What Are Good Muscle Endpoints for Translational Studies?

Jörn Rittweger(✉)1,2

1Institute of Aerospace Medicine, German Aerospace Center, Cologne, Germany
2Department of Pediatrics and Adolescent Medicine, University of Cologne, Cologne, Germany

© The Author

Abstract
Muscles matter to our health because of their size, their involvement in energy metabolism and their relevance for locomotion. Adequate selection of good endpoints is crucial for successfully designing translational studies. At least eight different muscle functions matter to health, namely the mechanical functions of exerting force, velocity, power, elastic storage and braking power, the two metabolic functions of substrate uptake (e.g. carbohydrates, lipids and amino acids) and substrate provision (e.g. lactate and amino acids) and secretory functions. However, specific endpoint tests have been validated for muscle force and power only. Walking speed and grip strength demonstrate good predictive value for hard clinical endpoints, such as disability, loss of autonomy and death. Vertical jump power also has good ecological validity and construct validity, and it depicts excellent test-retest reliability, which is an important advantage with regard to the study of power. Assessment of muscle mass, e.g. by magnetic resonance imaging, dual energy X-ray absorptiometry or bioelectrical impedance, should be considered as an important secondary endpoint to enhance construct validity. Further secondary endpoints should be included wherever they are likely to enhance the plausibility of the study outcome and assessment of test-retest reliability at baseline is always recommended. Well-established methods exist for three relevant muscular endpoints, namely power, strength and muscle mass, and these endpoints lend themselves to utilization in clinical studies. However, such validated methods lack a number of additional muscle functions that are scientifically only emerging. This applies foremost to the metabolic function of muscles but also to its role in storage and dissipation of mechanical energy.

Keywords: sarcopenia; frailty; musculature; immobilization; health

Introduction

Muscles matter to our health because a) the skeletal musculature constitutes the body’s biggest organ, b) because skeletal muscles are heavily involved in the organism’s streaming of energy and cytokine signaling and c) because a healthy musculature is a prerequisite for locomotion and thus for an individual’s autonomy and well-being. Moreover, it is widely recognized that dysmobility and frailty are precursors of the loss of autonomy, illness and death1,2. In an ideal world, hard clinical endpoints like loss of locomotor capacity, disease onset or death would be the primary endpoints for health intervention studies. However, capturing hard clinical endpoints is resource-demanding, as very large case numbers are usually required. Moreover, the traditional hard endpoints are somewhat crude and do not directly reflect the patients’ well-being3. Therefore, it has been suggested to focus on endpoints that should both matter to the patient and be strongly linked to hard clinical endpoints like mortality etc.3.

The present paper thus discusses which of the currently available muscle tests and measurements could serve as value-based endpoints for interventional health studies. In the methods, the paper explains the different functional roles that skeletal muscle has for our well-being; it then discusses criteria for a rationale selection of endpoints. In the results it reviews established muscle endpoints with regard to their suitability as endpoints for interventional health studies.
Which different functions does skeletal muscle have?

Muscles can primarily be seen as mechanical machines: they can shorten against external resistance. This process, also known as contraction, is based on the cross-bridge dynamics between the actin and myosin filaments within the muscle fibers. During a contraction, the fibers generate force (in Newtons) and velocity (in meters per second), the product of which is mechanical power (in Watts). From a mechanical point of view, five mechanical muscle functions (enlisted below) have to be distinguished (Fig. 1).

In addition, skeletal muscle plays an important role in the energy balance of the human body and it is increasingly recognized that muscle releases so-called myokines into the bloodstream, which have hormone-like effects on other organs. In the following, these different muscle functions are listed and briefly discussed.

Table 1 Different types of muscle contractions

<table>
<thead>
<tr>
<th>Type</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Concentric contraction</td>
<td>Contraction with muscle shortening</td>
</tr>
<tr>
<td>Eccentric contraction</td>
<td>Contraction with muscle lengthening</td>
</tr>
<tr>
<td>Isodynamic contraction</td>
<td>Contraction with constant force or torque</td>
</tr>
<tr>
<td>Isokinetic contraction</td>
<td>Contraction with constant velocity</td>
</tr>
<tr>
<td>Isometric contraction</td>
<td>Contraction without muscle shortening</td>
</tr>
</tbody>
</table>

Note: Most authors assess muscle strength under isometric conditions.

