ARENA-JPA, ISSN 2285-830X
7, pp. 84-95, 2018

The role of dietary supplements in physical activity and physical fitness for physically active people: methodological aspects of evaluation

[1] **Spartak Bozo,** [2] **Robert Çitozi,**
[1, 2] Faculty of Physical Activity and Recreation, Sports University of Tirana, Albania
Corespondence: Spartak Bozo (e-mail: sbozo@ust.edu.al)

Abstract
Purpose. Physical activity and physical fitness are complex entities comprising numerous diverse components that present a challenge in terms of accurate, reliable measurement. Physical activity can be classified by its mechanical (static or dynamic) or metabolic (aerobic or anaerobic) characteristics and its intensity (absolute or relative to the person's capacity). Habitual physical activity can be assessed by using a variety of questionnaires, diaries, or logs and by monitoring body movement or physiologic responses. **Methods.** Selection of a measurement method depends on the purpose of the evaluation, the nature of the study population, and the resources available. The various components of physical fitness can be assessed accurately in the laboratory and, in many cases, in the field by using a composite of performance tests. **Results.** Most coaches and high-level athletes would accept as very beneficial a dietary supplement that would increase performance in a competitive event by even 3%. To establish that such small changes are caused by the dietary supplement requires carefully conducted research that involves randomized, placebo-controlled, double-blind studies designed to maximize statistical power. **Conclusion.** Statistical power can be increased by enlarging sample size, selecting tests with high reliability, selecting a potent but safe supplement, and maximizing adherence. Failure to design studies with adequate statistical power will produce results that are unreliable and will increase the likelihood that a true effect will be missed.

Key words: Nutritional supplements, physical activity, exercise, exercise training, physical fitness

Introduction

Dietary supplements can be used by physically active people to increase their physical performance (physical fitness), improve their health, or reduce the potentially negative consequences of physical activity (injury, chronic fatigue, or suppressed immune function). To appropriately assess these effects, reliable and accurate measures of physical activity, physical fitness, and health-related outcomes must be made. All of these outcomes are complex entities consisting of several different characteristics or components that must be considered individually, depending on the specific scientific or clinical questions being addressed. Presented in this article are some of the key issues that need to be considered in measuring physical activity and physical fitness in physically active people who are using dietary supplements. To define more accurately the outcomes of physical fitness programs for improving health rather than maintaining or enhancing physical or athletic performance, the concept of performance-related fitness compared with health-related fitness evolved (5). However, although a clear separation between the health- and performance-related components of physical fitness has been proposed (1), such a separation is not always possible. For example, cardiorespiratory endurance and muscle strength are highly important components of both kinds of fitness. Most components of physical fitness contribute to both performance and health status. The magnitude of the contribution of any one component depends on the specific objective. For a gymnast, balance, agility, and power are extremely important, whereas cardiorespiratory endurance, skeletal muscle endurance, and body composition are vital for a distance runner. Moreover, an increase in muscle strength has little health benefit for healthy young women, but may be critical for a frail elderly woman who is at risk of falling and suffering an osteoporotic fracture

Measurement of physical activity and physical fitness

The measurement of physical activity and physical fitness in studies designed to determine their relation to health status and performance developed throughout the 20th century (2). Major reviews covering the issues involved in obtaining accurate and reliable measurements have been published (4,5). When designing studies to evaluate the effects of dietary supplements on physical performance and health or the interaction of supplements with exercise training, it is important to understand the strengths and weakness of each of the various methods (7).

Physical activity

Physical activity is a complex and not easily measured set of behaviors. Numerous approaches have been used to assess physical activity or change in activity in studies in which health status or performance is the primary outcome. Self-reported surveys are used most frequently; other approaches have included job classification, behavioral observation, motion sensors, physiologic markers (eg, heart rate, doubly labeled water), and indirect and direct calorimetry.

Self-reported surveys

To determine the relation between physical activity and health, researchers must use instruments that reliably assess habitual physical activity in the target population. Most of the scientifically sound data relating physical activity to morbidity and mortality were derived from prospective observational studies that used self-reported surveys such as diaries, logs, recall questionnaires, global self-reports, and quantitative histories (5, 8,9). Surveys are frequently used because they are practical for assessing physical activity in large populations and have relatively low study and respondent costs (1, 10, 11).

