SportRXiv

Part of the Society for Transparency, Openness and Replication in Kinesiology (STORK)

Preprint
not peer reviewed

Understanding Training Load as Exposure and Dose

For correspondence:
franco.impellizzeri@uts.edu.au

Franco M. Impellizzeri[1], Ian Shrier[2], Shaun J. McLaren[3,4], Aaron J. Coutts[1], Alan McCall[1,5], Katie Slattery[1], Annie C. Jeffries [1], Judd Kalkhoven [1]

[1], Human Performance Research Centre, School of Sport, Exercise and Rehabilitation, Faculty of Health, University of Technology Sydney, NSW, Australia; [2], Centre for Clinical Epidemiology, Lady Davis Institute, Jewish General Hospital, McGill University, Montreal Canada; [3], Newcastle Falcons Rugby Club, Newcastle upon Tyne, United Kingdom; [4], Department of Sport and Exercise Sciences, Durham University, Durham, United Kingdom; [5], Arsenal Performance and Research Team, Arsenal Football Club, London, United Kingdom

Please cite as:

Impellizzeri FM, Shrier I, McLaren JS, Coutts AJ, McCall A, Slattery K, Jeffries AC & Kalkhoven J. (2022). Understanding training load as exposure and dose. *SportRχiv*.

All authors have read and approved this version of the manuscript. This article was last modified on August 15, 2022.

Authors FMI @francoimpell, IS @IanShrier, SM @Shaun_McLaren1, AJC @AaronJCoutts, AM @Alan_McCall, @KS @ KatieSlattery3, ACJ @ AnnieJeffries6 and JK @ KalkhovenJudd can be reached on Twitter.

ABSTRACT

Various terms used in sport and exercise science, and medicine, are derived from other fields such as epidemiology and pharmacology. Conceptual and nomological frameworks have described training load as a multidimensional construct manifested by two causally related subdimensions: external and internal training load. In this article, we explain how the concepts of training load and its subdimensions can be aligned to classifications used in occupational medicine and epidemiology, where exposure can also be differentiated into external and internal dose. The meanings of these terms in epidemiology are explored from a causal perspective, and these terms and their underlying concepts are contextualised to the physical training process.

We also explain how these concepts can assist in the validation process of training load measures. Specifically, to optimise training (i.e., within a causal context), a measure of exposure should be reflective of the mediating mechanisms of the primary outcome. Additionally, understanding the difference between intermediate and surrogate outcomes allows the correct investigation of the effects of measures of exposure and their interpretation in research and applied settings. Finally, whilst the dose-response relationship can provide evidence of the validity of a measure, conceptual and computational differentiation between causal (explanatory) and non-causal (descriptive and predictive) dose-response relationships is needed.

Regardless of how sophisticated or “advanced” a training load measure (and metric) appears, in a causal context, if it cannot be connected to a plausible mediator of a relevant response (outcome), it is likely of little use in practice to support and optimise the training process.

1. Introduction

Various terms used in sport and exercise science, and medicine, are derived from epidemiology and other biomedical sciences, such as pharmacology. For example, many studies commonly utilise terms such as exposure and dose when referring to training and exercise.[1-4] Within the various areas of epidemiology and pharmacology, exposure and dose have been conceptualised and classified in many ways. Understanding these concepts is important for the selection of appropriate measures of exposure and dose, as well as the exploration of other related concepts such as the dose-response relationship. In this article, the meanings of various common terms within the physical training vernacular are reconciled with similar classifications and conceptualisations used in epidemiology and pharmacology.

The purpose of this article is to acknowledge recent discussions regarding the use of training load-related terminology and concepts including the interpretation of training load as dose,[5-9] and provide further clarification of the meaning and role of these terms within the training process framework by drawing parallels with epidemiology. Training can have various effects, and understanding what measures of exposure to use and the appropriate metrics (i.e., cumulative, volume, average and peak intensity, overall pattern, etc.), depends on the goal of the training, the responses of interest (primary outcomes), and the relevant aspects of training that the practitioners or researchers are interested in. Due to the numerous outcomes of interest (performance or health-related), each with their own set of mediators, there *cannot* be a single measure that reflects all the mediators of the various outcomes. We therefore also explain how key concepts presented in epidemiology can assist in the validation of measures of training load when the aim is to optimise the training process. In this context, "optimise" (and optimisation) specifically refers to the manipulation or variation of physical training parameters to obtain better outcomes (i.e., within a causal context).

2. Exposure and dose

Epidemiologists use different definitions of exposure (descriptive and operational) in different contexts. For the current article, we refer to the exposure definitions presented in a reference dictionary of epidemiology (Table 1).[10] At least two of these definitions can be contextualised to the physical training process proposed for sport and exercise science.[11,12]

As per definition 1 (Table 1), exposure is a variable whose causal effect is to be estimated, regardless of whether this variable is external or internal to the body. In other words, exposure refers to any variable of interest that might cause the outcome.[13] These variables can include behaviours, treatments, interventions, hazards, and traits,[14,15] as well as variables typically used in exercise physiology such as body mass index, heart rate, blood pressure, serum glucose and so on. In sport and exercise, we commonly want to estimate the causal effect of training on sporting performance or health-related outcomes.

Accordingly, training load has been also defined as the input variable that is manipulated to elicit the desired training response, which identifies training as the causal variable (as per definition 1).[16] Therefore, in the physical training context, "training" is the variable of interest, i.e., the exposure. However, this definition does not specify the scale of the variable.[17] By adding the term "load" (which indicates an amount) to "training", it is implied that we are generally referring to a continuous or at least ordinal, instead of categorical (e.g., yes/no) variable. It follows that, when applying definition 1 to physical training, "training load" acts as a generic term referring to the amount of the exposure variable irrespective of whether this exposure is external or internal to the body. However, within the physical training process framework, training load has been sub-categorised into internal and external (training) load,[14,15] the implications of which will be discussed in more detail in following sections.

Unlike definition 1, definition 3 (Table 1) more explicitly states that exposure is considered to be as a continuous (or ordinal) variable: "the *amount* of a factor to which individuals are exposed".[10] Exercise, physical activity and diet, and even events such as heading the ball in football or physiological/biomechanical constructs, are factors acting on the body i.e., factors to which individuals are "exposed".[18,19] This definition is arguably more consistent with the training load concept used in sports science as the term "training" acts as a qualifier identifying the factor of interest to which individuals are exposed, and the term "load" indicates the amount of this factor.[11,12,20] As per definition 1, when applying definition 3 to physical training, "training load" includes any measure of load without reference to whether it is external or internal to the body (i.e., whether it is a measure of external or internal load).