The measurement and interpretation of the force generating capacity of muscle is straightforward when using isolated muscle fibers from tissue samples. More caution is required when muscle strength is measured in humans with so-called dynamometers. Here, one mostly measures joint torque (instead of force). A point of criticism with such dynamometric assessments relates to the artificial contact points that are used (e.g., the shinbone for knee extension), which are perceived as unnatural by the test subject. This and other factors lead to exaggeration of strength readings in sequentially repeated testing sessions. These ‘learning effects’ have been interpreted as central nervous adaptations and one can only speculate as to how far they are also affected by the subject’s ability to tolerate (sub-conscious) discomfort in the unnatural setting of dynamometer testing. In any case, non-stationary results from repeated baseline testing can introduce undesirable bias into study designs. Most studies incorporate familiarization sessions in order to circumvent this issue. However, in my own experience, the ‘learning’ effects are effective for at least four subsequent sessions and it is unclear how far familiarization sessions prior to longer interventions (e.g., months or years) prevent bias at the end of the intervention.

Velocity

It is the muscle’s shortening that generates movement. The greater the contraction velocity is, the faster the movement. Muscle fibers of the fast type have a maximum shortening velocity, which is about 5 times greater than muscle fibers of the slow type and can therefore produce faster movements. Less well known but just as important is the fact that the maximum contraction velocity differs within a given fiber type and that the variation can be quite as much as between fiber types. This can be caused by oxidative damage and other post-translational protein modifications that affect functionality of the contractile apparatus of muscle cells and thereby reduce contraction velocity.

Force

The ability to generate force, also referred to as ‘muscle strength,’ is muscle’s best-known attribute. Of note, force generation without length changes (i.e. under isometric conditions, Table 1) in reality only matters when no movement is intended, for example when standing or holding objects. However, such activities are normally not limited by muscle strength, at least not in healthy people. In weakened people, however, standing can become difficult due to muscle weakness. Thus, isometric muscle strength is probably less important to people’s autonomy and well-being than often thought.
Maximally fast contractions against no resistance at all are physically impossible, given that bones and indeed muscles themselves have a mass. Thus, the accelerated mass is minimal when only a limb is moved relative to the body, e.g. a rapid lunge to prevent a fall. However, measuring the maximum rate of shortening is already difficult in muscle fiber preparations and there is currently no standard way to do it in humans.

### Power

Since contractions are normally performed against resistance, the most universal physical attribute of contractions is its mechanical power, which happens to be the product of force $F$ and velocity $v$. Power $P$ can also be seen as work $W$ per unit time $t$, according to the following equation where $s$ denotes the displacement

$$P = \frac{W}{t} = \frac{Fs}{t} = F \cdot v \quad \text{(Equation 1)}$$

It is important to understand that each engine has its own force-velocity characteristic curve and that the physiological curve for muscle is modelled by the hyperbolic relationship (Fig. 2) proposed by A V Hill almost a century ago. It should also be understood that the exact parameters of the hyperbolic curve are determined by several different factors, such as fiber type or post-translational modifications elicited by training status.

There are a number of ways in order to assess peak power output in humans, such as isokinetic dynamometers, the so-called Nottingham power-rig, or peak power in a vertical jump test. Of these tests, the vertical jump test has the advantage that it uses body mass as the source of resistance, thus using a normalization that is meaningful for real world outcomes (see ‘ecological validity’ below). The chair stand-up test (also called chair-rising test), which is frequently used in geriatrics, also constitutes a measure of muscle power, as the task is to lift the body’s center as rapidly as possible for a number of times (see Equation 1).

