European Journal of Physical and Rehabilitation Medicine 2017 October;53(5):759-74
DOI: 10.23736/S1973-9087.17.04218-6

SYSTEMATIC REVIEW

Quality of the tools used to assess aerobic capacity in people with multiple sclerosis

Maxime VALET [1, 2], Thierry LEJEUNE [1, 2] *, Jean C. HAKIZIMANA [2], Gaëtan STOQUART [1, 2]

[1]Neuro-Musculo-Skeletal Lab (NMSK), Université Catholique de Louvain, Institut de Recherche Experimentale et Clinique (IREC), Brussels, Belgium; [2]Service of Physical Medicine and Rehabilitation, Cliniques Universitaires Saint-Luc, Brussels, Belgium

*Corresponding author: Thierry Lejeune, Service of Physical Medicine and Rehabilitation, Cliniques Universitaires Saint-Luc, Avenue Hippocrate 10, 1200, Brussels, Belgium. E-mail: thierry.lejeune@uclouvain.be

ABSTRACT

INTRODUCTION: Assessments of physical fitness, including exercise tolerance functions, are valuable in persons with multiple sclerosis (MS). Many tools with widely varying advantages and disadvantages have been used to assess physical fitness in research and clinical practice. To date, there are no recommendations regarding the best tools to use for this purpose in persons with MS. This study aims to systematically review the psychometric properties of the tools used to assess exercise tolerance functions in persons with MS, and to propose recommendations regarding the best test to use.

EVIDENCE ACQUISITION: The literature was searched (PubMed, SPORTdiscus, PEDro, MEDLINE, Embase *via* Scopus, CINAHL, and PsycInfo) to identify the tools most frequently used to assess exercise tolerance functions. These tools were systematically analyzed.

EVIDENCE SYNTHESIS: Forty-eight articles were selected. Six tools or categories of tools concerning exercise tolerance functions were identified. Whole-body exercise tests combined with gas exchange analysis had the best psychometric properties (*e.g.*, validity, reliability) for assessing aerobic capacity in pwMS with mild to moderate disability (Expanded Disability Status Scale [EDSS] ≤6.5). Although sometimes used for this purpose, walk tests seemed to assess walking performance rather than exercise tolerance functions. The psychometric properties of other tests had scarcely been studied.

CONCLUSIONS: The tools vary widely in quality. Whole-body exercise testing combined with gas exchange analysis has the best psychometric properties of the reviewed tools. If gas exchange analysis is feasible, whole-body exercise tests combined with gas exchange analysis, with maximal exercise effort for pwMS with EDSS ≤4 and submaximal exercise effort for pwMS with EDSS ≥4.5, should be recommended to assess exercise tolerance, both in research and in clinical practice. A selection algorithm is proposed.

(*Cite this article as:* Valet M, Lejeune T, Hakizimana JC, Stoquart G. Quality of the tools used to assess aerobic capacity in people with multiple sclerosis. Eur J Phys Rehabil Med 2017;53:759-74. DOI: 10.23736/S1973-9087.17.04218-6)

Key words: Multiple sclerosis - Exercise test - Exercise tolerance - Physical fitness - International Classification of Functioning, Disability and Health.

Introduction

Multiple sclerosis (MS) is a dysimmune disease with unknown causes that affects the central nervous system (CNS) and is characterized by inflammation, demyelination and neurodegeneration.[1] MS affects approximately 2.5 million patients worldwide and is the leading non-traumatic cause of permanent disability in young adults.[2]

Physical fitness can be defined as 'a set of attributes that people have or may acquire that relates to the ability to perform physical activity'.[3] It has traditionally been divided into five components: cardiorespiratory fitness, body composition, muscular strength, muscular endurance and flexibility. People with MS are mainly deconditioned, especially regarding cardiorespiratory fitness, and this deficiency in physical fitness is associated with a number of negative outcomes, including fatigue, low health-related quality of life and impaired social participation and family life.[4] A recent meta-analysis found that, on average, the aerobic capacity of patients with MS, assessed by peak oxygen uptake (VO_{2peak}), was far below a threshold indicating a substantial reduction in all-cause mortality.[5] This physical deconditioning is partly due to physical inactivity, which is highly prevalent among persons with MS, and neurological disability.[6] Therapeutic exercise programs can improve cardiorespiratory fitness among patients with MS and, subsequently, other determinants of these patients' quality of life.[7]

The International Classification of Functioning, Disability and Health (ICF) is a framework for measuring health-related outcomes that was proposed by the World Health Organization in 2001.[8] This framework is recommended to assess the impact of MS in different domains.[9] The physical fitness components pertain to the body functions domain ("b"). In this classification, the classically called "cardiorespiratory fitness" component is defined by exercise tolerance functions (b455), which are divided into 4 functions: physical endurance (b4550), aerobic capacity (b4551), fatigability (b4552) and other exercise tolerance functions (b4558-b4559). However, among patients with MS, fatigability is considered a distinct entity and should therefore be separated from exercise tolerance functions.[10] In this review, "exercise tolerance functions" and "aerobic capacity" will be used interchangeably, as aerobic capacity is by far the most important component of exercise tolerance functions.