Other definitions presented in Table 1 along with alternative definitions that can commonly be found in the literature and relevant textbooks, often use the term "agent" interchangeably with "factor" and "variable". "Agent" is a generic term used to indicate substances (e.g., pollution and drugs) but can be extended to any other attributes. For example, exposure has been defined by Cordier and Stewart[21] "as a contact of an individual with an agent through any medium or environment." These authors clarified that exposure can be a chemical, biological, physical, or societal agent in the external environment, or characteristics of an individual (including weight and physical activity), susceptibility, exercise, diet and any other external or internal agent.[21] Others extend the definition to include any activity or action that can be directly or indirectly measured.[10] Accordingly, the epidemiological concepts and definitions of agents (or factors) have already been legitimately applied and adapted to physical activity and exercise.

Table 1. Definitions of exposure[10]

Def 1.	The variable whose causal effect is to be estimated.
Def 2.	Proximity and/or contact with a source of a disease agent in such a manner that effective transmission of the agent or harmful effects of the agent may occur.

Def 3.	The amount of a factor to which a group or individual was exposed, sometimes contrasted with dose, the amount that enters or interacts with the organism.
Def 4.	The process by which an agent comes into contact with a person or animal in such a way that the person or animal may develop the relevant outcome, such as a disease.

Within Table 1, definition 3 introduces another important concept: dose. Exposure and dose are often used interchangeably or even combined (e.g., dose of exposure) depending on the scientific context. However, in some areas such as environmental and occupational epidemiology, the concept of dose is differentiated from exposure. Definition 3 specifically states that exposure is sometimes contrasted with dose, where dose is defined as "the *amount* of a substance [agent or factor] that enters or interacts with the organism."[10] The reason for this contrast is that it is also common to think of exposure as an "agent in the external environment".[22] The key aspect of these dose definitions is that they refer to the actual stimulus "that enters or interacts with the body", which is what triggers biological responses and hence adaptations. In some contexts, causal effects occur when a dose exceeds a certain threshold. In other contexts, causal effects may increase as the dose increases, or be dependent on the total amount of exposure.

3. External and internal dose

The definitions for exposure and dose above can be conceptually applied to physical training in a sport and exercise context. However, these definitions do not clearly reflect the three elements of the training process framework, which include training load, external (training) load and internal (training) load (Figure 1).[12]

In a reflective model for the physical training process, training load is a multidimensional construct that is manifested by the two causally related sub-dimensions: external and internal load (Figure 1). Within (environmental and occupational) epidemiology and pharmacology (e.g., toxicology), there is a parallel classification of exposure, where the generic term "exposure" encompasses two causally related subdimensions. In this classification internal dose (amount individuals internally absorb, i.e., systemically available) [23-25] is a mediator of the effect of external dose (amount individuals are physically exposed) on the outcome/response (Figure 1).[24] The parallel between these classifications is therefore quite straightforward. Training load represents the exposure (overarching construct) where its two causally related sub-dimensions, external and internal training load (manifestations of the multidimensional construct), act as conceptual equivalents to external and internal dose. Training load/exposure can therefore be quantified by measures reflecting its two manifestations, i.e., by using measures reflecting external and/or internal (training) load/dose.

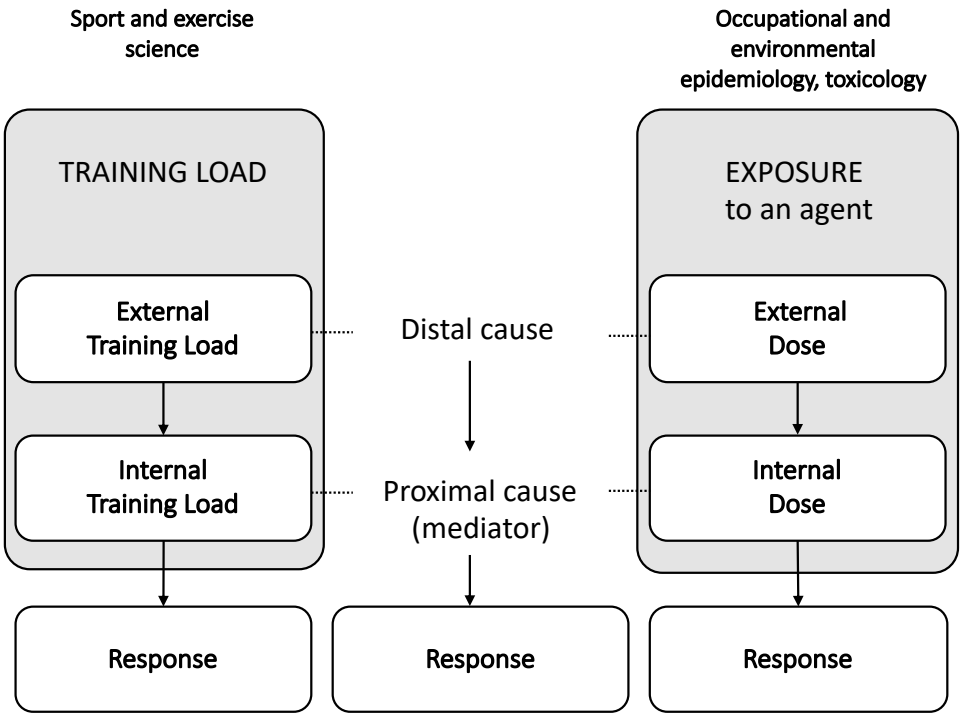

Figure 1: Diagram depicting the parallels exhibited between terms used in sports science and certain areas of epidemiology and pharmacology. The left and right diagrams represent classifications where the training load and exposure include two different but related subdimensions (external and internal load and dose). The central section provides the role of the two subcomponents in the causal pathway leading to a response. The internal dose and load are more proximal causes of the outcome of interest, i.e., they are mediators of (or mechanisms for) the effect of the external dose/load on the outcome.

