Gold standards for scientists who are conducting animal-based exercise studies

Frank W. Booth,1–3 Matthew J. Laye,2 and Espen E. Spangenburg4

1Departments of Biomedical Sciences, 2Medical Pharmacology and Physiology, and 3Nutrition and Exercise Physiology, University of Missouri, Columbia, Missouri; and 4Department of Kinesiology, University of Maryland, College Park, Maryland

IN RECENT YEARS, a disturbing trend has emerged in which scientific journals are accepting manuscripts in exercise biochemistry/physiology that have employed inappropriate methodologies to study exercise or exercise training (for example, Refs. 2, 10, 14, 17, 18, 27, 28). A lack of basic knowledge in exercise biochemistry is often evident and usually coupled with a failure to cite classic literature in the field. The outcome is the publication of questionable science that is subsequently subject to erroneous interpretation.

The purpose of this Viewpoint is to encourage a dialogue on the criteria needed to establish standards for studies that employ exercise and/or exercise training as an experimental intervention to permit clear, consistent, and appropriate evaluation of data by both investigators and readers. This Viewpoint will attempt to remind us, as scientists, that to maintain high quality exercise-based publications we must not ignore, but continue to build on, the experimental standards previously established by the leaders of our field. Thus this Viewpoint is divided into sections where important misconceptions have found their way into the current literature.

Important considerations in selecting methods for fiber typing and for testing of endurance times to exhaustion.

Fiber Types

Skeletal muscle is a heterogenous tissue composed of multiple fiber types that can be classified based 1) on contractile speed if the aim is contractile phenotype only or 2) on metabolic function if the goal is metabolic phenotype only. If the scientific aim is fiber type switching classification, then both contractile and metabolic markers must be determined separately, according to principles published by Gorza (9) and Edgerton and colleagues (21). Gorza (9) stated, “Nevertheless, the combined use of immuno- and enzyme histochemistry prevents incorrect fiber typing due to the interspecies variability of myosin ATPase activity among the correspondent fiber types, and completely modifies the presently used classification of mouse type 2 fibers.” Edgerton (21) wrote, “These findings demonstrate that a single muscle characteristic is not sufficient for discriminating between muscle fibres with different myosin heavy chain content.”

Unfortunately, misconceptions exist concerning use of appropriate methods for fiber type classification; they are as follows. First, misclassification of contractile phenotype due to an inappropriate myosin ATPase methodology, and second, the inappropriate inference of metabolic or contractile fiber type classification. Concerning the first issue of fiber type misclassification, Gorza (9) reported that preincubation at a single pH will not distinguish all fiber types. For example, in mice and rats, a single preincubation pH at 4.3, will not distinguish types IIa, IIx, and IIb fibers, while a single preincubation pH at 4.6 cannot distinguish between fiber types IIx and IIb in rats and I and IIx in mice. At a preincubation of 10.4, fiber types IIa and IIx cannot be distinguished in rats and mice. Thus at least two serial sections performed at a different preincubation pH must be done to obtain the contractile composition of muscle based on the myosin ATPase. Gorza (9) also reports that the degree of actomyosin ATPase inactivation may vary for the same fiber type in different species. At odds with Gorza, numerous papers appear to use a single preincubation pH on a single muscle section, citing human fiber type metachromatic staining methodology in their reports of “contractile” fiber type switching in mice (3, 18, 28).

Concerning the second issue of inappropriate fiber type classification, determination of metabolic fiber type from myosin ATPase methodology without metabolic classification of fibers in serial muscle sections can lead to inappropriate interpretations. For example, after spinal cord injury, a large increase in IIx fiber percentage occurs but concurrently is associated with an increase in the oxidative metabolic fiber type marker succinate dehydrogenase. Here assuming that a higher percentage of fast fibers equates with a less oxidative muscle is incorrect. While immunohistochemical analysis with myosin heavy chain specific antibodies allows for contractile fiber typing, it does not provide 100% certainty for metabolic phenotype because 1) of the continuous metabolic spectrum within fiber types (21 and 2) mitochondrial biogenesis and fiber type transformation can occur independent from each other (22).

Thus gold standards requiring dialogue are 1) whether Gorza’s and Edgerton’s fiber type standards are to be the gold standard and 2) whether definitive fiber type switching requires both a measure of contractile and metabolic phenotype change or if the inference of metabolic fiber type switching from contractile phenotype and vice vs. is acceptable.

Measuring Animal Exhaustion in Treadmill Running

Numerous investigations are using exhausting exercise on the treadmill to quantify endurance capacity of a rodent without proving exhaustion with biochemical measures. Exhaustion is defined as “the inability to continue exercise” (5) and must
include empirical data proving the animal is fatigued to claim “exhaustion.” Previously established criteria used to prove fatigue in rodent studies include the inability of the animal to right itself (1, 24), blood glucose <76 mg/100 ml blood (7), or low skeletal muscle and liver glycogen concentrations (11).