Among persons suffering from MS, aerobic capacity has been linked to relevant parameters from all ICF domains and to patients' quality of life.[5] It is thus worth assessing exercise tolerance functions among patients with MS.[11] Evaluating physical fitness helps identify the most suitable treatment for a patient. Frequent assessments of patient's fitness allow for the effectiveness of implemented treatments to be determined and for interventions to be adapted when necessary. In a research setting, this could help determine the effect of codified interventions on physical fitness and its correlations with other important determinants of quality of life. Therefore, it is important to use measurement tools that are valid, reliable, accurate, responsive, acceptable and easy to administer to perform high-quality evaluations.

When examining the clinical and research practices, it is worth noting that many different tools and protocols are used to assess physical fitness, particularly aerobic capacity. The quality of these tools varies widely, and it appears that there is no evidence-based consensus regarding what instrument should preferentially be used to assess the exercise tolerance of patients with MS. The first aim of this review was thus to determine what tools are commonly used to assess aerobic capacity among these patients. Additionally, this study aimed to clarify which tools have a sufficient quality to be used in clinical trials and in routine clinical practice. Following this analysis, our final aim was to propose recommendations for choosing a tool to evaluate the exercise tolerance functions of patients suffering from MS.

Evidence acquisition

Our study protocol has not been previously published. The Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines were applied when relevant.[12] To identify tools used to assess aerobic capacity, a systematic search of the literature was conducted. First, two authors identified studies involving assessment of exercise tolerance functions among patients with MS (M.V. and J.C.H.).

Studies had to:

— use a well-defined tool that aimed to assess exercise tolerance functions or an assimilated concept (*e.g.*, aerobic capacity, physical endurance, cardiorespiratory fitness);

— include only patients with MS (with or without healthy controls), regardless of the severity or disease pattern;

— be published after 2005 in a peer-reviewed journal, in order to include recently used tools only;

— be written in English or French.

We searched 7 databases (PubMed, SPORTdiscus, PEDro, MEDLINE, EMBASE (*via* Scopus), CINAHL and PsycInfo) with the following terms (as established for PubMed): "multiple sclerosis" AND ("endurance" OR "aerobic capacity" OR "cardiorespiratory fitness" OR "exercise tolerance"). The last search was conducted on 20 October 2016. Interventional and observational original studies were included, as were expert recommendations, whereas systematic reviews and meta-analyses were excluded. The reference section of each meta-analysis was screened to ensure that no potentially relevant study was excluded (MV). When several articles referred to a single study and included the same assessment tool, only one report was included, based on completeness of the description of the tool. Tools that aimed to assess aerobic capacity were then systematically extracted and described (M.V. and J.C.H.).

We subsequently analyzed the psychometric and clinical properties of the tools or types of tools identified. A second unrestricted search of databases was performed to identify the psychometric and clinical properties of the extracted tools (MV). This analysis was based on the following criteria, adapted from Andresen (Table I).[13-18]

ICF domain considered

Does the tool really assess exercise tolerance functions (b455)? Does it assess a specific subfunction? Does the tool assess another body structure or function, activity, or participation?

Standards and norms

Is there any standardized protocol? Are there published norms for the general population and for patients with MS?

Floor and ceiling effect

A floor effect occurs when the lowest score of a measurement is achieved by an excessive number of subjects (*e.g.*, >15%) and when this tool is therefore