While the structure[11,12] of the training load classification may appear "artificial" at first, it actually aligns with available classifications used in other fields and it is a relatively simple bi-dimensional construct where the directionality of the causal relationship of its sub-dimension is intuitive. This may explain why this classification and related framework is widespread in sports and exercise science. This structure is helpful in both practice and research settings to 1) understand how and what components of the training process we can measure, 2) provide a reference framework for the validation of the measures of external and internal load, and 3) help to understand the role of specific measures (and the characteristic they reflect) in the causal pathway leading to the response/outcome of interest. The external dose is therefore the amount of training (agent/factor) to which individuals are exposed, and is quantified using measures reflecting the amount of physical training (activities and actions) performed by athletes.[11,12,20] The internal dose is the internal (to the body) training load that an individual experiences to cope with the demands

imposed by the external load, and is quantified using measures reflecting the psycho-physiological stress during the training/exercise, or any other internal load measures. In epidemiology and pharmacology, external dose is expected to cause different internal doses between and within individuals as internal dose is influenced by factors such as genetics, metabolism, susceptibility, changes in states, etc.[18,26] Similarly, within the physical training process, a given external (training) load can correspond to different internal loads between and within individuals at different times.[11,12]

The central section of Figure 1 illustrates the additional terminology of proximal and distal causes that is sometimes used to describe the causal relationships between variables. This simple causal structure is also important from a practical perspective because it emphasises the mediating role of the internal (training) load and it highlights that to alter the internal load, we need to manipulate the external load (or the environment which constitutes an external, to the body, factor/agent). For the remainder of this article, we will use the term "exposure" and "training load" when referring generically to both (1) external dose or load and (2) internal dose or load concepts. We will use the terms external dose/load and internal dose/load as per the aforementioned definitions.

3.1 Why not use exposure and dose instead of load?

According to what has been presented in the previous section, the use of the terms "training exposure", "external training dose" and "internal training dose" (i.e., "exposure" and "dose" instead of "load") would be reasonable and maybe even technically appropriate. However, the term "load" is firmly entrenched within the sports science literature, with the first reports using the term training load to indicate the amount of training presenting back in the '80s,[27-29] and a vast quantity of research utilising the term presenting thereafter. We believe that the continued use of the term "load" in sports science is appropriate as this term communicates the amount of something and concurrently conveys the idea of a demand imposed on the organism. These connotations have contributed to the successful adoption of "load" in other scientific areas (e.g., allostatic load, cognitive load, etc). While the meaning of the term "load" may vary between contexts, polysemy is common and due to the longstanding use of the term load in sports science, we find it very unlikely that practitioners and sports scientists will confuse the meaning of this term with alternative meanings found in other contexts (e.g., mechanical).[8] Accordingly, we maintain that keeping and continuing to use "load" instead of "exposure" and "dose" is reasonable. Ultimately, however, it is up to the researchers and the scientific community whether they use "load" or "exposure/dose" as either is acceptable.

4. Identifying the measures of exposure (external and internal dose)

Every exposure includes multiple components, where different components may affect different outcomes. If we are interested in understanding how manipulating an exposure will affect a specific

outcome (i.e. within a causal context), the measure of exposure (either external or internal) should reflect, directly or indirectly, the component of the exposure that causes (i.e. the mechanism or mediator of) that particular outcome of interest.[13,22] For example, let's assume a researcher (or practitioner) is interested in understanding how to manipulate small-sided games to obtain improvements in sprinting (hypothesised to be one of the determinants of match performance). The training load of small-sided games can be quantified in a variety of manners including measures of external load (e.g., time or distance covered running at various speeds, number of accelerations, etc.), and internal load (e.g., heart rate, lactate or perceived exertion, etc.). Therefore, if it is considered that the neuromuscular stimulus provided by the small-sided games is the mediating mechanism that improves sprinting ability, we need a measure of exposure that reflects this neuromuscular stimulus. However, the presented examples of internal load measures cannot reasonably be expected to reflect the neuromuscular stimulus. Consequently, a measure of external load may be used as a substitution (surrogate) of the internal mediating mechanism (neuromuscular stimulus). For example, if it is hypothesised that running above a certain percentage of maximal speed provides an adequate reflection of the neuromuscular stimulus mediating the response of interest, we can select this external load measure. Similarly, if it is anticipated that the accelerations completed in a certain range adequately reflect the neuromuscular stimulus, we can select this measure as an alternative (or additional) measure. Ideally, the appropriateness of a measure of external training load as a reflection of the neuromuscular stimulus should be supported by evidence from studies. In scenarios where a lack of research exists, the selection of a measure should at least be based on a well-thought-out hypothesis linking the measure to the mechanistic stimulus. Examining the extent to which measures of training load reflect the components causing the mediating mechanisms of an outcome is the goal of the validation process. While the mechanisms reflected by new measures or metrics are often inferred by their label, unfortunately, many labels appear to be selected for marketing purposes rather than being mechanistically motivated and validated (see section 7 on the validation process).

To summarise the above, in sport and exercise, if a measure of training load (external and internal load) is to be used for the purpose of optimising training, this measure should be reflective of the mediating mechanisms through which the specific physical training (factor/agent) is anticipated to cause the effects/responses of interest; i.e., changes in sporting performance or health-related outcomes. We again highlight that the term "optimise", very commonly used in sports science, implies the active manipulation of the external load to generate a specific internal load that ultimately influences (in a positive way) the outcome. Accordingly, optimisation implies causation, and therefore a causal exposure-outcome framework (known or hypothesised) is needed to guide the identification of the measure of exposure.[13,19] Basic (e.g., physiological) research and frameworks are useful in guiding our understandings of potential mediating mechanisms and the corresponding measures. An example of applying this process to training load and hypothesised mechanisms of injury is provided in two recent papers.[20,30] These frameworks can also be utilised to inform and develop causal directed acyclic graphs,[31] which can then be used to guide appropriate statistical analyses. Importantly, these frameworks are proposals and revising or replacing

frameworks when updated knowledge is presented is a normal step in the scientific process. Regardless, frameworks have an important role in presenting causal assumptions transparently, and avoiding or limiting ad hoc and post hoc explanations, thus reducing bias and cherry picking.

5. Metrics

Once an appropriate measure of exposure is identified, we need to determine the relevant dimensions (metrics) for the quantification.[22] There are a variety of possible operationalisations leading to different metrics and the choice depends on the exposure/dose-response process of interest.[32] Typical (static) metrics in epidemiology are average, peak, duration and cumulative exposure (CE).[26,32] In this section, we elaborate on CE because similar approaches are commonly used within sports science and medicine. Within this metric, exposure (external and internal dose) is usually operationalised according to two dimensions: intensity (I) and duration (T).[26] Duration is the time period during which the amount of a substance (or training, in a sports context) is delivered. Intensity, according to Checkoway et al.,[26] "represents the magnitude of the amount of a substance that potentially can enter the body and be delivered to the biological target(s)." It can also be measured as a rate i.e., "the rate at which a substance is brought in contact with the body".[33] Mathematically, the rate is simply the slope of the magnitude versus duration relationship, which may or may not vary over time depending on the context.

CE is calculated as CE = I (intensity) x T (duration), or more generically, if intensity varies over time, as a time integral of exposure intensity depicted in equation 1:

$$CE = \int_{t1}^{t2} I(t)\, dt \qquad \text{(eq. 1)}$$

where the duration is the interval (t_1, t_2), *I* is intensity and *t* is time. The formula for the cumulative dose is obtained by substituting *I* with I_D (dose intensity, where dose refers to internal dose)[33].