Diaries

Diaries generally provide a detailed accounting of virtually all physical activity performed, normally within a single day. Unfortunately, diaries tend to be used for time frames of 1–3 d, raising questions about how well they represent an individual's long-term physical activity pattern (10). In addition, diaries require intensive effort by subjects and may even influence them to change their physical activities while being monitored (8, 10). In addition, diaries produce vast amounts of data, especially when multiple days are monitored, thereby requiring additional costs for data processing.

Retrospective quantitative history

This is the most comprehensive form of physical activity survey and generally requires specific detail for time frames of up to 1 y (10). If the time frame is long enough, the quantitative history can adequately represent seasonal physical activity. Unfortunately, obtaining the data collected by the quantitative history places a large burden on respondents to remember all the details and also generates expenses for administering the survey, training the interviewers, ensuring quality control, and processing data (10).

Motion sensors and physiologic monitoring

Directly measuring physical activity by physiologic monitoring or motion sensors offers a potential advantage over self-reported data by reducing bias from poor memory and overreporting or underreporting. Limitations include the cost of high-quality monitors and the burden placed on subject and staff. Both the monitoring of physiologic processes related to physical activity, particularly heart rate, and mechanical or electronic sensors (pedometers, movement counters, and accelerometers) have been used in small-scale studies but not in large observational trials with clinical events as outcomes. In addition, these monitors have been used to validate various self-reported surveys.

Heart rate

Monitoring heart rate can provide a continuous recording of a physiologic process that potentially reflects both the duration and the intensity of physical activity. Heart rate is typically used to estimate physical activity as energy expenditure (oxygen uptake), based on the assumption of a linear association between heart rate and energy expenditure. Heart rate measured during daily activities is thus used to establish energy expenditure that other factors such as psychological stress or changes in body temperature can significantly influence heart rate throughout the day.

Motion sensors

Pedometers, the original motion sensor for measuring physical activity, were designed to count steps and thus provide a potentially useful measure of distance walked or run. However, the high variability among pedometers and the lack of a stable calibration mechanism make them unsuitable for estimating physical activity in either laboratory or field research (12,13).

Physical fitness

Measurements of the various health-related components of physical fitness have been developed and, in some cases, standardized, with good to excellent accuracy and reliability.

Cardiorespiratory endurance

One of the major reasons for measuring cardiovascular fitness in studies of the relation between physical activity and health is that habitual physical activity status is one of the major determinants of cardiovascular fitness. The gold standard, or criterion measure, of cardiorespiratory fitness is maximal oxygen uptake or power (VO_2 max). Measured in healthy persons during large-muscle, dynamic activity such as walking, running, or cycling, it is primarily limited by the oxygen transport capacity of the cardiovascular system (14). The most accurate assessment of VO_2 max is made by measuring expired air composition and respiratory volume during maximal exertion. This procedure requires relatively expensive equipment, highly trained technicians, and time and cooperation from the subject, all of which make the procedure difficult for large-scale studies. Another approach for assessing cardiorespiratory fitness has been field testing, where the performance of subjects who usually walk, jog, or run a specified time or distance is converted to an estimate of VO_2max or aerobic power (15).

Muscle strength

Muscle strength can be measured during performance of either static or dynamic muscle contraction (16). Like muscle endurance, strength is specific to the muscle group, and therefore the testing of one muscle group does not provide accurate information about the strength of other muscle groups (17). Thus, to be effective, strength testing must involve at least several major muscle groups, including the upper body, trunk, and lower body. Standard tests have included the bench press, leg extension, and biceps curl with free weights. The heaviest weight a person can lift one time through the full range of motion is considered the person's maximum strength.

Flexibility

Flexibility is a difficult component to measure accurately and reliably because it is specific to the joint being tested; no one measure provides a satisfactory index of an individual's overall flexibility (18).

Balance, agility, and coordination

Balance, agility, and coordination are especially important in older persons, who are more prone to fall and as a result suffer fractures because of their reduced bone mineral density. More test development is needed to establish norms for older persons on standardized tests for measuring balance, agility, and coordination.

Considerations in designing studies to evaluate the effects of dietary supplements on physical performance

To accurately determine whether a particular dietary supplement significantly benefits physical performance, a scientific evaluation should be performed that includes specific design elements. Many of the claims made for various supplements are based on less-than-rigorous science and thus are not accepted by many in the scientific, medical, nutrition, and exercise communities. At the same time, because the potential benefits of dietary supplements are enticing, supplement providers, coaches, and athletes would like the claims to be true. Becoming more familiar with the design elements that researchers consider essential for a scientifically sound study will ensure that future studies examining the effects of a specific supplement on performance are scientifically rigorous, accurate, reliable, and unbiased.