TABLE I.—- *Criteria used to assess the quality of the tools.*

General criterion	Subcriterion	Definition	Thresholds
ICF domain [13, 14]	–	Does the tool assess exercise tolerance functions (b455)? Does the tool asses a more precise field inside this function?	–
Standardization [13]	–	Is there any standardized protocol?	–
Norms [13]	–	Are there norms for the general population and for persons with MS?	–
Floor effect [15]	–	The lowest score of a measurement is achieved by an excessive number of subjects and when this tool is therefore unable to assess the level of the worst-performing subjects	>15% (+ technical limitations)
Ceiling effect [15]	–	The highest score of a measurement is achieved by an excessive number of subjects and when this tool is therefore unable to assess the level of the best- performing subjects	>15% (+ technical limitations)
Burden [13]	–	Is the tool easy to implement in research and clinical practice? Are the cost in money and time reasonable? Is the test acceptable for the patient and for the assessor? Are the guidelines simple to understand? Are there other logistical obstacles? Is the result easy to interpret?	–
Reliability [16]	Intra-evaluator test-retest Inter-evaluator	Is the tool able to reliably measure the variable across two tests performed consecutively in the same patient? Stability of data between two evaluators or more assessing the same group of subjects	ICC>0.90 (individual use) ICC>0.70 (group use) ICC≥0.75: excellent 0.60-0.74: good 0.40-0.59: fair to moderate <0.40: poor
Validity [16, 17]	Content validity Criterion validity Construct validity	Does the tool really measure what it is supposed to do? Is it well correlated to a “gold standard”? Is it well correlated to other measures evaluating the same concept and is it distinct from other measurement tools that do not assess the same concept	Clear description of the measurement aim, target population and the concepts that are being measured. Pearson’s correlation coefficient (r) >0.90: very high 0.70-0.90: high 0.50-0.70: moderate 0.30-0.50: low <0.30: negligible
Responsiveness [18]		Is the tool responsive to a real change in the parameter?	
	MDC	What is the smallest change corresponding to a real individual variation?	
	MCID	What is the smallest change that could be considered clinically important by the patient or the assessor?	

ICC: intraclass correlation coefficient; MDC: minimal dectectable change; MCID: minimal clinically important difference.

unable to assess the level of the worst-performing subjects.[19]

A ceiling effect occurs when the highest score of a measurement is achieved by an excessive number of subjects (*e.g.*, >15%) and when this tool is therefore unable to assess the level of the best-performing subjects.[19]

Assessor and patient burden

Is the tool easy to implement in research and clinical practice? Is the cost in money and time reasonable? Is the test acceptable for the patient and for the assessor? Are the guidelines simple to understand? Are there other logistical obstacles? Is the result easy to interpret? The answers to these questions allowed us to subjectively evaluate the patient and assessor burden of the test.

Reliability

Intra-evaluator (test-retest): Is the tool able to reliably measure the variable across two tests performed consecutively in the same patient by the same assessor? (For individual case use, an Intraclass Correlation Coefficient (ICC) >0.90 is recommended, whereas for the assessment of a large group, an ICC > 0.70 is generally considered sufficient).[19]

Inter-evaluator: Do the data remain stable between two evaluators or more assessing the same group of subjects? (ICC≥0.75 indicates excellent reproducibility; 0.75>ICC≥0.40 indicates good reproducibility; ICC<0.40 indicates poor reproducibility).[13]

Validity

Does the tool measure what it is supposed to measure (content validity)? Is it well correlated with a gold standard (criterion validity)? Is it well correlated with other measures evaluating the same concept, and is it distinct from other measurement tools that do not assess the same concept (construct validity)?

Responsiveness

Is the tool responsive to a true change in the parameter? What is the smallest change corresponding to a real individual variation (minimal detectable change, MDC)? A commonly used statistical method to evaluate MDC is as follows:

$$\text{Standard deviation (SD)} \times \sqrt{}\,(1 - \text{ICC}_{\text{test-retest}}) \times 2.77$$

What is the smallest change that could be considered clinically important by the patient or the assessor (minimal clinically important difference [MCID]? The MCID is less objective than the MDC and can be estimated in a so-called reproducible test by the following formula:[19]

$$\text{MCID} = 0.5\ \text{SD}$$

The values for which a significant change (P<0.05) was detected in interventional studies reflected the tool's sensitivity to change at the group level.

Evidence synthesis

Six hundred and forty-five articles were identified, 270 of which remained after the duplicates were removed. Of these 270 articles, 208 were excluded based on their title and abstract. The full-texts of the 62 remaining articles were screened, and 14 were excluded (11 because of redundancy of the studies and 3 for not assessing exercise tolerance or an assimilated concept) (Figure 1). Forty-eight articles were retained, and the tools used to assess exercise tolerance functions were extracted.

