--675.
21 Hessman FV. Figure _ Calibration ImageJ plugin. 2009 (last accessed 20 Mar 2018):http //www.astro.physik.uni--goettingen.de/\~ hessman.
22 Post TM, Freijer JI, Ploeger BA, Danhof M. Extensions to the Visual Predictive Check to facilitate model performance evaluation. *Journal of Pharmacokinetics and Pharmacodynamics.* 2008;35(2):185--202.
23 Hostrup M, Kalsen A, Auchenberg M, Bangsbo J, Backer V. Effects of acute and 2-week administration of oral salbutamol on exercise performance and muscle strength in athletes. *Scandinavian Journal of Medicine & Science in Sports.* 2016;26(1):8-16.
24 World_Anti-Doping_Agency. *Urine Sample Collection Guidelines.* https://www.wada-ama.org/sites/default/files/resources/files/wada_guidelines_urine_sample_collection_2014_v1.0_en.pdf October 2014 2014.
25 World_Anti-Doping_Agency_Laboratory_Expert_Group. *Decision limits for the confirmatory quantification of threshold substances* https://www.wada-ama.org/sites/default/files/resources/files/2016-12-13_td2017dl.pdf2017.
26 Berges R, Segura J, Ventura R, et al. Discrimination of prohibited oral use of salbutamol from authorized inhaled asthma treatment. *Clin Chem.* 2000;46(9):1365-1375.
27 Mu S, Ludden TM. Estimation of population pharmacokinetic parameters in the presence of non-compliance. *J Pharmacokinet Pharmacodyn.* 2003;30(1):53-81.
28 Mareck U, Guddat S, Schwenke A, et al. Determination of salbutamol and salbutamol glucuronide in human urine by means of liquid chromatography-tandem mass spectrometry. *Drug Testing and Analysis.* 2011;3(11-12):820-827.

FIGURE 1 DIAGRAM REPRESENTING FINAL MODEL STRUCTURE Arrows represent the flow of drug amounts. Bioavailability for lung and gut absorption after inhalation is represented as percentages. Lung-absorbed amount is directly introduced into the central compartment and consists of 100% parent salbutamol. Of the amount absorbed from the gut, 50% is parent salbutamol and 50% is directly converted to sulphated salbutamol metabolite (first-pass effect).

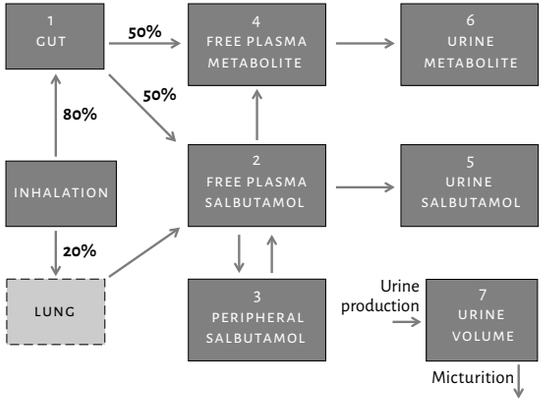


FIGURE 2 VISUAL PREDICTIVE CHECK OF SIMULATED CONCENTRATIONS AFTER AN INHALATION OF 1600 MCG SALBUTAMOL Black solid line: median predicted concentrations. Grey dashed lines: 95% concentration prediction interval. Grey points: observations extracted from Haase *et al.*[6] Left panel: plasma concentrations. Right panel: urine concentrations

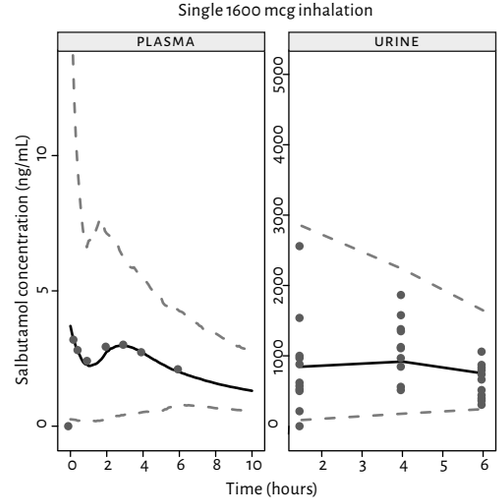


FIGURE 3 SIMULATED URINE CONCENTRATIONS OVER TIME, AFTER ADMINISTRATION OF THE ACCEPTED INHALATION OF 800 MCG SALBUTAMOL BI-DAILY AT STEADY-STATE (*LEFT PANEL*) AND THE RESULTING SPREAD IN MEASURED CONCENTRATIONS WHEN TIME IS NOT TAKEN INTO ACCOUNT IN DOPING CONTROL (*RIGHT PANEL*) Note: the left panel shows the concentrations that would be measured at a certain time after administration when the bladder is voided for the first time since dose administration, with a constant urine production rate. Black solid line: median predicted concentrations. Grey dashed lines: 99.9% concentration prediction interval. Bar plot (right panel): median (black point) and 99.9% prediction interval (bar). Note that the upper limit of this bar is lower than the upper limit peak concentrations in the left panel due to depicting a 99.9% prediction interval of an untimed sample in the right panel.