Measures of maximal power depict learning effects that are substantially smaller than in measures of muscle strength. Moreover, the age-related decline in muscle power seems to amount to about 1% per year, with good consistency across studies and assessment methods. In the same studies, measures of muscle strength depicted lower reproducibility, and also more diverging age effects than power measures. In summary, therefore, power seems an entity that matters to habitual activity, which can conveniently be assessed in humans.

### Elastic properties

Skeletal muscles connect to bones via tendons and skeletal muscle also has its own connective tissue, namely epimysium, perimysium and endomysium. This arrangement seems essential to permit elastic storage of energy, which is a prerequisite for efficient bouncing movements, such as running or walking. Another advantage of elastic storage is the buffering of impact forces and thereby the reduction of peak forces within the human body. The physical parameters that describe elastic spring systems are the stiffness and amount of energy return. Two methods have been proposed, namely sinusoidal perturbation and quick release methods. For sinusoidal perturbation testing, a series of different vibration frequencies are externally superimposed onto a constant-force contraction and this is repeated at different constant-force levels. Under the assumption of e.g. a Hill-type muscle model (contractile element, parallel-elastic and series-elastic elements), the viscous and elastic properties can be extracted from the fitted models. However, the method has not yet been validated and accuracy and trueness are unknown. Thus, the sinusoidal perturbation method is far from being a suitable clinical endpoint. For the quick release method, viscous properties are negated and elastic recoil is measured through changes in acceleration and in position during the initial 20 ms of a quickly released muscle following a constant-force contraction.

### Energy dissipation

Slowing down a moving object or indeed the human body itself, is as important as accelerating it. To do so, kinetic and elastic energy have to be converted into heat. Although only few muscle researchers have studied energy dissipation, it could theoretically be of great importance, particularly in relation to falls and fractures. Namely, when the potential energy of a fall,
transformed into kinetic energy at impact on the ground, cannot be dissipated within the right time, energy will be transferred to the bone, which can lead to its destruction. Some support for this idea is provided by a large prospective training study, in which resistance training has halved the number of fractures, without any change in bone or in the frequency of falls\(^1\). Of course, the exact reasons for this surprising finding remain elusive, but a training-induced effect that directly (e.g. via alterations in the eccentric force-velocity relationship)\(^2\) or indirectly (e.g. via muscle mass or prolongation of the impact) improves energy dissipation would offer a plausible explanation.

Currently there are no established methods for measuring the capability of muscles to dissipate energy. In addition, one should also note that energy dissipation is related to heat generation, which could in itself constitute another function of skeletal muscle (that is not considered here).

### Substrate utilization

To cover their energy expenditure, muscles absorb fats and sugars from the by-passing blood. At rest, skeletal muscle is a rather moderate energy consumer, but it can become the biggest producer in the human body when working at maximum capacity. This is because of the large fraction that muscles constitute of our body and also because of the excellent oxygen extraction capability of the working muscle. However, skeletal muscle affects the body’s energy metabolism not only during physical work but also after completion of an exercise bout (so-called excess post-exercise oxygen consumption)\(^3\). Of course, the exact reasons for this surprising finding remain elusive, but a training-induced effect that directly (e.g. via alterations in the eccentric force-velocity relationship)\(^4\) or indirectly (e.g. via muscle mass or prolongation of the impact) improves energy dissipation would offer a plausible explanation.

Measurement of the human body’s energy turnover by oxygen consumption (also called indirect calorimetry) is straightforward. When measuring carbon dioxide dissipation at the same time, one can also differentiate between carbohydrate and lipid utilization. Assessment of substrate utilization relies on the respiratory quotient (RQ), which is obtained by dividing carbon dioxide dissipation by oxygen uptake. The rationale is that lipid combustion requires more oxygen per carbon dioxide than carbohydrate or amino acid combustion. Assessment of the RQ, notably, is done under the assumption of steady-state conditions and it therefore requires measurements over at least 20 min. Thus, RQ must not be confused with the respiratory exchange ratio (RER), which technically is obtained by a seemingly identical division. However, alterations in the RER are profoundly linked to the emergence of lactate during exercise, which by definition must be prevented for RQ measurements.