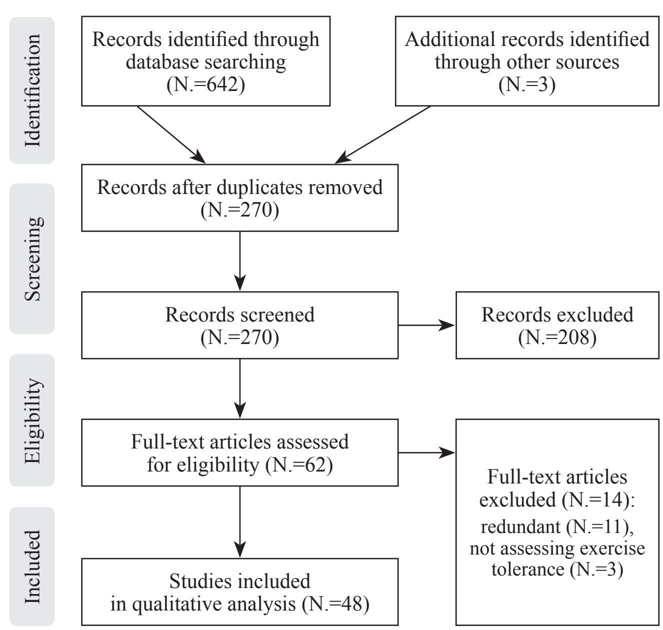

Figure 1.—Flow chart diagram of the review.

The following tests or categories of tests were identified from the selected articles:

1) whole-body maximal exercise testing with gas exchange analysis:
   a) Cycle ergometer;[6, 20-46]
   b) Recumbent stepper;[47-50]
   c) Arm crank ergometer;[50-52]
   d) Treadmill;[27, 53, 54]
2) Whole-body maximal exercise testing without gas exchange analysis (cycle ergometer);[55-57]
3) Whole-body submaximal exercise testing with gas exchange analysis (cycle ergometer);[58-62]
4) 6-Minute Walk Test (6MWT);[4, 63]
5) Modified Canadian Aerobic Fitness Test (mCAFT);[31, 64]
6) Ruffier-Dickson Test.[65]

These tests were further analyzed based on the aforementioned criteria.

Whole-body maximal exercise testing with gas exchange analysis

Of the 48 studies included in the review, a vast majority (*i.e.*, 36) used whole-body maximal exercise testing to assess aerobic capacity. Based on the text, the authors used this type of test to assess "exercise capacity," "physical fitness," "aerobic capacity," and "cardiorespiratory capacity." Four different types of machines were used: cycle ergometers (N.=28), recumbent steppers (N.=4), treadmills (N.=3), and arm crank ergometers (N.=3).

The main parameter measured was the maximal relative (*i.e.*, based on weight, expressed as mL/kg/min) or absolute (expressed as mL/min) oxygen uptake (VO_2) measured during the test (maximal oxygen uptake [VO_{2max}] or peak oxygen uptake [VO_{2peak}], as explained below). Other measurements included heart rate (HR), rate of perceived exertion (RPE, assessed by the Borg Scale,[66] expressed between 6 and 20), power output or work rate (WR) and peak work rate (WR_{peak}). Seven studies included measurements of plasma lactate. The protocol used to reach maximal exercise effort differed depending on the study. Warm-up was often performed for a duration of between 2 and 5 minutes, always without loading. The tests generally started without loading, but a few studies used a test that started with a minimal loading of between 8 and 30 W. The load increased either by step at regular intervals (of between 1 and 3 minutes), or continuously. The rate of increase ranged between 2.5 and 20 W/min, with a mean of 10 W/min.

Direct measurement of VO_{2max} by gas exchange analysis during a stepwise or continuous incremental maximal exercise effort on a cycle ergometer or treadmill is considered the 'gold standard' for assessing aerobic capacity in healthy subjects.[43, 67] However, there is no consensus regarding the standard protocol that should be used. Most studies have used a plateau in VO_2 as the primary criterion, but there is no consensus on what should be considered a plateau, and few authors have clearly defined it.[5] Thus, this criterion is often subjectively determined. The maximal pattern of a test is also frequently based on 3 secondary criteria: a respiratory exchange ratio (RER) (VCO_2/VO_2) equal to or greater than 1.10 or 1.15; an HR of less than 10 bpm under the age-predicted maximum (220 – age or 208 – 0.7 × age); and a RPE of 15 to 18 on the Borg Scale (17 is equivalent to a "very hard" level reported by the patient). Effort is generally considered maximal when at least two of these three criteria are met. Some authors also used a lactate criterion, the so-called fourth secondary criterion, and considered a blood lactate level above 8 mmol/L the threshold for maximal exercise.[30, 32, 58, 68] When maximal effort is reached, the maximum VO_2 measured during the test is more frequently denoted "VO_{2peak}."