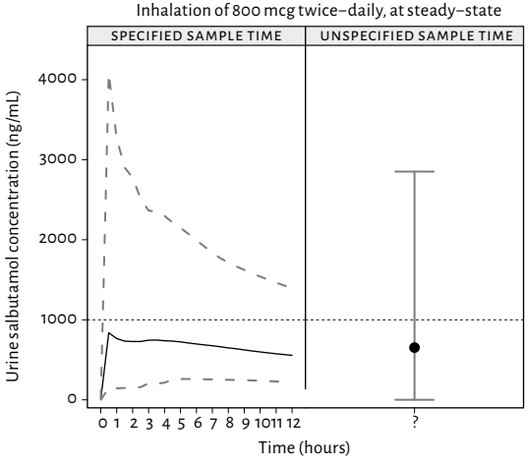


FIGURE 4 SIMULATED URINE SALBUTAMOL CONCENTRATIONS AFTER THE LAST DOSE OF A TWO-WEEK TREATMENT WITH 8 MG ORAL SALBUTAMOL TABLETS, AT REGULAR MICTURITION INTERVALS OF 8 H WITH A CONSTANT URINE PRODUCTION RATE The last dose of steady-state dosing is shown at 0 h, illustrating that levels decline to below WADA levels of 1000 ng/mL (horizontal dashed black line) well within 48 h of the last dose taken for 99.9% of the simulated study subjects. Black solid line: median predicted concentrations. Grey dashed lines: 99.9% concentration prediction interval.

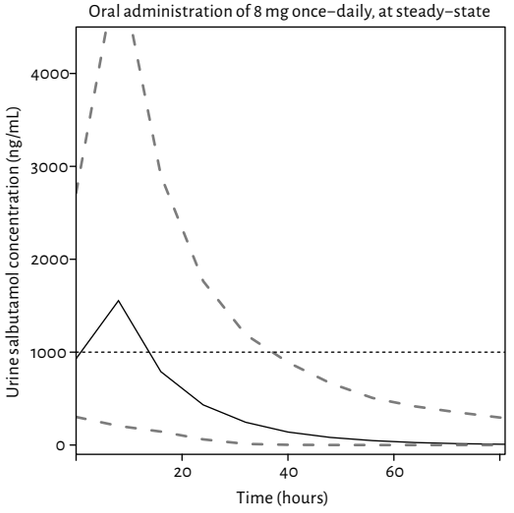


TABLE 1 MODEL PARAMETERS USED IN THE SIMULATIONS, WITH THEIR SOURCE

Parameter	Typical value	cv%	Source
Cardiac output (x weight$^{0.79}$) (L/min)	0.166	23%[b]	Holt[10]
Haematocrit	0.409	7%	Mørkeberg[16]
Bio-availability Gut (%)	80	23%[b]	Auclair[7]
Absorption constant gut (h^{-1})	0.5[a]	57%	Auclair[7]
Absorption lag gut (h)	1.5	83%	Auclair[7]
Renal clearance salbutamol (L/h)	17.5	25%	Morgan[5]
Renal clearance S-SAL (L/h)	5.91	25%	Morgan[5]
Central volume salbutamol (L/kg)	1.12[a]	63%	Auclair[7]
Peripheral volume salbutamol (L/kg)	1.92[a]	50%	Auclair[7]
Intercompartmental clearance (L/h)	0.56[a]	37%	Auclair[7]
Proportional error salbutamol (%)	23		
Proportional error S-SAL (%)	23		

a. Adjusted to better correspond to data from Haase et al.[6]