An easier way to present the same method is shown in equation 2[26]:

$$CE = \sum_{i=1}^{N} C_i \cdot t_i \qquad \text{(eq. 2)}$$

where C represents the exposure level (i.e., intensity) and *t* corresponds to each time interval as shown in Figure 2 where the sum ($\sum$) of the black rectangles, one for each time interval (from t_1 to t_{20}) corresponds to the CE.

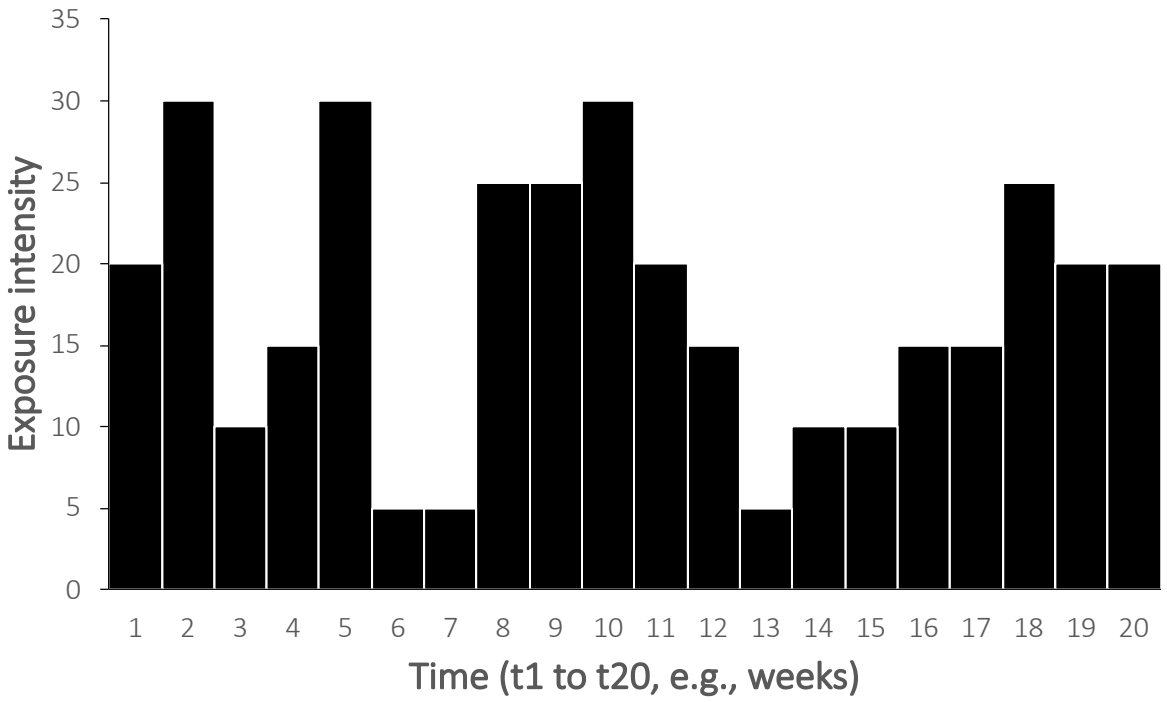

Figure 2. Hypothetical exposure history from the start at, t_1, to the end, t_{20} (adapted from Checkoway et al.[26]). The sum ($\sum$) of the black rectangles, one for each time interval (from t_1 to t_{20}) corresponds to the measure of cumulative exposure.

This method for CE resembles the manner in which training load has historically been quantified. For example, common measures such as training impulse (TRIMP)[34-36] and session-Rating of Perceived Exertion[37] are based on the multiplication of intensity (eventually time-weighted) by duration. For the American College of Sports Medicine, the weekly volume formula (for aerobic activities) is the product of frequency, session duration and intensity.[38] In resistance training, volume is sometimes defined as the product of sets and repetitions, or the product of set, repetition and load lifted (absolute or relative).[39-41] These variations are all reasonable measures of cumulative exposure and their choice depends on the practical or research goal.

Recent opinion articles have highlighted that combining exercise intensity and duration to form a single cumulative metric has limitations.[7,42,43] Many of these limitations have also already been acknowledged in epidemiology.[44] To elaborate, this approach provides an estimate that represents the average effect across all individuals with the same cumulative exposure, and does not take into account the pattern of exposure over time. As a practical example, consider the exposure occurs over time as it does in smoking. Using a simplified summary measure of the dose over time (e.g., pack-years of smoking) or any cumulative dose over a recent period, we obtain an average estimate across individuals who smoke 1 pack/day for 1

year and individuals who smoke ½ pack/day for 2 years.[42] If we are interested in the differences between these two groups, then such a simplification is not justified and we might consider using intensity, duration *and* their interaction as separate variables in one statistical model.[43,44] More complicated questions might require biological models of (disease) processes that are dynamical in nature (time-varying) and account for the response of biological systems to changes in exposure and internal conditions (for more details see *Chapter 16* in Smith et al.[45]).

It follows that the cumulative measures of exposure and any other metrics used in epidemiology and in sport and exercise cannot be considered right or wrong *per se*, as their validity and appropriateness depend on specific details such as what aspects of the exposure the practitioners or researchers are interested in, the acceptability of the limitations of the metric, and whether the assumptions on which the metric are based are reasonable and eventually testable. There are instances where the simple duration of the exposure is the only variable at our disposal, and while limited, this information may still be appropriate and meaningful in many contexts.[46]

6. Dose-response

Dose-response (sometimes referred to as exposure-response depending on the context) is another important concept common in epidemiology that describes the relationship between the amounts of a factor/agent (i.e., dose) and the responses (e.g., risk of an outcome). This relation can be monotonic, which in an epidemiological context means the risk of a disease always increases as the intensity or duration of exposure increases,[47] or it can exhibit other patterns.[48] In sport, a well-known theoretical dose-response function is the parabola (or "inverted U" shape). This function is based on the concept that performance improves with increases in training load up until a maximum level is reached (vertex), and beyond this point, further increases in training load result in a deterioration of performance (overreaching and overtraining).[49] Other more complex methods of modelling the cumulative effects of time-dependent exposures (such as training load) on continuous and binary outcomes have been proposed in epidemiology, and these models can also be applied to sport.[48,50,51]

6.1. Defining the response (outcome) of interest: surrogate and intermediate outcomes

In sport and exercise or medicine, there are two main responses of interest: performance and health-related outcomes. When measures of these two main responses are unavailable, surrogate outcomes or intermediate responses that are considered to be determinants of performance or health-related outcomes, may be utilised. For example, we may be interested in the effect of aerobic training (or any other intervention) on cycling performance, and we use a simulated time trial as a surrogate of actual on-road performance. Or we may want to examine the relationship between an appropriate measure of resistance training (external or internal dose/load) and strength under the assumption that a higher level

of strength affects (i.e., mediates) sporting performance or health-related risks (outcome/responses of interest). In this case, strength is a mediator of the effect of the exposure (resistance training) on performance or health and can be used as an intermediate outcome.