Another method that directly assessed cardiorespiratory endurance by an incremental exercise test on a cycle ergometer was used: the Oxygen Uptake Efficiency Slope (OUES).[42] This method is based on the logarithmic relationship between minute ventilation (VE) and VO_2:

$$VO_2\ (\text{mL/min/kg}) = \alpha \times \log VE\ (\text{L/min}) + \beta.$$

The slope of this equation (α) is the OUES, a dimensionless value reflecting "ventilatory efficiency." The accuracy of this measure (assessed by linear regression between logVE and VO_2) is excellent (r=0.94).

Psychometric properties

The American College of Sports Medicine (ACSM) has published norms for VO_{2max} by gender, age and type of test used, including cycle ergometer.[67] Normative values for the OUES score for Caucasian subjects aged

20 to 60 years have been established.[69] However, the presence of any floor or ceiling effect is unknown. No learning effect has been observed for this tool.

A maximal exercise test on an ergometer takes between 20 minutes and one hour, depending on the protocol used, the environment, the assessor and the patient. The equipment required for the test is relatively expensive (ergometer, HR monitor, electrocardiogram, spirometer, etc.). Direct measurement of VO_2 during exercise, both in clinical practice and in research, can thus be inconvenient for the patient and the assessor because it is time consuming and requires certain skills.[43] Moreover, the mask worn during the exercise period can sometimes be annoying for the patient.

The reliability of direct measurements of VO_{2max} by gas exchange analysis among patients with MS is excellent ($ICC_{test\text{-}retest}$=0.98).[32] The reliability of the OUES is also excellent ($ICC_{test\text{-}retest}$=0.89).[45] The criterion validity of VO_{2max} is maximal, as it is considered the gold standard. The OUES is highly correlated with VO_{2max} (r=0.86).[42] Based on the aforementioned formula, we can estimate that among patients with MS, the MDC of the relative VO_{2max} would be between 1.16 and 1.72 mL/kg/min (*i.e.,* between 84 and 114 mL/min in absolute terms). According to the same estimation model, the MCID would be between 2.9 and 4.3 mL/kg/min (210-285 mL/min). Heine *et al.* computed MDC values of 4.6 mL/min/kg and 367 mL/min, respectively, (*i.e.*, 21% of baseline values).[45] Successful interventional studies showed significant improvements in VO_{2max} of 2.9 to 4 mL/kg/min in relative terms [21, 34] and 200 mL/min in absolute terms.[30] The MDC of the OUES was estimated to be 6.6 in fatigued patients with MS.[45]

Whole-body maximal exercise testing without gas exchange analysis

Three studies used a cycle ergometer test without spirometry.[55-57] They all used a maximal exercise testing protocol. Two studies recorded the time of exercising to volitional exhaustion and WR_{peak},[55, 56] whereas the third study assessed heart rate and RPE at the peak.[57]

Psychometric properties

These tests have not been standardized, and no normative values are available. The presence of floor or ceiling effects has not been investigated. A learning effect is unlikely. This type of test takes between 20 and 45 minutes to perform; the burden is similar to the whole-body maximal exercise testing with gas analysis, except that no gas exchange measures are needed, and the costs and patient's and assessor's burden are thus significantly lower.

The reliability of these tests has not been investigated. The criterion validity of WR_{peak} is excellent, as shown by its correlation with VO_{2max} (r=0.94),[43] but its responsiveness is unknown.

Whole-body submaximal exercise testing with gas exchange analysis

Five studies conducted by the same team used a submaximal exercise test performed on a cycle ergometer. They consistently used the same protocol across their studies. Patients were instructed to pedal for 6 minutes at a constant rate against a resistance corresponding to 25% of their theoretical maximal power output. Patients then had to rest while seated for 6 minutes before they performed the same 6-minute bout. In this protocol, exercise tolerance was expressed as the Mean Response Time (MRT). The MRT is defined as the ratio between the calculated oxygen deficit at the end of the 6-minute bout and the difference between resting and steady-state VO_2, expressed in seconds. The lower the MRT is, the better the participant's aerobic capacity.

Furthermore, one study evaluated the validity of the OUES derived from submaximal exercise ($OUES_{submax}$) compared to the same index derived from the total duration of a maximal exercise test.[42]

Psychometric properties

These tests lack a strict standardization, as well as norms, and the presence of a floor, ceiling or learning effect is not known. The burden is similar to that of whole-body maximal exercise testing with gas exchange analysis, except that maximal effort is not necessary, which adds safety and comfort for the patient. The time needed to perform these tests ranges between 30 and 60 minutes.

The reliability has not been reported. The MRT presented a moderate criterion validity (r=0.67), but responsiveness has not been investigated.

Six-minute Walk Test

Of the 48 articles included, the 6MWT was used in one clinical study to assess aerobic capacity and has been recommended by an expert consensus for that purpose.[4, 63] The authors used this test to assess "functional aerobic capacity (cardiovascular fitness)."