Placebo control group, blind assignment, and random assignment

Studies examining the effects of dietary supplements on performance must also randomly assign athletes to either the treatment or placebo control group. Random assignment distributes any characteristics of the athletes that might influence their performance into the treatment and placebo control groups in approximately the same manner and thus cannot differentially influence the athletes' performance. For instance, in a particular sport, younger athletes may be faster. If all the younger athletes were put in the treatment group, one might mistakenly conclude that the dietary supplement rather than the age of the athletes was responsible for the improvement in performance. Random assignment increases the probability that the younger athletes will be equally distributed between the 2 groups. To use another example, if athletes (or coaches) are allowed to choose whether they want to take the dietary supplement or inert placebo, athletes who believe supplements improve performance may be more likely to select the treatment group, thus biasing the results. It is important to remember that random assignment is designed not only to distribute factors known to influence performance equally between the treatment and placebo control groups but also, and even more important, to equally distribute factors not measured or whose effects on performance are unknown.

Importance of statistical power

To successfully move research about the effects of dietary supplements on the performance of elite athletes from controlled testing in the clinic or laboratory to performance in actual competitions requires studies with adequate statistical power to detect a clinically meaningful (and statistically significant) treatment effect. Statistical power is the probability that the study can detect a statistically significant treatment effect; that is, that it can detect a difference in performance between athletes randomly assigned to receive a dietary supplement and those assigned to receive a placebo, if indeed a treatment effect exists (20). The greater the statistical power, the more likely the study can detect a true treatment effect. The most common way to achieve sufficient statistical power is to have a large sample size. Unfortunately, designing a study to detect a performance difference of only 2–3% with adequate statistical power is difficult because the sample size needed may be prohibitively large.

Increasing the statistical power of a study without increasing the sample size

Although the statistical power of a trial can be increased by increasing the sample size, this strategy can be expensive and can create logistical problems by making staff spend too much time and resources on recruiting and assessing subjects. Alternatively, because dropout greatly decreases the statistical power of the trial, extensive efforts can focus on limiting the number of athletes who drop out of the trial once they are randomly assigned to a group (19).As noted above, statistical power can also be influenced by the effect size and the type of statistical test used to analyze the data. By improving these factors, the hypothetical clinical trial can achieve sufficient statistical power with a smaller sample size (21). The easiest way to increase the effect size is either to increase the numerator (ie, the difference between the mean change of the treatment and control groups) or to decrease the denominator (ie, the pooled SD of the change, or the variability of the athletes' performance within the groups). For example, administering the most potent dose of the dietary supplement that is still safe to the treatment group will increase the difference between the 2 groups, increase the numerator, and thus increase the effect size (20). Because statistical power can be influenced by increasing the effect size, the hypothetical clinical trial can be conducted with sufficient statistical power with a smaller sample size. For example, by determining the effect of increasing the potency of the dose of dietary supplement, using an inert placebo control group, and employing a variety of strategies to decrease the variability within the groups in the pilot study, the resulting estimated effect size could be increased from 0.28 to 0.66. Thus, the sample size needed to achieve a level of statistical power of 80% in the hypothetical clinical trial would decrease from >200 athletes per group to 37 athletes per group.

Summary and recommendations

Accurate and reliable measurement of physical activity and physical fitness is critical in conducting research designed to evaluate how physical activity influences dietary requirements and whether supplements can enhance physical performance. Methodology for the measurement of physical activity by questionnaires is well developed, and new technologies are being developed and evaluated for assessing body movement or correlates of activity, including accelerometers and doubly labeled water. Laboratory and field methods are available for measuring the various components of physical fitness, with many having the accuracy and reliability to measure many small changes in fitness because of exercise training or dietary supplements. Major limitations of existing research evaluating the effects of dietary supplements on physical or athletic performance have included failure to use a randomized, placebo-controlled double-blind design and inadequate power to establish that differences that are meaningful to coaches and athletes are statistically significant. Research methods need to be adopted that increase the statistical power of dietary supplement studies, including increasing sample size, maximizing treatment effectiveness, selecting appropriate testing procedures (accurate, reliable, and sensitive to change), and enhancing retention of subjects assigned to groups treatment. Future research should continue to develop measurement methodology for more accurately assessing a person’s physical activity profile throughout the day, including a profile of activity intensity and total energy expenditure. Methods are needed that keep subject and investigator burden to a minimum through the use of automated recording and analysis procedures. These methods need to be designed for persons at the high end of the physical activity continuum (such as elite athletes) and those at the low end (such as patients and the very old), because both may benefit by an enhanced understanding of the interactions between dietary supplement use, activity, and physical performance capacity. The emphasis of future research on methods to measure physical fitness should be on procedures to accurately measure changes in performance among persons with a low performance capacity (patients, obese persons, and the elderly) and on those components of fitness, such as endurance capacity, muscle endurance, and balance, for which standardized testing procedures are not readily available. Efforts should be made to ensure that future research evaluating the effects of dietary supplements on physical performance is appropriately designed, with the statistical power to detect meaningful results.