A commonly published criterion for rodent exhaustion has now become the number of electric shock apparatus contacts (6, 14, 18, 27). To our knowledge, no studies have proven that the frequency that a rodent contacts an electric shock is associated with any of the other aforementioned measurements of exhaustion. A confounding variable to employing contact number to electric shocks alone is rodents exhibit different running styles with some styles causing premature stopping before biochemical exhaustion. For example, in Copp et al. (4), some rats that turn and fight the treadmill belt for an extended time before demonstrating a natural running gait at the beginning of a test will likely fatigue earlier because of excess energy expended, and they recommend excluding these animals from exhaustion tests.

Sufficient information is often omitted so that efficacy of exhaustion tests and exercise training cannot be determined. Some publications fail to provide the following information in the methods to allow confidence in the claim of exhaustion. 1) Thorough and carefully conducted acclimatization protocol to running on the treadmill to eliminate rodents whose running style would bias running distance (4); 2) with exhaustion, animals visually exhibit changes in body gait that can only be detected by an experienced monitor (4); and 3) the intensity (grade and speed) that animals ran, since exhaustion is most repeatable in rats running at 70–90% of \( \text{VO}_2 \text{peak} \) (4).

Thus a gold standard requiring dialogue is whether exhaustion from forced treadmill running can be determined with 100% fidelity without proving biochemical or central fatigue.

**Proper Application of Exercise Training Concepts**

Documentation of endurance training by an increase in classical markers of metabolic adaptation, such as skeletal muscle citrate synthase, mitochondrial DNA or protein concentrations/activities, and/or peak \( \text{VO}_2 \) uptake, must be demonstrated to indicate a sufficient training stimulus (reviewed in Ref. 12). Furthermore, investigators need to define their exercise paradigm, by providing the reader with the duration, intensity, and frequency of the exercise protocol. When skeletal muscle strength training models are applied, typical markers should include changes in muscle mass and total protein per whole muscle (reviewed in Ref. 25). Another error is a failure to either know the classic literature or to be unable to place findings in proper context. For example, “exercise” is illustrated to increase type I fibers (Fig. 3 of Ref. 15). Holloszy (8) and colleagues commented, “The only conversion of fiber types in response to endurance exercise that is well documented is conversion of type IIb to IIx to type IIa fibers . . . It is possible that the extreme training programs...that involve many hours of intense exercise daily for years may cause an increase in slow (type I) muscle fibers, but this remains to be documented.” These concepts to demonstrate a traditional training adaptation are particularly important when comparing a drug to exercise (18).

Thus a gold standard requiring dialogue is whether proof of an established training adaptation is required to prove training efficacy in healthy populations for endurance, strength, neural, and/or flexibility.

**Use of the Word “Exercise”**

The type of exercise [e.g., endurance, strength, neural, flexibility (19)] must precede the term “exercise” to permit readers to understand the results, since it has been known for decades that each type of exercise has its own unique set of acute and chronic adaptations/phenotype (13, 20). Such terminology is now critical. Recently the term “exercise mimetic” appeared (18), implying all adaptations from endurance, strength, coordination, and flexibility training from therapeutic intervention. For example, recombinant erythropoietin (in 1988)(16) and AICAR (in 2008)(18) may increase endurance performance and yet these drugs do not totally “mimic” all adaptations and both contain significant pathological side effects (23, 29), unlike endurance exercise training. No scientific precedence existed prior to 2008 to associate “exercise” with words like “mimetic” or “pill”, such usage, if used, would require careful consideration of all exercise adaptations.

The gold standard requiring dialogue is whether it is scientific and ethical to claim a chemical induces all the adaptations of exercise in healthy humans when measuring only a small fraction of exercise adaptations and when this chemical produces detrimental changes not associated with exercise.

**Overall Conclusions**

To properly report exercise responses and adaptations, it is critical that exercise scientists reexamine the methodological gold standards that were established through the elegant work of our predecessors (26). The proper use of these gold standards ensures reliable data, appropriate interpretation, and comparison with other exercise literature.

It is the goal of this Viewpoint to begin an in-depth, lengthy debate on the issues raised. The deliberations should lead to a consensus position statement from some society such as the Environmental and Exercise Physiology section of the American Physiological Society or the American College of Sports Medicine. A position statement would provide appropriate animal exercise research standards so that other journals and funding agencies perform proper scientific reviews of exercise animal research using standards developed from those with valid training in exercise biochemistry and who have proven such with long-track records in their peer-reviewed journals.

**REFERENCES**