The 6MWT is a submaximal test. The patient is instructed to walk for six minutes along a simple pathway on a hard, flat indoor surface (ideally back and forth in a 30 m hallway, marked every 3 m, according to the American Thoracic Society (ATS) Guidelines).[70] The objective is to cover the longest distance possible at a comfortable pace determined by the patient. The patient uses his/her usual walking aids during the test. Standard phrases of encouragement should be provided every minute. The patient is free to stop if he/she wants but should resume walking as soon as possible. The six-minute walking distance (6MWD) is the result of the test and is expressed in meters.

Psychometric properties

The 6MWT is standardized, and normative values are available based on gender, age and weight.[71] The 6MWD has no floor effect, but a ceiling effect, likely linked to the maximum achievable speed without running, has been suggested.[72] There is no learning effect for this test among patients with MS.[73] The administrator's and patient's burden for the 6MWT is low. The test takes less than ten minutes and does not require expensive materials. This test is well tolerated in patients with MS, even when presenting severe gait impairments.[73] Furthermore, it is a safe test regarding locomotor and cardiovascular issues. Finally, the test is easy to understand and interpret.

The 6MWT is a reliable and reproducible test for patients with MS ($ICC_{test\text{-}retest}$=0.96; $ICC_{intraevaluator}$=0.94; $ICC_{interevaluator}$=0.91-0.93).[73, 74] Regarding criterion validity, 6MWD is weakly correlated with absolute VO_{2max} (r=0.25) and moderately correlated with relative VO_{2max} (r 0.48-0.62).[31] The MDC calculated according to the aforementioned formula, assuming an SD of 100 m and an ICC of 0.96, is 55 m.

Modified Canadian Aerobic Fitness Test

Of the 48 studies included, 2 used the mCAFT.[31, 64] One estimated the validity of this tool in predicting VO_{2peak} among patients with MS, while the other described a protocol for a clinical trial.[64] The mCAFT is a multistage submaximal test performed on a double 20.3-cm step. Patients are instructed to step at a regular rhythm and are guided by a musical cadence. The stepping rate is regularly increased. The test stops when the subject reaches 85% of his/her age-predicted maximum heart rate. An equation estimating VO_{2max} based on the stage reached, sex, age and body weight has been developed and validated in healthy subjects (R^2=0.77).[75] Estimated VO_{2max} is considered the main variable of this test.

Psychometric properties

The protocol is well standardized. To the author's knowledge, no normative values have been specifically developed for this test. However, as the final result is expressed in estimated VO_{2max}, normative values of VO_{2max} could be used. The presence of a floor effect, as well as of a learning effect, is unknown. A ceiling effect is likely present in this test, but its magnitude for patients with MS has not been determined. Overall, this test takes approximately 15 to 30 minutes. It requires a specific double step and a CD containing the pacing rhythm. The instructions are quite easy to understand and to apply. The submaximal nature of this test lowers the patient's burden in comparison with a maximal test.

The reliability of the test is unknown, but the criterion validity is good, as indicated by its good correlation with VO_{2max} (r=0.70).[31] The responsiveness of the test has not been determined.

Ruffier-Dickson Test

One study used the Ruffier-Dickson test to assess cardiorespiratory performance.[65] The Ruffier-Dickson test is a submaximal test based on heart rate measurements. Patient's heart rate is measured at rest (HR0), and the patient is then instructed to perform 30 squats in 45 seconds; heart rate is measured immediately afterward (HR1) and after a seated rest of one minute (HR2). Two indices can be calculated from this test: the Ruffier Index and the Ruffier-Dickson Index. The Ruffier Index is calculated as:

$$\text{Ruffier Index} = \frac{HR0 + HR1 + HR2 - 200}{10}$$

the lower the index is, the better the exercise tolerance. Different cut-off scores have been proposed: a score below 0 denotes a very good effort adaptation capacity, while a score over 15 denotes a poor effort adaptation capacity. The Ruffier-Dickson Index is calculated as:

$$\text{Ruffier-Dickson Index} = \frac{(\text{HR1}-70)+2\times(\text{HR2}-\text{HR0})}{10}$$

Again, lower scores indicate better exercise tolerance. A score below 0 denotes excellent effort adaptation capacity, while a score over 10 indicates poor adaptation. It is not clear which index was used by the authors.

Psychometric properties

This test is standardized, and the norms are categorical. The presence of any floor, ceiling or learning effects is unknown. The test takes 5 minutes to perform and has no specific requirements except for a chronometer. As it is a submaximal test, the patient's burden is limited.