Recent results from bed rest studies suggest that immobilization leads to reduced flexibility in the selection of substrate species, likely reflecting immobilization-related shifts in metabolic pathways within skeletal muscle cells\(^5\). Moreover, the ability of muscles for amino acid uptake is affected by age and training state\(^6\). However, there is currently no standard approach to assess the role of muscles as a substrate consumer within the organism’s metabolism.

### Provision of substrates

It is well-established that the skeletal muscle releases energy equivalents in the form of lactate to other organs such as the heart, brain and kidney\(^7\). This ‘lactate shuttle’ is constitutively active and supports gluconeogenesis in the liver and kidney, and the shuttling is enhanced when oxygen becomes unavailable to the working muscle. Furthermore, skeletal muscle can also provide amino acids for gluconeogenesis (so-called glucose-alanine cycle), especially in phases of carbohydrate deficiency. Further candidates for energy shuttles to other organs are certain forms of lipoproteins.

The measurement of the lactate concentration in the blood is widely established. However, this does not directly reflect energy flows from the skeletal muscle to other organs because accumulation of lactate in the peripheral blood could either be due to increased lactate accrual from the muscle or due to reduced lactate utilization in the receiving organs. Disentangling this would be technically possible with stable isotopes\(^8\), which seems feasible in the field of fundamental human physiology studies but difficult in translational studies.

### Secretory function

In addition to energy equivalents, skeletal muscle also releases signaling agents into the bloodstream as a means to convey information to other organs. For example, interleukin-6 (IL-6) is produced in response to exercise in the muscle and secreted into the blood stream\(^9\), where it signals the muscle’s need for carbohydrates\(^10\). Next to IL-6, there are also other so-called myokines, such as myostatin, brain-derived neurotrophic factor (BDNF) and potentially many others that are yet to be discovered. Another example is the recently discovered irisin, which promotes the browning of white adipose tissue in response to exercise\(^11\) and also impinges on bone\(^12\). It seems obvious that the secretory function of muscle could be extremely important for human well-being. However, research in this area is still in its infancy, thus it is currently difficult to develop valid clinical endpoints on skeletal muscle secretory function.

### What makes a good muscle endpoint?

#### Ecological validity

Good endpoints for health intervention studies should be relevant to everyday life, so that they allow prediction of a patient’s well-being and fate. To give an example, the short performance physical battery (SPPB) can impressively predict...
hospital admissions and death\(^1\). Similar good predictions are possible on the basis of the real world gait speed\(^{2,29}\). Ecological validity of endpoints is therefore demanded by regulatory authorities, e.g. for drug approval studies. Thus, hard clinical endpoints or value-based endpoints\(^2\) should be chosen as primary endpoints for translational studies (Table 2).

**Table 2 A non-inclusive list of hard clinical endpoints for exercise intervention studies**

<table>
<thead>
<tr>
<th>Examples of hard clinical endpoints</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mortality</td>
</tr>
<tr>
<td>All-cause cardiovascular mortality</td>
</tr>
<tr>
<td>Falls</td>
</tr>
<tr>
<td>Fractures</td>
</tr>
<tr>
<td>Onset of walking disability</td>
</tr>
</tbody>
</table>

Construct validity

It generally enhances the acceptance of a therapeutic strategy, if scientists and practitioners understand the mechanism of action. In other words, good studies should ‘make sense.’ To demonstrate this, a study should include explanatory parameters that can explain the study outcome. In the current context, the increase in Real World Gait Speed could, for example, be explained by an increase in muscle mass or muscle power. Therefore, including measurements of muscle mass and muscle power as secondary endpoints into the study design enhances construct validity. This is not only helpful to improve confidence in the outcome but can also help to individualize the therapeutic intervention on the side of the practitioner.