b. Fixed to 23% (omega = 0.05 in NONMEM*) due to limited data.*

9
DISCUSSION AND CONCLUSIONS

Pharmaceutical research and development has mainly been driven by the discipline of medicine, and for good reason: there is a large societal benefit to be gained in being able to use pharmacological agents to treat diseases. This societal motivator is perhaps less strong in other disciplines where drugs are used, such as doping in sports, but the experience in medicine could nonetheless be applied to develop knowledge and evidence in these disciplines as well. One lesson learnt from clinical pharmacology in medicine is that biological systems are complex and that performing successful pharmacological interventions is therefore not straightforward. This complexity is for example reflected in the low success rate for market approval of new drugs in drug development, which is currently around 10-15%.[1] For these few successful drugs, there is sufficient evidence that there is a beneficial effect in a certain population and that this benefit outweighs the potential harm of the treatment (positive benefit-risk assessment). This benefit-risk assessment is the cornerstone for the approval of drugs, both for the FDA and EMA,[2,3] and factors such as study design and choice of endpoints determine the strength of evidence and the uncertainties that drive this assessment. This is relevant as most doping substances are therapeutics, often approved for medical use (see the World Anti-Doping Agency (WADA) Prohibited List[4]), but for none of these substances the assessment was focused on improving sports performance. Sports performance, a product of biological systems, is also determined by a complex system of factors (take for example endurance performance[5]), meaning that the evidence available on the medical application of a drug is not sufficient for understanding its specific effects in sports or its benefit-risk ratio in that setting. In addition, where in medicine we can target the sub-optimally functioning system causing a disease, the physiological system of athletes has usually optimised (to some extent) for the specific task due to training. This makes improving the trained system possibly even more difficult, not in the least because it is not evident to determine which factor(s) in that system could be successfully targeted with a pharmacological intervention. Moreover, even in medicine seemingly simple and well-understood symptoms or diseases can prove complex to treat. The Cardiac Arrhythmia Suppression Trial (CAST) for example was designed to suppress premature ventricular complexes (PVC) after myocardial infarction to reduce mortality. Although antiarrhythmics encainide and flecainde indeed reduced occurrence of PVC,[6] mortality actually increased in the randomised, placebo-controlled trial that followed.[7] And in chronic kidney disease, where erythropoietin production is reduced causing anaemia, restoring haemoglobin to almost normal levels using recombinant human erythropoietin has been shown to cause only modest improvement in patient-reported fatigue and more importantly, increase the risk of stroke compared to placebo treatment.[8] Such (counter-intuitive) evidence shows that clinical pharmacology can be more complex than perhaps anticipated, and that randomised, placebo-controlled trials, using relevant clinical endpoints are needed to reveal the true effect of a treatment. Or to put it differently: it has proven difficult enough to pharmacologically improve physiological systems when they are in a diseased state, improving the system beyond the normal state requires at least a similar degree of thoroughness. However, the first two chapters of this thesis show that the evidence for performance enhancing effects of substances that are termed doping is limited. Evidence for effects of combinations of substances is virtually absent. For most substance classes no robust studies have been performed, for some classes there is evidence that there is no positive effect on performance, and for only 5 out of 23 classes there is evidence of a beneficial effect. Moreover, a meagre 11 studies including a total of 266 subjects form the total evidence-base for pharmacological performance enhancement, a large contrast with the amount of studies and subjects that substantiate the benefit of therapies in medicine. Despite this, the conviction of the positive effect of (the combination of) doping substances is strong, which is the driving factor for athletes to abuse these drugs. The discrepancy between this belief and the available evidence is striking, and an effort to generate this missing evidence is therefore necessary. As in regular medicine, this should start by studying effects of single substances in well-designed studies, and only if deemed relevant based on this evidence and the potential pharmacological mechanism, for combinations of substances. For one substance class, recombinant human erythropoietin (rHuEPO), we describe in this thesis a contribution to this effort. Interestingly, when we applied a randomised, placebo-controlled study design, we showed that although increases in haemoglobin lead to an improvement in several surrogate performance markers in athletes, this did not improve clinical outcome, i.e. endurance performance. It is unknown whether these findings also apply to professional athletes, as these could not be included in the study due to

current WADA regulation. However, our study was the first to study effects on clinically relevant outcome measures in trained athletes, and there was no indication that the highest performing participants showed stronger effects of rHuEPO treatment. These findings emphasise, as we have seen in therapeutic studies, that pharmacological interventions induce complex effects which may be difficult to predict on a clinical level. Therefore, the selection of endpoints is critical for the interpretation and implications of findings, and our study in this respect touches upon another important shortcoming in doping research. Effects of a treatment should be evaluated based on relevant clinical outcome measures or validated surrogate markers, which is currently not always the case. One such often used surrogate marker in sports medicine, the lactate threshold, we have evaluated in this thesis, including the issues associated with the marker. By determining repeatability and the (limited) relation with actual sports performance, the value of such a marker as a surrogate endpoint can be better assessed. Finally, apart from evidence on the potential positive effects of doping substances, it is also relevant to gain knowledge on their safety profile. For rHuEPO we describe such a detailed assessment in this thesis, showing that endothelial activation occurs both through the pharmacological intervention alone and in combination with exercise, indicating a potential risk for athletes abusing these drugs. Summarizing these findings, we conclude that there is incomplete information and a clear lack of evidence for performance enhancement for the majority of substance classes on the Prohibited List, which could be considered alarming. A more systematic effort using well-designed research into doping substances could address these information gaps and thereby potentially reduce the unwanted use of substances that show no benefit on performance and/or evidence for negative effects. An additional problem that results from the lack of information, is that for all substances on the List there is a need for detection strategies and assays to catch cheating athletes. Not only is this an enormous operation that is very expensive and burdensome for both athletes and anti-doping agencies; in the final chapters of this thesis we show that the applied detection methods and procedures themselves are not without flaws and shortcomings either. The doping detection regulation that is in force therefore on top of all this does not unquestionably protect fair playing athletes or catch cheating athletes. Anti-doping efforts would greatly benefit from a more clinical pharmacological approach in this respect too, by making use of the available (medical) knowledge and techniques. Overall, the combination of all of these uncertainties and shortcomings leads to a doping system that is not fully transparent, nor consistent, damaging both athletes and sports in general.