Other psychometric properties have not been studied among patients with MS.

Psychometric and clinical properties of the tools are summarized in Table II.

Discussion

The aim of this review was to identify and evaluate the tools used to assess exercise tolerance functions in patients suffering from MS. The search for articles on exercise tolerance in these patients led to the identification of 48 relevant articles of variable quality, reporting a total of 6 tools or types of tools. In this final section, their qualities and limitations will be discussed, allowing for a proposal of recommendations for best-choice tests in patients with MS.

Table II.—*Psychometric properties of the most frequently used and potentially interesting tools to assess cardiorespiratory fitness in MS patients.*

	6MWT	WBMET with gas exchange analysis	WBMET without gas exchange analysis	OUES	WBSET with gas exchange analysis (MRT)	mCAFT	Ruffier-Dickson Test
Unit	m	mL/kg/min or mL/min	W	No unit	s	Estimated mL/kg/min	No unit
ICF domain	Activity	Body function	Body function	Body function	Body function	Body Function	Body function
ICF code	d450 (d4500)	b455 (b4550/b4551)	b455 (b4550/b4551)	b455 (b4550/b4551)	b455 (b4551)	b455 (b4550/b4551)	b455
Standardization	+	≠	±	±	±	+	+
Maximal (M)/ submaximal (S)	S	M	M	M/S	S	S	S
Norms	+	+	–	+	–	±	+ (?)
Floor/ceiling effect*	–/+	?	?	?	?	?/+	?
Learning effect	+	–	–	–	?	?	?
Subjective total burden	–	+++	++	++	++	+	–
Total time taken, min	~10	~30 to 60	~20 to 45	~25 to 60	~30 to 60	~15 to 30	~5
$ICC_{test\text{-}retest}$	0.96	0.98	?	0.89	?	?	?
$ICC_{intraevaluator}$	0.94	?	?	?	?	?	?
$ICC_{interevaluator}$	0.91-0.93	?	?	?	?	?	?
Content validity	?	+	+	+	+	+	?
Criterion validity (VO_{2max})	r=0.48-0.62	Gold standard	r=0.94	r=0.86	r=0.67	r=0.70	?
Construct validity	–	+++	+	+	+	?	?
MDC	20-55 (?)	1.16-4.6 (84-367) (?)	?	1.44-6.6 (?)	?	?	?
MCID	50	2.9 – 4.3	?	?	?	?	?

Data presented in parentheses refer to healthy subjects.
WBMET: whole-body maximal exercise testing; OUES: Oxygen Uptake Efficiency Slope; WBSET: whole-body submaximal exercise testing; MRT: whole-body submaximal exercise testing with gas exchange analysis; mCAFT: modified Canadian Aerobic Fitness Test; ICF: International Classification of Functioning, Disability and Health; ICC: Intraclass Correlation Coefficient; MDC: minimal detectable change; 6MWT: 6-Minute Walk Test; MRT: Mean Response Time; ?: unknown; (?): poor evidence or inconsistent data; +: existing, present; +++: heavy; –: absent; ≠: different standardized protocols used.
*In MS patients with EDSS≤6.5

Whole-body maximal exercise testing with gas exchange analysis

The vast majority of studies assessing exercise tolerance functions used VO_{2max} as the preferred metrics. VO_{2max} is the gold standard for assessing aerobic capacity among healthy subjects.[67] Cycle ergometers were the most commonly used device, but recumbent steppers, arm crank ergometers and treadmills were also used for this purpose. In the current MS literature, "VO_{2max}" and "VO_{2peak}" were used interchangeably.[32] Given the observed consensus on the aim of these exercise tests, the maximal tests conducted on an ergometer were actually designed to assess aerobic capacity, which is clearly included in the body functions domain of the ICF.

Langeskov-Christensen *et al.*, in a study of 9 men and 11 women with MS (EDSS 0-6), studied the validity of the criteria used to determine whether maximal exercise effort had been reached. The primary endpoint (*i.e.*, a VO_2 plateau despite an increase in power output (defined in this study as an increase in VO_2 of less than 50% of the expected increase for the increase in power attained) was achieved by 40% of patients with MS. Regarding the secondary criteria, 65% of patients achieved at least the age-predicted maximal heart rate (Tanaka's formula: HR_{max} = 208 – 0.7 × age) minus 10 bpm; 95% achieved an RER greater than or equal to 1.15; 95% achieved a score equal to or greater than 17 on the Borg Scale; and 75% had post-exercise lactate levels greater than 8 mmol/L. The difference with healthy subjects regarding the rate of attainment of these criteria was not significant (P>0.05). A total of 95% of MS patients (19/20) achieved at least 3 of the 5 criteria.[32] In a similar study, including 56 patients with MS (EDSS from 1 to 6.5), Heine *et al.* found that a RER of 1.10 was achieved by 69.6% of participants; a heart rate equivalent to 90% of the age-predicted maximum heart rate (HR_{max}) was achieved by 48.2%; and 23.2% of patients rated their effort as 18 or higher on the Borg scale.[41] They also showed that the ability to achieve the HR or VO_2 plateau criteria was significantly lower among patients with a moderate level of disability (EDSS 4.5-6) than among those with a mild level of disability (EDSS≤4). They concluded that aerobic capacity assessments by a maximal exercise test are valid in patients with a mild level of disability and that could be limited by motor symptoms in patients with moderate disability.