A surrogate *outcome* is especially useful when the primary outcome (commonly termed "endpoint" in clinical research) is difficult to measure or requires a very long follow-up, impacting the feasibility and costs of studies.[52-56] This is also applicable to sports science and medicine as actual performance or health-related endpoints are often difficult to measure, in both research and applied setting. However, it is important to distinguish between surrogate and intermediate outcomes[57] because their interpretation and implications differ.

Consider a surrogate outcome (S) for the primary outcome Y. By knowing the effect of the exposure (A, e.g., training) on S, we can predict the effect of A on Y, which is the response of interest (e.g., performance).[52-54] To be considered a strong surrogate outcome (i.e., as *substitution* of the primary outcome), S should *fully* (or *mostly*) mediate Y.[55] This is shown by the directed acyclic graph (DAG) presented in Figure 3a, where S is the only mediator in the causal path. However, although a scenario where S fully mediates Y is ideal, it is rare. A more realistic DAG is represented in Figure 3b, where there are unmeasured confounders (U) of the relation between S and Y and another causal path may exist between A and Y, independent from S. This other path can be mediated by other mechanisms (M_1 to M_3) as shown in DAG 3c, where S is just one of various mediators used as intermediate outcomes (M_4) and not a surrogate outcome. The DAGs in Figure 3 are simplified scenarios presented for educational purposes. The causal structures in studies are often much more complicated and can include additional confounders, colliders and effect modification.

The previous paragraph introduced some important considerations. In a scenario such as DAG 3c, if we measure only the effect of A on the intermediate outcome (M_4) and not the other mediators (M_1 to M_3), no accurate inferences can be made about the magnitude of the overall effect of A on Y. To emphasise the importance of accounting for all mediators, scenarios such as presented in DAG 3c can lead to the so-called surrogate paradox. This paradox occurs when other mediators besides S (M_4 in DAG 3c) exist (S is incorrectly identified as a surrogate outcome when it is actually an intermediate outcome) and the effects of the other mediators influencing Y are in the opposite direction compared to the effect of S on Y. For example, the effect of A on S (M_4) may increase Y, but the effect of A on M_1/M_2/M_3 may decrease Y. If we incorrectly considered S to be appropriate surrogate for Y, we would conclude A is beneficial when the true overall effect of A on Y is harmful, resulting in a paradox (false positive result).[52,54,58] It follows that a surrogate outcome should be *consistent*, meaning that a change in a particular direction of S should always reflect a directionally consistent change in the primary outcome (Y).[52] To illustrate this with an applied sporting example, a practitioner or researcher can find that an exposure has a positive effect on strength, which is assumed to be the main mediator of a given improvement in performance outcome. However,

the actual measured effect of A on performance outcome suggests A negatively affects performance. If we know, however, that strength is only one of the many mediators, there will not be expectations regarding the direction of the changes in the performance outcome, i.e., we are simply examining the effects of training on one mediator. It is clear that understanding or hypothesising the role of the measures of exposure and response (outcome) in the causal pathway drastically changes their interpretation.

To further clarify, using the above-mentioned example in sport, this means that we need studies showing that the time trial (surrogate outcome) is predictive of the actual (on-road) performance of a cyclist or that strength (mediator used as intermediate outcome) is causally related to the real performance of interest. However, in sports and exercise, these assumptions are commonly speculative, and some evidence is needed to support them.

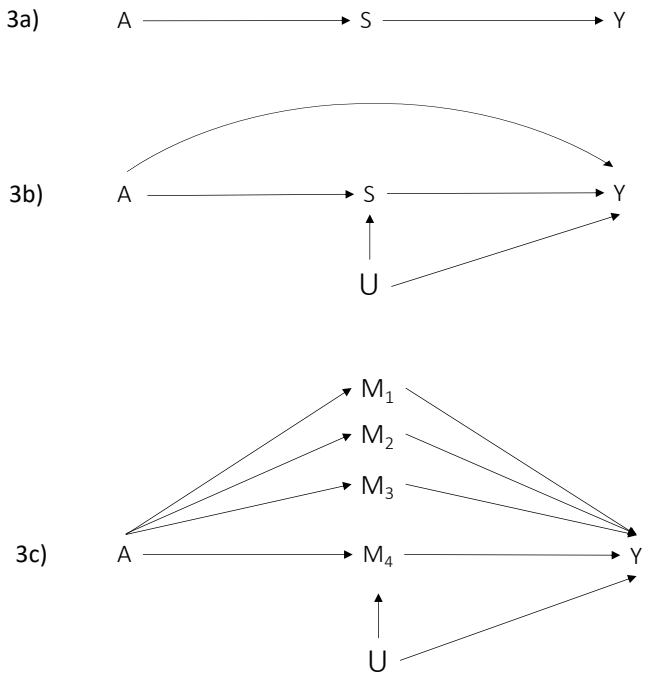

Figure 3. Directed acyclic graph representing 3a) surrogate outcome (S) fully mediating the effect of A (e.g., training) on the response of interest, Y (e.g., performance), 3b) surrogate outcome (S) with an unmeasured confounder (U) of the S – Y path and another causal path between A and Y, independent from S, and 3c) hypothetical additional mediators (M_1 to M_3) of the A-Y causal path, where the surrogate outcome is actually just another mediator (M_4) that can be used as an intermediate outcome.

7. Validity of the measures of exposure or dose

As explained in section 4, within a causal context (i.e., when we want to manipulate the exposure to alter the outcome), the measure of exposure should reflect a component that is the mediating mechanism for the exposure-outcome causal path. Consider the question posed in section 4: Does a measure of the time spent running above a given percentage of maximal speed, during small-sided games, provide an appropriate reflection of the neuromuscular stimulus experienced to improve sprint performance? There are several challenges to consider.

First, our external dose is a measure of running time above a speed threshold (measured), our internal dose is the neuromuscular stimulus (unmeasured) and our intermediate outcome of interest is sprint performance (measured), under the assumption (unverified) that sprint performance is a determinant of match performance (unmeasured). If we want to make inferences about neuromuscular stimulus on sprint performance, we can think of the running time measure as a "surrogate" for internal dose.[19] This must be justified with empirical or at least theoretical evidence.