Accuracy

Accuracy of a measurement is the closeness of the observed value to the true value\(^3\). It consists of two components: true-

\[
t = \frac{\bar{X}_D}{S_D / \sqrt{n}}
\]

(Equation 2)

Where \(\bar{X}_D\) and \(S_D\) are the average and standard deviation of the pre-post differences and \(n\) is the number of independent observations. \(S_D\) itself is composed of the measurement-related variation \(u_M\) and the inter-individual variation in the intervention response \(u_{IR}\). Under the assumption of normal distribution they are related by

\[
s_D = \sqrt{2 \cdot u_M^2 + u_{IR}^2}
\]

(Equation 3)

Solving equation 2 for \(n\) and substituting \(S_D\), we find that

\[
n \propto 2 \cdot u_M^2 + u_{IR}^2
\]

(Equation 4)

It thus becomes apparent that the measurement precision is mathematically more relevant to the required study sample size \(n\) than the inter-individual variation to the response. Note that the factor 2 in equation 4 results from the fact that \(u_M\) affects pre- and post-measurement. This can be relieved by obtaining 2 or more measurements at baseline, which has the additional benefit of assessing test-retest reliability.

Bias

There are many ways in which endpoint assessments can be systematically affected and this can lead to misinterpretation of results. For example, assessment of the maximum isometric muscle strength has low precision unless strong verbal encouragement is provided. However, verbal encouragement can easily introduce a systematic error, also called bias, e.g. if the examiner has an interest in the study outcome. To reduce bias, raters and examiners should be blinded against group assignment of a given individual. Another type of bias consists of learning effects. If, for example, study participants train on the same machine on which the endpoint is measured, then the
learning effect leads to study bias that will be misinterpreted as the effectiveness of a training method. To exclude this, endpoints should be assessed with a method that is different from the intervention method.

**Scaling problems**

Some muscle endpoints must be considered as extensive variables. For example, it seems obvious that people with greater body height have a larger muscle mass than shorter people. However, this does not automatically mean that tall people are healthier just because of their larger muscle mass. Although ways of normalization have been proposed and established, the decision for a given normalization method is always somewhat arbitrary. Therefore, it seems preferable to use such endpoints that are indifferent to body dimensions and therefore do not require further normalizations.

**Economic and other boundary conditions**

Finally, the resources required for an assessment (e.g. staff, hours, space and hardware, analysis hours) are important criteria for endpoint selection. Equally important is the acceptance of the measurement method by the study participants, quite as much as availability and applicability of the assessment within.

**Overview of muscle endpoints and their suitability for translational studies (author’s choice)**

The selection of endpoints naturally depends on the hypotheses of a given study. Thus, it is within the investigator’s responsibility to identify endpoints that challenge these hypotheses in the most convincing way. However, two rules may generally apply. Firstly, one normally wants to include measures of muscle mass as well as measures of muscle function, so that effects can be attributed to muscle quantity and/or quality. Secondly, one would mostly prefer endpoints that have better precision than others, given that this will enhance the study power. One needs to bear in mind that precision for a given test may vary between labs and also between study cohorts. Moreover, it is always a good idea to perform two repeated baseline measures, firstly in order to assess precision within a given study and secondly in order to explore the presence of learning effect. In the following, some additional aspects regarding specific muscle endpoints are highlighted.

**Short physical performance battery (SPPB)**

This instrument aims at assessing physical capacity in older and frail people, with the subcategories balance, walking speed and chair-rising test. It thus assesses measures that are obviously involved in the physical functioning and its score predicts disability, falls, loss of autonomy and death. The test’s ecological validity and the construct validity therefore rate high. Moreover, it has excellent precision and does not involve any obvious scaling problems. Regarding limitations, one has to realize that the SPPB is a composite score, which means that different categories are lumped together to a compounded construct. In the case of the SPPB, it turns out that its predictive value is mostly driven by the category ‘walking speed’.

Accordingly, the composite nature of the score has been challenged, in particular with regard to the categories ‘balance’ and ‘chair-rising.’ In addition, both the chair-rising and the walking category of SPPB are associated with sizable learning effects.