The conclusion that therefore must be drawn from all this is that sports medicine, just as regular medicine, needs to move towards a more evidence-based approach. This means implementing state-of-the-art study designs, with randomised, placebo-controlled, double blind set-ups and developing and using relevant biomarkers in these studies that measure and/or predict the pharmacological effect and accompanying impact on physiological performance and finally actual sports performance. The generated knowledge and evidence will benefit anti-doping efforts, improving decision-making and optimizing implemented procedures. For example, if there is strong evidence a substance class does not enhance performance, the class does not have to be prohibited. This evidence can in turn be used to prevent use through education, and detection efforts for these substances could be dropped. On the other hand, if there is strong evidence a substance class does enhance performance, the substance class should be prohibited, and all efforts should go into developing/improving reliable and robust detection methods, based on clinical pharmacological principles. If in addition there is also strong evidence of a health risk of such a class, this again can be used to deter use through education and targeted measurements of risk factors. The proposed shift would require, among others, revision of the WADA Code,[9] but in many cases it would be possible to apply these evidence-based principles in the doping setting, perhaps with some minor modifications compared to regular medicine. If this approach is to be successful, a few obstacles may need to be overcome, and we will deal with three of these.

One evident obstacle is the amount of resources available for doping research, as the proposed studies and detection method development require funding. To be implemented structurally, a redistribution of funding will therefore be required. Some resources could be made available by focusing doping testing efforts only to the proven effective substances; a more structured and evidence-based approach to compiling the Prohibited List, would for example very likely result in a reduction of the enormous amount of substances on the List that need testing. Another source of funding might need to come from the business of sports itself. Astronomical sums are involved with many sport disciplines through media rights, salaries and sponsor

contracts. In order to keep sports fair, safe and sustainable, these parties need to be convinced of the motivation for the proposed changes, possibly opening the possibility to organise a part of these resources to flow to anti-doping efforts. For a sponsor for example, taking part in such a system might be interesting because of the publicity value of being associated with clean sports.

A second obstacle to implementing an evidence-based approach would be that new substances could be used by athletes while research is not yet concluded. We would argue that depending on how likely such a substance could enhance performance (based on the mechanism of action and (pre-)clinical studies), it can be placed on the Prohibited List until robust evidence is available, after which re-evaluation should take place. This evidence should contain information on the effect on sports performance as well as safety evaluations, generated in well-designed clinical studies using relevant outcome measures and taking into account the pharmacological mechanism of the substance. In the organised system we propose, this should not take more than a few years. So, if we look at the substances currently on the List, for many of these the required evidence could already have been available, as they have been on the Prohibited List for many years.

Finally, a potential obstacle would be that large trials are needed to power studies in order to detect very small beneficial effects that might be relevant for cheating athletes. As in regular medicine, this should be addressed by first establishing what a relevant effect would be and performing a power calculation for the number of subjects needed in a trial to detect such an effect. If it then turns out difficult to access a sufficiently large study population, combined and dedicated efforts should be made to deal with this problem, as is often the case in regular medicine. One consideration could be changing the current rules preventing elite athletes to participate in these trials.

Overall, there are challenges of implementing an evidence-based approach in doping research, but as described here they are not insurmountable. At the same time, we have shown in this thesis that the current approaches to doping regulation are undesirable. Moving towards an evidence-based approach in doping research will create better focused anti-doping efforts and will lead to more rational, transparent and objective doping regulation. This will only benefit athletes and spectators, and therefore in our view is the way forward.

REFERENCES

1 Hay M, Thomas DW, Craighead JL, Economides C, Rosenthal J. Clinical development success rates for investigational drugs. *Nat Biotechnol* 2014; 32(1): 40-51.

2 Committee for Medicinal Products for Human use T. Report of the CHMP working group on benefit-risk assessment models and methods, 2 007.

3 Food and Drug Administration US. Benefit-Risk assessment in drug regulatory decision-making. 2018.

4 World Anti-Doping Agency T. Standard prohibited list 2018. Montreal: World Anti-Doping Agency, 2018.

5 Joyner MJ, Coyle EF. Endurance exercise performance: the physiology of champions. *J Physiol* 2008; 586(1): 35-44.

6 Cardiac Arrhythmia Pilot Study I. Effects of encainide, flecainide, imipramine and moricizine on ventricular arrhythmias during the year after acute myocardial infarction: the CAPS. *Am J Cardiol* 1988; 61(8): 501-9.

7 Echt DS, Liebson PR, Mitchell LB, et al. Mortality and morbidity in patients receiving encainide, flecainide, or placebo. The Cardiac Arrhythmia Suppression Trial. *N Engl J Med* 1991; 324(12): 781-8.

8 Pfeffer MA, Burdmann EA, Chen CY, et al. A trial of darbepoetin alfa in type 2 diabetes and chronic kidney disease. *N Engl J Med* 2009; 361(21): 2019-32.