If an observational study with no biases finds that our measure of running time above a certain threshold is associated with improved performance, there remain two possibilities. Either the exposure does provide a neuromuscular stimulus that improves performance, or the exposure causes a neuromuscular stimulus but its effect on performance occurs completely through an independent mechanism. In other words, the selected measure of exposure may appropriately reflect an identified mechanism of interest, however, this mechanism is irrelevant (or marginally relevant) to the outcome. In short, this measure of exposure may be a valid reflection of an incorrectly identified mechanism.

To summarise, the validation process within a causal context includes two-steps. The first step consists of providing evidence that the selected measure is a valid reflection of the putative causal mechanism linked to the outcome i.e., evidence about its construct validity where the construct (latent variable) is the hypothesised mechanism. The second step requires the examination of whether this putative mechanism is actually causally related to the outcome of interest. From a practical perspective, to complete this process, a measure of the mechanism is needed, which is the measure validated in the first step. Then, this measure is used in a purposely made study to examine the causal role (mediator) of the proposed mechanisms on the outcome of interest.

Finally, we want to highlight that validity is *not* an absolute concept and is context dependent.[59,60] As an example, heart rate (or VO_2) can only be used to quantify one aspect of training load, the stimulation of "one part" of the cardio-respiratory system. Accordingly, heart rate (or VO_2) is only valid for the cardio-respiratory component of training load related to its specific mechanistic effects, and cannot be expected to reflect other relevant mechanisms pertinent to athletic training (such as the stimulation of the

anaerobic, or neuromuscular system). Therefore, heart rate (or VO_2) is a valid measure of exposure for certain questions related to the cardio-respiratory component of training load, but would be an invalid measure of exposure for questions related to muscle damage or neuromuscular stimulus. In conclusion, understanding "validity for what purpose" and the context is essential for a correct interpretation.

7.1. Dose-response as evidence of validity

Since a valid surrogate for our question requires particular causal relationships, appropriate causal inference methods are needed[61-63] to examine whether the dose-response relationship is indeed causal.[62,64] Some investigators infer a measure is a valid surrogate (of the internal dose or outcome) if there is a dose-response relationship (biological gradient) because it is one of the Bradford Hill criteria of causality. However, associations between the dose of exposures and responses may be due to confounding and not causality,[14,65,66] or multiple independent effects of the exposure.

Therefore, at the forefront, the well-known differentiation between causal and non-causal (e.g., predictive) relations[62,67-69] becomes relevant when interpreting the dose-response as evidence of validity. If the relationship is not causal, we can still use the dose-response relationship to explore or to predict. For exploration, the relationship can be utilised to develop hypotheses. For prediction, the validity only concerns the ability to accurately (or acceptably) predict the likelihood of an outcome or future event (forecasting), and the prediction model cannot automatically be considered causal or reflective of causal factors.[62,67,68] In our previous example, our measure of running time might have had dose-response relationship with neuromuscular status and with performance even though neuromuscular status did not mediate the effect. If true, developing new interventions to change neuromuscular status would not improve performance. This mandates that it is not appropriate to attempt to manipulate a measure of exposure simply because it is one of the features of a prediction model.

Similarly, the absence of a dose-response relation may occur for reasons other than the non-validity of the measure of exposure or the non-causality of the mechanism (i.e., causation is still possible). For example, the identification of a dose-response relationship can be more difficult when a substantial effect occurs above a threshold (of exposure) or when there is a ceiling effect in the response. While a causally potent biological etiology is certainly desirable, and detailed causal knowledge provides arguably the strongest foundations for the validation process, such an approach has many challenges. For example, etiological factors may function and interact in complex ways, and their causal roles can sometimes only be understood in terms of the larger system in which they are embedded. That is various mechanisms may act simultaneously and interact to elicit a response of interest.[70] Developing an appropriate reference causal structure to inform the analysis and modelling is necessary but can be very challenging.

8. Conclusion

Sport and exercise science & medicine have adopted many terms from other scientific fields. By highlighting the parallels between them, we have illustrated how the concepts of external and internal training load are coherent and consistent with notions from some fields of epidemiology (and pharmacology). Training load is a term reflecting the general concept of exposure, while the separation of training load into external and internal training load closely mimics the separation of exposure into external and internal dose. These subdimensions allow for the differentiation of measures of training-related behaviours from the internal psycho-physiological stimulus (internal dose) induced by these behaviours (i.e., formalisation of the causal relationship between the training load sub-dimensions).

Validity is not an absolute concept; it is context-dependent and is derived from a variety of methods and sources which contribute to the research base for and against the theoretical framework supporting a construct. Additionally, the validation process depends on the appropriate selection of the measures of exposure and the response of interest, as per the context. Ideally, a measure should reflect the mechanisms that, at least theoretically, link the exposure to the targeted effects/responses. However, etiological pathways facilitating sporting performance or health-related outcomes are complex and may be dependent upon a variety of mechanisms acting concurrently and potentially interacting with each other. It is therefore unlikely that a single measure can reflect all the mechanisms mediating the response of interest.

9. Practical applications

The fundamental goals of training are to improve athletic performance, reduce the risk of injury or improve health. Any measure of exposure should provide information to support the planning and execution of a training program aimed to improve characteristics that are causally linked to performance (or to any outcome of interest). The take-home message of this article is that the measures of training load (to support and optimise the training process) should be chosen wisely based on a plausible relationship with a mechanism of interest and the evidence supporting it.

A recent phenomenon of concern is the development, introduction and adoption of several new metrics that lack conceptual support: i.e., there is no explicit theoretical framework that reasonably links these metrics to the mechanism that it is supposed to reflect, nor whether these mechanisms are reasonably related to the outcome of interest (e.g., performance enhancement or better health). In other words, these metrics are not theory-driven, and the “burden of proof” has been reversed; it appears that measures of exposure are presented, and it is then left to others to try and understand what it actually measures and whether it can be useful. It does not matter how sophisticated or “advanced” a metric appears. If it cannot be connected to a plausible mechanism (or relevant responses), it is likely of little use

to support and optimise the training process in practice. However, such metrics can still be used for exploration and to generate hypotheses from a research perspective.

Finally, we would like to emphasise that with this and previous articles[8,9,11,12,71] we are not claiming that the provided classification and framework *must* be used; we have simply presented several arguments and "informal" conceptual analyses to explain why we believe this conceptualisation of training load and its components can be used and is useful to scientifically investigate the physical training process.

Conflict of Interest: The authors declare no financial or non-financial conflicts of interest with the contents of the current article.

Financial support: This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Data and Supplementary Material Accessibility: no data were used.