**Real world gait speed (RWGS)**

Wearable devices with 3-dimensional mobile accelerometry have in the past decade been pushed to frontier from lab-based assessment towards measurements within people’s habitual environment. Whilst most lab-based muscle tests aim at maximal capabilities, the paradigm is now shifting from ‘can-do’ to ‘do-do’, meaning that real world assessments address the actual exposure to physical activity. In this way, RWGS is different from the walking-category within SPPB, with the latter being a lab-based maximal test. Construct validity and ecological validity of RWGS are high, precision and trueness are excellent, the measurement is unbiased and there seem to be no scaling problems. The main limitation of this method is that one may underestimate the operational effort in sending and receiving recording boxes.

**Vertical jump test**

This test computes peak power that is developed during a counter-movement jump. It thus tests the entire chain of anti-gravity muscles, which is required in many habitual physiological activities. Ecological validity and construct validity therefore are high. Its precision is superior to SPPB, chair-rising test and maximal gait speed and it does not depict any sizable learning effect. Moreover, the vertical jump test can be applied across a wide range of ages and abilities. When normalized to body mass, there is no further scaling adjustment required.

As a side note, people unfamiliar with the test often feel that ‘old and sick people cannot jump’ and that accordingly one could not use this test in clinical or frail populations. However, this is not true, as vertical mechanical power can be assessed even when the jump is small or indeed also when there is no lift-off at all. Moreover, in my own experience approximately 50% of those frail people who are unable to perform a chair-rising test can still perform a vertical jump test and one rarely sees people who can perform a chair-rising test but not a jump test.
course, it sometimes requires guidance by an experienced operator to convince frail people that they can indeed jump. In any case, one can argue that the vertical jump test is easier to perform than the chair-rising test, which is a standard assessment in geriatric medicine.

Dynamometric testing

There is a great quantity of devices commercially available to test for almost all muscle groups of the human body, with grip strength, knee extension, plantar flexion and back extension being the mostly used modalities. Most devices can be set to either isometric mode (i.e. with no movement) or to isokinetic mode (i.e. with movement at given angular velocities), which then yields a measurement of power. It is important to note that single joint movements are not as directly related to habitual physical activity as walking or jumping. Moreover, as explained above, isometric force capacity does not seem to be so relevant to habitual activities compared to muscle power. Thus, construct validity seems to be somewhat limited for muscle strength. In line with this, ecological validity is usually better for dynamometric measures of power than for muscle strength. However, for unknown reasons grip strength turns out to be an independent predictor of hard clinical endpoints.

For all dynamometric testing there is the issue that it requires strong verbal encouragement, which introduces rater-dependent bias and which may also explain why precision of the method is often only moderate to low.

Muscle mass

There are several established methods, including magnetic resonance imaging (MRI), computed tomography (CT), B-mode ultrasound (US; assesses 2-dimensional crosssections as a surrogate for 3-dimensional mass), dual energy X-ray absorptiometry (DEXA; assesses fat-free mass as a surrogate of muscle mass) and BIA. Ample evidence demonstrates their ecological validity, and given the crucial role of muscle for locomotion and autonomy, construct validity is also rating high. Precision and trueness vary for different methods but are generally good. A disadvantage associated with US, is that it is rater-dependent and labor-intensive. On the other hand, analyzing MRI and CT data for muscle volume can also be very time-consuming.

Sarcopenia scores

Loss of muscle mass (sarcopenia) and deterioration of muscle function have long been recognized as causal factors of frailty. To be able to clinically diagnose sarcopenia, the European Working Group on Sarcopenia in Older People (EWGSOP) has defined a framework for diagnostic criteria. The original algorithm has recently been updated and now explicitly defines cut-off values for grip strength, for the chair-rising test, for gait speed and for appendicular muscle mass, and cut-off values have also been defined for the vertical jump test. This framework can be regarded as an ideal starting point for the selection and definition of endpoints in translational studies. Where justifiable, e.g. by considerations of precision or validity, the spectrum of endpoints naturally can be customized to best address the hypotheses in a given study design.

Metabolism

There is no general consensus on how to assess the involvement of muscles in the body’s metabolism. As mentioned above, indirect calorimetry is widely established and can assess energy turnover at rest and during work and it can also assess substrate utilization with regard to carbohydrates and lipids. It should also be considered that assessment of resting metabolic rate is operationally somewhat demanding, as test subjects have to be fasted and not have performed previous exercise for this test.