9 World_Anti-Doping_Agency. World Anti-Doping Agency Code. 2015. https://www.wada-ama.org/en/resources/the-code/world-anti-doping-code (accessed August 8 2018).

NEDERLANDSE SAMENVATTING

De klinische farmacologie van prestatiebevordering en doping detectie in sport

Al sinds het ontstaan van competitieve sport proberen deelnemers vals te spelen om beter te presteren. Dat kan op allerlei manieren, zoals het afsnijden van een hardloopparcours, de trein nemen tijdens een lange wielerwedstrijd, of doen alsof er een overtreding wordt gemaakt bij voetbal ('schwalbe'). Mits goed uitgevoerd en onopgemerkt door de officials levert zulk vals spel een voordeel op; de deelnemer verandert (en vergemakkelijkt) hiermee immers de vereisten van de sport. Voor een ander soort poging tot prestatiebevordering is het echter minder vanzelfsprekend dat de sporter er daadwerkelijk voordeel mee doet: dopinggebruik, oftewel het gebruik van stoffen of methoden die door het Wereld Anti-Doping Agentschap (WADA) verboden zijn binnen de sport. Dergelijk vals spel kan namelijk alleen prestatie bevorderend zijn als het gebruik van de middelen de relevante fysieke vereisten van een sport positief kan beïnvloeden. En juist dat laatste is maar de vraag; de stoffen op de dopinglijst hebben een bepaald farmacologisch effect, het merendeel wordt zelfs als behandeling voor medische aandoeningen ingezet, maar hebben deze middelen ook voordelige effecten op de gezonde en getrainde sporter? Waar er voor geneesmiddelen hard bewijs vereist is, afkomstig uit robuust opgezet klinisch onderzoek, om te kunnen concluderen dat er een positief effect is bij de behandeling van patiënten, is dat in de doping *Code* van de WADA niet nodig om een middel op de verboden lijst te zetten. We kunnen dus niet zomaar aannemen dat de middelen op de dopinglijst ook echt prestatie bevorderen. Desalniettemin is er altijd een sterk geloof geweest in het prestatie bevorderende effect van een hoop stoffen; het gebruik van middelen met een farmacologische werking wordt al sinds de oude Griekse Olympische Spelen toegepast. Ook in de huidige tijd wordt er nog steeds doping gebruikt, af te leiden uit het regelmatig verschijnen van bekentenissen van sporters en hun staf en van bevindingen van anti-doping instanties en politie. Om dit verboden gebruik tegen te gaan en op te sporen werken anti-doping instanties aan voorlichting over en bestrijding van doping, en worden (top)sporters gecontroleerd door middel van dopingcontroles. Dergelijke werkzaamheden vragen veel onderzoek, tijd en geld en hebben behoorlijke impact op het leven van de sporter. Desondanks zijn er (blijkbaar) sporters die doping gebruiken en is het systeem van controles niet waterdicht waardoor dopinggebruikers onopgemerkt blijven of sporters vals worden beschuldigd. Kortom, er is een duidelijk probleem met doping in de sport. En gezien het belang van sport in de maatschappij (sociaal, financieel, emotioneel) is dit ook een maatschappelijk probleem, onderstreept door de hoeveelheid berichten over doping in de media. Het doel van dit proefschrift is daarom bij te dragen aan een oplossing voor dit probleem door de kennis van de (klinische) farmacologie toe te passen op het omgaan met het gebruik en de detectie van doping.

Zoals gezegd zijn de middelen die op de dopinglijst staan zijn voor het overgrote deel geneesmiddelen, en om een stof als geneesmiddel te kunnen registreren zijn er strenge eisen. Een van die eisen is dat er robuust bewijs is dat een middel een positief effect geeft, dat wil zeggen dat dit aangetoond moet zijn in gerandomiseerde, placebogecontroleerde, dubbelblinde klinische studies. Bij doping en prestatiebevordering gaat het om dezelfde middelen en gelden dezelfde (biologische, farmacologische) principes, dus geldt dat eenzelfde mate van bewijs benodigd is om te bepalen of een middel de prestatie daadwerkelijk bevordert. En omdat de belangrijkste (of misschien wel de enige) reden dat een sporter doping gebruikt het bevorderen van zijn prestatie is, is het absoluut van belang om te bepalen of een middel een effect heeft op sportprestatie. Is er geen positief effect van een middel dan kan die informatie het gebruik ontmoedigen, en is er een duidelijk positief effect dan is het noodzakelijk het gebruik ervan te voorkomen en gebruikers op te kunnen sporen. Als er geen bewijs wordt geleverd zou elk middel potentieel prestatie bevorderend kunnen werken en wordt de lijst van verboden middelen alleen maar langer. En dat is exact wat er gebeurt. Op de huidige WADA dopinglijst staan 23 klassen van middelen (bij elkaar enkele honderden stoffen), en zoals we in dit proefschrift laten zien is er maar voor 5 klassen robuust bewijs dat ze de prestatie bevorderen. Voor positieve effecten op de duurprestatie is er zelfs voor geen enkel middel bewijs. Voor de overige 18 klassen is er geen solide bewijs dat er een positief effect is op prestatie, sterker nog, voor 6 klassen op de lijst is er bewijs dat ze de sportprestatie niet bevorderen. Onder die 6 klassen valt ook de klasse van een van de meest bekende en beruchte vormen van doping, EPO, oftewel erythropoietine. Hoewel de bestaande onderzoeken naar recombinant humaan EPO (rHuEPO) een effect laten zien op specifieke metingen die iets met prestatie te maken hebben, is er geen bewijs dat daadwerkelijke sportprestatie, zoals een tijdrit in een wielerwedstrijd, positief beïnvloed wordt. Hier speelt namelijk een