CRediT author statement: FMI: Conceptualization, Writing- Original draft preparation; IS: Conceptualization, Writing- Reviewing and Editing; SM: Writing- Reviewing and Editing; AJC: Writing- Reviewing and Editing; AM: Writing- Reviewing and Editing; KS: Writing- Reviewing and Editing; ACJ: Writing- Reviewing and Editing; JK: Conceptualization, Writing- Reviewing and Editing.

References

1. Damsted C, Glad S, Nielsen RO, Sørensen H, Malisoux L. Is there evidence for an association between changes in training load and running-related injuries? A systematic review. *Int J Sports Phys Ther.* 2018;13(6):931-942.
2. Schoenfeld BJ, Ogborn D, Krieger JW. Dose-response relationship between weekly resistance training volume and increases in muscle mass: A systematic review and meta-analysis. *J Sports Sci.* 2017;35(11):1073-1082.
3. Fitzpatrick JF, Hicks KM, Hayes PR. Dose-Response Relationship Between Training Load and Changes in Aerobic Fitness in Professional Youth Soccer Players. *Int J Sports Physiol Perform.* 2018:1-6.
4. Wasfy MM, Baggish AL. Exercise Dose in Clinical Practice. *Circulation.* 2016;133(23):2297-2313.
5. Staunton CA, Abt G, Weaving D, Wundersitz DWT. Reply to: "The 'training load' construct: Why it is appropriate and scientific". *J Sci Med Sport.* 2022;25(5):449-450.
6. Staunton CA, Abt G, Weaving D, Wundersitz DWT. Misuse of the term 'load' in sport and exercise science. *J Sci Med Sport.* 2022;25(5):439-444.

7. Passfield L, Murias JM, Sacchetti M, Nicolo A. Validity of the Training-Load Concept. *Int J Sports Physiol Perform.* 2022:1-8.
8. Impellizzeri FM, Jeffries AC, Weisman A, et al. The 'training load' construct: Why it is appropriate and scientific. *J Sci Med Sport.* 2022;25(5):445-448.
9. Impellizzeri FM, Jeffries AC, Weisman A, et al. Letter to "The 'training load' construct: Why it is appropriate and scientific": Authors' reply. *J Sci Med Sport.* 2022;25(5):451-452.
10. Porta M. *A Dictionary of Epidemiology.* New York, NY, US: Oxford University Press; 2016.
11. Impellizzeri FM, Marcora SM, Coutts AJ. Internal and External Training Load: 15 Years On. *Int J Sports Physiol Perform.* 2019;14(2):270-273.
12. Jeffries AC, Marcora SM, Coutts AJ, Wallace L, McCall A, Impellizzeri FM. Development of a Revised Conceptual Framework of Physical Training for Use in Research and Practice. *Sports Med.* 2021.
13. Lee TA, Pickard AS. Exposure definition and measurement. In. *Developing a protocol for observational comparative effectiveness research: a user's guide*: Agency for Healthcare Research and Quality (US); 2013.
14. Rothman KJ, Greenland S, Lash TL. *Modern epidemiology. 3rd edition.* Philadelphia: Wolters Kluwer 2008.
15. Lash TL, VanderWeekle TJ, Haneuse S, Rothman KJ. *Modern epidemiology. 4th edition.* Philadelphia: Wolters Kluwer 2021.
16. Coutts AJ, Crowford S, Kempton T. Developing athlete monitoring systems: theoretical basis and practical applications. In: Kellmann M, Beckmann J, eds. *Sport, recovery, and performance: interdisciplinary insights.* London: Routledge; 2017.
17. Stevens SS. On the Theory of Scales of Measurement. *Science.* 1946;103(2684):677-680.
18. Ahrens W, Pigeot I. *Handbook of epidemiology.* Vol 451. New York, NY: Springer; 2014.
19. White E, Armstrong BK, Saracci R. *Principles of Exposure Measurement in Epidemiology: Collecting, Evaluating, and Improving Measures of Disease Risk Factors.* 2 ed. New York, NY, US: Oxford University Press; 2008.
20. Kalkhoven JT, Watsford ML, Coutts AJ, Edwards WB, Impellizzeri FM. Training Load and Injury: Causal Pathways and Future Directions. *Sports Med.* 2021;51(6):1137-1150.
21. Cordier S, Stewart PA. Exposure Assessment. In: Ahrens W, Pigeot I, eds. *Handbook of Epidemiology.* New York, NY: Springer New York; 2014:613-638.
22. Cordier A, Stewart PA. Exposure Assessment. In: Ahrens W, Pigeot I, eds. *Handbook of epidemiology.* New York, NY: Springer; 2014:437-462.
23. Food and Agriculture Organization of the United Nations, World Health Organization. *Chapter 5: dose-response assessment. Principles and methods for the risk assessment of chemicals and food* Stuttgart, Germany: World Heath Organization; 2009.
24. Boslaugh SML-ATG. *Encyclopedia of epidemiology.* Farmington Hills, Mich.: Thomson Gale; 2008.

25. Slob W, Zeilmaker MJ, Hoogenveen RT. The Relationship Between Internal and External Dose: Some General Results Based on a Generic Compartmental Model. *Toxicol Sci.* 2020;177(1):60-70.
26. Checkoway H, Pearce N, Kriebel D. *Research methods in occupational epidemiology.* Vol 34. New York, NY, US: Oxford University Press; 2004.
27. Roedde S, MacDougall JD, Sutton JR, Green HJ. Supercompensation of muscle glycogen in trained and untrained subjects. *Can J Appl Sport Sci.* 1986;11(1):42-46.
28. Morgan WP, Brown DR, Raglin JS, O'Connor PJ, Ellickson KA. Psychological monitoring of overtraining and staleness. *Br J Sports Med.* 1987;21(3):107-114.
29. Daniels JT, Yarbrough RA, Foster C. Changes in VO2 max and running performance with training. *Eur J Appl Physiol Occup Physiol.* 1978;39(4):249-254.
30. Kalkhoven JT, Watsford ML, Impellizzeri FM. A conceptual model and detailed framework for stress-related, strain-related, and overuse athletic injury. *J Sci Med Sport.* 2020;23(8):726-734.
31. Shrier I, Platt RW. Reducing bias through directed acyclic graphs. *BMC Med Res Methodol.* 2008;8(1):70.
32. Kriebel D, Checkoway H, Pearce N. Exposure and dose modelling in occupational epidemiology. *Occup Environ Med.* 2007;64(7):492-498.
33. Checkoway H, Pearce N, Crawford-Brown DJ. *Research methods in occupational epidemiology.* New York, NY, US: Oxford University Press; 1989.
34. Calvert TW, Banister EW, Savage MV, Bach T. A systems model of the effects of training on physical performance. *IEEE Trans Syst Man Cybern Syst.* 1976(2):94-102.
35. Morton RH, Fitz-Clarke JR, Banister EW. Modeling human performance in running. *J Appl Physiol.* 1990;69(3):1171-1177.
36. Banister EW, Calvert TW, Savage MV, Bach TM. A systems model of training for athletic performance. *Aust J Sports Med.* 1975;7:57-61.
37. Foster C, Hector LL, Welsh R, Schrager M, Green MA, Snyder AC. Effects of specific versus cross-training on running performance. *Eur J Appl Physiol Occup Physiol.* 1995;70(4):367-372.
38. American College of Sports Medicine, Riebe D, Ehrman JK, Liguori G, Magal M. *ACSM's guidelines for exercise testing and prescription.* Tenth edition ed. Philadelphia: Wolters Kluwer; 2018.
39. Kraemer WJ, Ratamess NA. Fundamentals of resistance training: progression and exercise prescription. *Med Sci Sports Exerc.* 2004;36(4):674-688.
40. American College of Sports Medicine position stand. Progression models in resistance training for healthy adults. *Med Sci Sports Exerc.* 2009;41(3):687-708.
41. McBride JM, McCaulley GO, Cormie P, Nuzzo JL, Cavill MJ, Triplett NT. Comparison of methods to quantify volume during resistance exercise. *J Strength Cond Res.* 2009;23(1):106-110.