Furthermore, the homeostatic model assessment (HOMA) index needs to be mentioned, which is well-established for the assessment of insulin resistance and only requires serum levels of insulin and glucose. Similarly well-established but somewhat more resource-demanding is oral glucose tolerance testing (OGTT). Although all of these metabolic tests are not directly linked to muscle, they currently seem best to study skeletal muscle impact on metabolism in translational human studies.

Conclusion

Of the eight muscle functions that matter for human autonomy and well-being, established endpoint tests exist for two functions only, which happen to be mechanical functions (force and power). Table 3 summarizes the main features of the most established muscle endpoints. Well-established metabolic tests exist, which however are not specific to muscle function. This is currently limiting our understanding of the mechanisms by which muscle is involved in clinical metabolic disorders. It can be very rewarding to fill these important technological gaps and to develop and establish methods that assess the braking power of muscle, the elastic storage capacity, the utilization and release of substrates by the musculature, as well as signaling through myokines. For application in translational studies, primary endpoints should be selected that have good ecological validity and construct validity and for which accuracy is commensurate to the study sample size. Finally, it is always good practice to assess test–retest reliability.
Table 3  Overview of the main pros and cons of the most established muscle endpoints

<table>
<thead>
<tr>
<th>Test</th>
<th>Pros</th>
<th>Cons</th>
</tr>
</thead>
<tbody>
<tr>
<td>SPPB: Composite score of categories balance, chair rising and maximal walking speed</td>
<td>Predictive of dysmobility, loss of autonomy and death; Excellent precision</td>
<td>Composite score; Predictive value mostly driven by category, maximal walking speed; Categories, chair rising and maximal walking speed depict learning effects; Probably not sensitive in people with good physical competence (e.g. athletes)</td>
</tr>
<tr>
<td>RWGS: Assessment of habitual walking speed</td>
<td>Predictive of dysmobility, loss of autonomy and death; Excellent precision; Inobtrusive</td>
<td>Effort in logistics often underestimated; Probably not sensitive in people with good physical competence (e.g. athletes)</td>
</tr>
<tr>
<td>Vertical jump test: Main endpoint is peak power in a single countermovement jump</td>
<td>Predictive of frailty and sarcopenia; Excellent precision; Sensitive in a wide range, from athletes to frail people</td>
<td>Predictive value for dysmobility, loss of autonomy and death still unclear; Apprehension exists to do the test in some frail people</td>
</tr>
<tr>
<td>Muscle strength: Mostly defined as maximal force during isometric contractions</td>
<td>Grip strength predictive of dysmobility, loss of autonomy and death</td>
<td>Predictive value debatable for strength tests other than grip strength; Outcome operator-dependent (verbal encouragement); Sizable learning effects</td>
</tr>
<tr>
<td>Muscle mass: by MRI, DEXA, US or BHA</td>
<td>Predictive of dysmobility, loss of autonomy and death; Good precision</td>
<td>Large effort in analysis of MRI, DEXA or US images</td>
</tr>
</tbody>
</table>

Note that scores for the definition of sarcopenia require an assessment of muscle mass and at least one measure of muscle power or force. SPPB: short physical performance battery; RWGS: real world gait speed; MRI: magnetic resonance imaging; DEXA: dual energy X-ray absorptiometry; US: B-mode ultrasound; BIA: bio-electrical impedance analysis.

References

12 Davies CTM, Rennie R. Human power output. *Nature* 1968, 217:770. DOI: 10.1038/217770a0
glucose intolerance even when fat mass is stable. *J Clin Endocrinol Metab* 2018, 103: 1910-20. DOI: 10.1210/je3.2017-02267


40 Hardesty SA, Gregson CL, Rittweger J, Crabtree N, Ward K, Tobias JH. Jump power and force have distinct associations with cortical bone parameters: findings from a population enriched by individuals with high bone mass. *J Clin Endocrinol Metab* 2014, 99: 266-75. DOI: 10.1210/jc.2013-2837