ander principe uit de klinische farmacologie: is de gemeten uitkomstmaat relevant voor hetgeen je uiteindelijk wilt beïnvloeden (namelijk sportprestatie)? Alleen dan kan het effect op een zogenaamde surrogaat marker worden gezien als bewijs voor een daadwerkelijk relevant effect. Omdat dit bewijs voor rHuEPO naar onze visie ontbrak hebben we vervolgens een gerandomiseerde, placebogecontroleerde, dubbelblinde studie uitgevoerd met 48 getrainde wielrenners, waarbij niet alleen de al bekende surrogaat markers (zoals maximale zuurstofopname en maximaal vermogen) werden gemeten, maar ook de voor wielrennen meer representatieve maten van prestatie in een tijdrit en in een wegwedstrijd op de Mont Ventoux. Net als voorgaande studies vonden we een positief effect van rHuEPO op de surrogaat eindpunten, maar prestatie op de tijdrit of de wegwedstrijd verbeterde in de rHuEPO groep niet ten opzichte van de placebo groep. Deze studie laat daarmee zien dat rHuEPO in getrainde wielrenners geen effect heeft op wielerprestatie in een tijdrit of een wegwedstrijd bergop. Daarnaast toont de studie aan dat het mogelijk is om doping te onderzoeken volgens de bestaande principes van de klinische farmacologie, daarmee een wetenschappelijke basis leggend voor anti-doping activiteiten voor andere middelen.

Behalve het effect van een middel op de prestatie is ook kennis over het mogelijke gezondheidsrisico voor de sporter van belang. Volgens de WADA *Code* kan een risico op de gezondheid namelijk ook een reden zijn een middel op de dopinglijst te zetten, ter bescherming van de sporter. Daarnaast schrikt voorlichting over bekende gezondheidsrisico's potentiele gebruikers misschien af. Net als voor het farmacologische effect op prestatie is de juiste methode voor het verkrijgen van bewijs van een dergelijk nadelig effect de besproken robuuste klinische studie. Daarom konden we in dezelfde studie onderzoeken of de behandeling met rHuEPO een risico geeft op gezondheidsproblemen. In onze studie zien we geen bijwerkingen van de behandeling, maar wel geeft rHuEPO een verhoging van de bloedmarkers E- en P-selectine, die duiden op een mogelijk verhoogd risico op vaatproblemen. Daarnaast laat onze studie zien dat intensieve inspanning, zoals wielrenners dat doen, effect heeft op een grote hoeveelheid markers betrokken bij stolling en de functie van de vaatwand. De combinatie van rHuEPO en inspanning geeft zelfs een extra stijging van enkele markers. Alhoewel het risico nog altijd erg laag is, geeft dit wel aanwijzing voor een mogelijk verhoogd risico op vaatproblemen zoals trombose bij rHuEPO gebruik, mogelijk verergerd wanneer het wordt gecombineerd met inspanning.

Ook deden we in dezelfde studie een uitgebreide analyse van een van de surrogaat markers voor prestatie, de zogenaamde lactaatdrempel. Door beter inzicht te krijgen in de reproduceerbaarheid, betrouwbaarheid en relevantie van dergelijke surrogaat markers, en de beperkingen ervan in kaart te brengen, kan de waarde en de relatie met daadwerkelijke sportprestatie van een dergelijke uitkomstmaat beter worden beoordeeld. Voor het bepalen van de lactaatdrempel blijken vele verschillende methodes, en onze analyse laat zien dat niet al deze methodes even goed zijn. Maar zelfs de betere methodes zijn maar deels voorspellend voor sportprestatie, een bevinding die het nadeel van het gebruik van surrogaat markers onderstreept.