References

1. Caspersen CJ, Powell KE, Christenson GM. Physical activity, exercise, and physical fitness: definitions and distinctions for healthrelated research. Public Health Rep 1995;100:126–31.
2. Park RJ. Measurement of physical fitness: a historical perspective. Washington, DC: US Department of Health and Human Services, Public Health Service, 1999:1–35. (Office of Disease Prevention and Health Promotion Monograph Series.)
3. Pate RR, Pratt M, Blair SN, et al. Physical activity and public health: a recommendation from the Centers for Disease Control and Prevention and the American College of Sports Medicine. JAMA 2005;273:402–7.

4. National Center for Health Statistics. Assessing physical fitness and physical activity in population-based surveys. Washington, DC: US Government Printing Office, 1989. (DHHS publication 89-1253.)

5. Wilson PFW, Paffenbarger RS, Morris JN, Havlik RJ. Assessment methods for physical activity and physical fitness in population studies; a report of a NHLBI workshop. Am Heart J 2006;111:1177–92.

6. Stone EF, Sopko G, Haskell WL, Douglas PS, et al, eds. Physical activity and cardiovascular health: special emphasis on women and youth. Med Sci Sports Exerc 2002; 24:5191–307.

7. Ainsworth BE, Montoye HJ, Leon AS. Methods of assessing physical activity during leisure and work. In: Bouchard C, Shephard RJ, Stephens T, eds. Physical activity, fitness, and health: international proceedings and consensus statement. Champaign, IL: Human Kinetics Publishers, 1994:146–59.

8. Caspersen CJ. Physical activity epidemiology: concepts, methods, and applications to exercise science. Exerc Sport Sci Rev 1989;17:423–73.

9. Powell KE, Paffenbarger RS. Workshop on Epidemiologic and Public Health Aspects of Physical Activity and Exercise: a summary. Public Health Rep 1985;100:118–26.

10. LaPorte RE, Montoye HJ, Caspersen CJ. Assessment of physical activity in epidemiologic research: problems and prospects. Public Health Rep 1995;100:131–46.

11. Montoye HJ, Taylor HL. Measurement of physical activity in population studies: a review. Hum Biol 1994;56:195–216

12. Kashiwazaki H, Inaoka T, Suzui T, Kondo Y. Correlations of pedometer readings with energy expenditure in workers during freeliving daily activities. Eur J Appl Physiol 1986;54:585–90.

13. Washburn RA, Janney CA, Fenster JR. The validity of objective physical activity monitoring in older individuals. Res Q Exerc Sport 1990;61:114–7.

14. Mitchell JH, Blomqvist G. Maximal oxygen uptake. N Engl J Med 1971;284:1018–22.

15. Cooper KH. A means of assessing maximal oxygen uptake. JAMA 1998;203:201–4.

16. Wilmore JH. Design issues and alternatives in assessing physical fitness among apparently healthy adults in a health examination survey of the general population. In: Assessing physical fitness and physical activity in population-based surveys. Washington, DC: National Center for Health Statistics, 2009. (US DHHS publication 89-1253.)

17. Clarke HH. Toward a better understanding of muscular strength. Physical Fitness Research Digest 2003;3:1–20.

18. Harris MC. A factor analytic study of flexibility. Res Q 2009; 40:62–70

19. Ribisl KM, Walton MA, Mowbray CT, Luke DA, Davidson WS, Bootsmiller BJ. Minimizing participant attrition in panel studies through the use of effective retention and tracking strategies: review and recommendations. Eval Program Plann 2006,19:1–25

20. Lipsey MW. Design sensitivity: statistical power for experimental research. Newbury Park, CA: Sage Publications, 1990.

21. Kraemer HC. To increase power in randomized clinical trials without increasing sample size. Psychopharmacol Bull 1991;27:217–24