42. Renfree A, Casado A, McLaren S. Re-thinking athlete training loads: would you rather have one big rock or lots of little rocks dropped on your foot? *Res Sports Med.* 2021:1-4.
43. Shrier I, Wang C, Stokes T, Trejovargas J, Stovitz SD, Steele RJ. Causal effects, workload and injury risk: The importance of specifying the research question. *J Sci Med Sport.* 2022;25(7):574-578.
44. de Vocht F, Burstyn I, Sanguanchaiyakrit N. Rethinking cumulative exposure in epidemiology, again. *J Expo Sci Environ Epidemiol.* 2015;25(5):467-473.
45. Smith TJ, Kriebel D. Current Topics in Occupational Epidemiology. In: *Biologically based exposure assessment for epidemiology.* Croydon, UK: Oxford University Press; 2013.
46. Burstyn I, Barone-Adesi F, de Vocht F, Gustafson P. What to Do When Accumulated Exposure Affects Health but Only Its Duration Was Measured? A Case of Linear Regression. *Int J Environ Res Public Health.* 2019;16(11).
47. Aschengrau ASGR. *Essentials of epidemiology in public health.* Burlington, MA: Jones and Bartlett Publishers, Inc; 2018.
48. Ritz C, Baty F, Streibig JC, Gerhard D. Dose-Response Analysis Using R. *PloS one.* 2016;10(12):e0146021.
49. *Overtraining in sport.* Champaign, IL, US: Human Kinetics; 1998.
50. Danieli C, Sheppard T, Costello R, Dixon WG, Abrahamowicz M. Modeling of cumulative effects of time-varying drug exposures on within-subject changes in a continuous outcome. *Stat Methods Med Res.* 2020;29(9):2554-2568.
51. Wagner M, Grodstein F, Leffondre K, Samieri C, Proust-Lima C. Time-varying associations between an exposure history and a subsequent health outcome: a landmark approach to identify critical windows. *BMC Med Res Methodol.* 2021;21(1):266.
52. Vanderweele TJ. Surrogate measures and consistent surrogates. *Biometrics.* 2013;69(3):561-569.
53. Pearl J. [Discussion on "Surrogate Measures and Consistent Surrogates"]. *Biometrics.* 2013;69(3):575-577.
54. Joffe MM, Greene T. Related causal frameworks for surrogate outcomes. *Biometrics.* 2009;65(2):530-538.
55. Prentice RL. Surrogate and Mediating Endpoints: Current Status and Future Directions. *J Natl Cancer Inst.* 2009;101(4):216-217.
56. Lassere MN. The Biomarker-Surrogacy Evaluation Schema: a review of the biomarker-surrogate literature and a proposal for a criterion-based, quantitative, multidimensional hierarchical levels of evidence schema for evaluating the status of biomarkers as surrogate endpoints. *Stat Methods Med Res.* 2008;17(3):303-340.
57. DeMets DL, Psaty BM, Fleming TR. When Can Intermediate Outcomes Be Used as Surrogate Outcomes? *Jama.* 2020;323(12):1184-1185.
58. Yin Y, Liu L, Geng Z, Luo P. Novel criteria to exclude the surrogate paradox and their optimalities. *Scand J Stat.* 2020;47(1):84-103.

59. Sechrest L. Validity of measures is no simple matter. *Health Serv Res.* 2005;40(5p2):1584-1604.
60. American Educational Research Association, American Psychological Association, Education NCoMi. *The standards for educational and psychological testing.* American Educational Research Association; 2014.
61. Pearl J. *Causality.* 2 ed. Cambridge: Cambridge University Press; 2009.
62. Hernán MA, Hsu J, Healy B. A Second Chance to Get Causal Inference Right: A Classification of Data Science Tasks. *CHANCE.* 2019;32(1):42-49.
63. Díaz I, van der Laan MJ. Targeted data adaptive estimation of the causal dose–response curve. *J Causal Inference.* 2013;1(2):171-192.
64. Pearl J. Causal inference. *Causality: objectives and assessment.* 2010:39-58.
65. Rothman KJ, Greenland S. Causation and Causal Inference in Epidemiology. *Am J Public Health.* 2005;95(S1):S144-S150.
66. Shimonovich M, Pearce A, Thomson H, Keyes K, Katikireddi SV. Assessing causality in epidemiology: revisiting Bradford Hill to incorporate developments in causal thinking. *Eur J Epidemiol.* 2021;36(9):873-887.
67. Ramspek CL, Steyerberg EW, Riley RD, et al. Prediction or causality? A scoping review of their conflation within current observational research. *Eur J Epidemiol.* 2021;36(9):889-898.
68. Shmueli G. To explain or to predict? *Stat Sci.* 2010;25(3):289-310.
69. Stovitz SD, Verhagen E, Shrier I. Distinguishing between causal and non-causal associations: implications for sports medicine clinicians. *Br J Sports Med.* 2019;53(7):398-399.
70. Zachar P, Jablensky A. Introduction: The concept of validation in psychiatry and psychology. In: Zachar P, Stoyanov DS, Aragona M, Jablensky A, eds. *Alternative perspectives on psychiatric validation: DSM, IDC, RDoC, and Beyond.* New York, NY, US: Oxford University Press; 2014.
71. Impellizzeri FM, Rampinini E, Marcora SM. Physiological assessment of aerobic training in soccer. *J Sports Sci.* 2005;23(6):583-592.