Het zal daarmee duidelijk zijn dat er voor de meeste middelen op de dopinglijst een tekort aan informatie en een gebrek aan bewijs is over de effecten op sporters. Dit belemmert de voorlichting aan sporters, en bemoeilijkt het beleid van anti-doping inspanningen. Door systematischer onderzoek te doen naar doping, gebruikmakend van goed-gecontroleerde studies, kunnen deze tekortkomingen beter aangepakt worden, zoals we in dit proefschrift laten zien. Daarnaast kan dat ook bijdragen aan een beter onderbouwde, op bewijs gebaseerde (en dus mogelijk kortere) dopinglijst. Dat kan vervolgens invloed hebben op doping controles, omdat er meer gericht gewerkt kan worden aan de middelen die daadwerkelijk het stempel doping verdienen. En dat een grotere aandacht voor de detectiemethodes nodig is, laten de laatste hoofdstukken van dit proefschrift zien. De detectiemethodes en bijbehorende procedures die we onderzochten hebben tekortkomingen, waardoor ze de sporter die doping gebruikt lang niet altijd zullen opsporen, en daarnaast niet altijd de onschuldige sporter zullen beschermen. De dopingassays voor rHuEPO die momenteel worden voorgeschreven door de WADA konden we evalueren door in de eerder beschreven studie met rHuEPO urine en bloedmonsters af te nemen en te analyseren bij een geaccrediteerd doping lab volgens officiële WADA procedures. Alhoewel er na het uitvoeren van een tweede, zogenaamde bevestigingstest, geen vals positieve bevindingen waren voor rHuEPO in urine, zijn er onzekerheden over waarom enkele placebo-monsters bij de primaire analyse wel positief waren. Daarnaast blijkt het detectievenster voor rHuEPO behoorlijk nauw en werd een substantieel deel van de urines dus vals negatief gerapporteerd. Dat betekent dat gebruikers maar een relatief korte periode het risico lopen positief te testen. Een andere detectieprocedure van de WADA die we onderzochten, voor het bepalen

van overschrijdingen van de toegestane dosis salbutamol, heeft fundamentele tekortkomingen. We laten in onze studie zien dat met deze procedure het in wezen onmogelijk is om met zekerheid te zeggen of een sporter een overtreding van de dopingregels heeft begaan, en ook hier zullen dus schuldige sporters vrijuit gaan en onschuldige sporters soms beschuldigd worden. De aanpak van doping controles is dus ook niet volledig op orde, en een betere wetenschappelijke basis zal bijdragen aan het verhelpen van deze tekortkomingen in de strijd tegen doping. Samengevat leiden de genoemde onzekerheden en onduidelijkheden van zowel de effecten als de detectie van dopingmiddelen tot een anti-doping systeem dat niet voldoende transparant, consistent en betrouwbaar is. En dat is uiteindelijk weer schadelijk voor de sporters, toeschouwers en de sport.

De enige conclusie die op basis hiervan daarom mogelijk is, is dat sportgeneeskunde en anti-doping werk in het bijzonder, net als de geneeskunde al heeft gedaan, een meer bewijs-gedreven weg moeten inslaan. Met goed opgezette klinische studies en gedegen wetenschappelijk onderzoek kunnen het bewijs en de onderbouwing worden geleverd die een beter gestructureerd anti-doping systeem verwezenlijken. Hiermee zullen anti-doping inspanningen efficiënter, effectiever en uiteindelijk eerlijker worden, en dat komt de sporters, de toeschouwers, en dus de sport, ten goede.

LIST OF PUBLICATIONS

Heuberger JA, Cohen Tervaert JM, Schepers FM, Vliegenthart AD, Rotmans JI, Daniels JM, Burggraaf J, Cohen AF. Erythropoietin doping in cycling: lack of evidence for efficacy and a negative risk-benefit. *Br J Clin Pharmacol* 2013; 75: 1406-21.

Heuberger J, Schmidt S, Derendorf H. When is protein binding important? *J Pharm Sci* 2013; 102: 3458-67.

Sy SK, **Heuberger J**, Shilbayeh S, Conrado DJ, Derendorf H. A Markov chain model to evaluate the effect of CYP3A5 and ABCB1 polymorphisms on adverse events associated with tacrolimus in pediatric renal transplantation. *AAPS J* 2013; 15: 1189-99.

Kleinloog D, Stevens J, **Heuberger J**, Spinhoven P, van GJ. The influence of personality on the sensitivity to subjective effects of Delta9-tetrahydrocannabinol. *Psychiatry Res* 2014; 220: 945-53.

Heuberger JA, Guan Z, Oyetayo OO, Klumpers L, Morrison PD, Beumer TL, van Gerven JM, Cohen AF, Freijer J. Population pharmacokinetic model of THC integrates oral, intravenous, and pulmonary dosing and characterizes short- and long-term pharmacokinetics. *Clin Pharmacokinet* 2015; 54: 209-19.

Vilums M, **Heuberger J**, Heitman LH, IJzerman AP. Indanes-Properties, Preparation, and Presence in Ligands for G Protein Coupled Receptors. *Med Res Rev* 2015; 35: 1097-126.

Guan Z, Klumpers LE, Oyetayo OO, **Heuberger J**, van Gerven JM, Stevens J. Pharmacokinetic/pharmacodynamic modelling and simulation of the effects of different cannabinoid receptor type 1 antagonists on Delta(9)-tetrahydrocannabinol challenge tests. *Br J Clin Pharmacol* 2016; 81: 713-23.

Peters E,* **Heuberger JA***, Tiessen R, van EA, Masereeuw R, Arend J, Stevens J, Pickkers P. Pharmacokinetic Modeling and Dose Selection in a Randomized, Double-Blind, Placebo-Controlled Trial of a Human Recombinant Alkaline Phosphatase in Healthy Volunteers. *Clin Pharmacokinet* 2016; 55: 1227-37.

Sy SK, Beaudoin ME, Zhuang L, Loblein KI, Lux C, Kissel M, Tremmel R, Frank C, Strasser S, **Heuberger JA**, Mulder MB, Schuck VJ, Derendorf H. In vitro pharmacokinetics/pharmacodynamics of the combination of avibactam and aztreonam against MDR organisms. *J Antimicrob Chemother* 2016; 71: 1866-80.

Heuberger J, Rotmans JI, Gal P, Stuurman FE, van 't Westende J, Post TE, Daniels JMA, Moerland M, van Veldhoven PLJ, de Kam ML, Ram H, de Hon O, Posthuma JJ, Burggraaf J, Cohen AF. Effects of erythropoietin on cycling performance of well trained cyclists: a double-blind, randomised, placebo-controlled trial. *Lancet Haematol* 2017; 4: e374-e86.

van Amerongen G, Kanhai K, Baakman AC, **Heuberger J**, Klaassen E, Beumer TL, Strijers RL, Killestein J, van Gerven J, Cohen A, Groeneveld GJ. Effects on Spasticity and Neuropathic Pain of an Oral Formulation of Delta9-Tetrahydrocannabinol in Patients With Progressive Multiple Sclerosis. *Clin Ther* 2017.

Heuberger J*, van Dijkman SC*, Cohen AF. Futility of current urine salbutamol doping control. *Br J Clin Pharmacol* 2018.

Kovalchuk MO*, **Heuberger JA***, Sleutjes BT, Ziagkos D, van de Berg LH, Ferguson TA, Franssen H, Groeneveld GJ. Acute effects of riluzole and retigabine on axonal excitability in patients with ALS: A randomized, double-blind, placebo-controlled, cross-over trial. *Clin Pharmacol Ther* 2018.

Muehlan C, **Heuberger J**, Juif PE, Croft M, van Gerven J, Dingemanse J. Accelerated Development of the Dual Orexin Receptor Antagonist ACT-541468: Integration of a Microtracer in a First-in-Human Study. *Clin Pharmacol Ther* 2018.

Heuberger J, Cohen AF. Review of WADA Prohibited Substances: Limited Evidence for Performance-Enhancing Effects. *Sports Med* 2018.

Heuberger J, Gal P, Stuurman FE, de Muinck Keizer WAS, Mejia Miranda Y, Cohen AF. Repeatability and predictive value of lactate threshold concepts in endurance sports. *PLoS One* 2018; 13: e0206846.

Birkhoff WAJ, **Heuberger J**, Post TE, Gal P, Stuurman FE, Burggraaf J, Cohen AF. Recombinant human erythropoietin does not affect several microvascular parameters in well-trained cyclists. *Physiol Rep* 2018; 6: e13924.

Kletzl H, Marquet A, Gunther A, Tang W, **Heuberger J**, Groeneveld GJ, Birkhoff W, Mercuri E, Lochmuller H, Wood C, Fischer D, Gerlach I, Heinig K, Bugawan T, Dziadek S, Kinch R, Czech C, Khwaja O. The oral splicing modifier RG7800 increases full length survival of motor neuron 2 mRNA and survival of motor neuron protein: Results from trials in healthy adults and patients with spinal muscular atrophy. *Neuromuscul Disord* 2019; 29: 21-29.

CURRICULUM VITAE

Jules Heuberger was born on March 15th, 1989 in Maastricht, The Netherlands. After completing secondary school at the *Trichter College / Bonnefanten College* in Maastricht in 2007, he studied Bio-Pharmaceutical Sciences at Leiden University. He halted his studies for one year to take the position of Commissioner of General Affairs of the Royal Dutch Pharmacy Student Association (K.N.P.S.V.) in 2009. During his master's he performed a pharmacometrics internship at the Centre for Human Drug Research (CHDR) and an internship at the department of Pharmaceutics of the University of Florida in Gainesville, Florida, US. He obtained his doctoral degree (*cum laude*) in 2013 and started working as a clinical scientist at CHDR. In this position under supervision of dr. G.J. Groeneveld his focus was on clinical studies related to disorders of the central nervous system, with a special interest in neuromuscular disorders. During this period at CHDR he also performed the research described in this thesis under supervision of prof. dr. A.F. Cohen. Since 2018 he holds a position as senior clinical scientist at CHDR in the neurology and pain group. In 2019 he became a board certified clinical pharmacologist. Jules lives together with his partner Evelien.