Regarding the OUES, in patients with HR reaching 90% of their age-predicted maximum heart rate (27/56), the OUES derived from the total duration of maximal exercise ($OUES_{max}$) correlated very well with $OUES_{50}$, measured halfway through the exercise (submaximal exercise) (r=0.93). However, there was a slight bias, as the $OUES_{50}$ underestimated the $OUES_{max}$ by 1.8 (*i.e.*, 7%) on average.[42] Thus, in clinical practice, the $OUES_{max}$ is preferred when maximal exercise effort can be achieved. The average OUES measured in patients with MS was 24.2±7.2 *versus* 46±9 in healthy subjects.[42, 76] Submaximal OUES is therefore a promising method of assessing aerobic capacity without requiring maximal exercise effort. To date, no interventional studies in persons suffering from MS have used this tool.

Langeskov-Christensen *et al.* demonstrated that direct measurements of VO_{2max} by gas exchange analysis had excellent reliability ($ICC_{test\text{-}retest}$=0.98) among patients with MS when a gradual maximal exercise test was performed on a cycle ergometer, using two measurements taken at an interval of 2 to 8 days.[32] Among healthy subjects, Van Laethem *et al.* demonstrated that OUES showed excellent intratest reliability (*i.e.*, logVE was highly correlated with VO_2 in a single test) and had a test-retest reproducibility that was similar to that of VO_{2peak}. However, they also noted that the reproducibility of the OUES became higher when calculated from increasing levels of achieved exercise intensity.[76]

The literature does not mention a floor or ceiling effect for the direct measurement or estimation of VO_{2max} by ergometer. We can hypothesize that for any patient able to pedal against a resistance and therefore able to generate a working rate greater than 0 W, the test does not have a floor or ceiling effect, but this has not been formally demonstrated.

A maximal exercise test on an ergometer takes between 20 minutes and one hour, depending on the protocol used, the environment, the assessor and the patient. The equipment required is significantly more expensive than that required by the methods presented above (ergometer, HR monitor, electrocardiogram, spirometer, etc.). Direct measurement of VO_2 during exercise, both in clinical practice and in research, can thus be inconvenient for the patient and the assessor because it is a time-consuming tool that requires certain skills.[43] Moreover, the mask worn during the exercise period can sometimes be annoying for the patient. The power generated during a maximal test on a cycle ergometer is not strongly influenced by the internal energetic cost of the movement, and gait and balance problems do not interfere with the testing, as could occur during a walk or treadmill test. Therefore, from this perspective, the cycle ergometer may be an ideal tool for patients suffering from MS.[43] In patients presenting with more severe neurological disability (EDSS>6), an arm crank ergometer seems to be the device of choice. However, in healthy subjects, upper limbs exercise elicits only 70% of the VO_{2peak} achievable by exercise involving lower limbs. In patients with milder disability, this type of test could therefore be inappropriate, and a cycle ergometer or recumbent stepper should thus be preferentially used. Direct measurement of VO_{2max} is scored on an easy-to-interpret scale in both clinical practice and research. A major disadvantage of maximal tests when applied in patients with MS is the need to reach maximum exercise effort, which is not always possible for these patients, due to their multiple disabilities. In this context, submaximal tests constitute a safer, more reliable alternative that might be more acceptable for patients with MS.[77]

Unfortunately, there is no standardized protocol used to perform maximal exercise testing with gas exchange analysis. Based on the protocols used and the level of fitness of patients with MS,[5] a 3-minute warm-up period with a minimal load (0 to 10 W) seems reasonable. The test should start at the load determined for the warm-up and be followed by an increment of 10 W/min (for women) or 15 W/min (for men) in minimally disabled patients (EDSS 0-4) or 7.5 W/min in all patients with more severe disability (EDSS>4). The increment rate should be even lower (*e.g.,* 5 W/min) if an arm crank ergometer is used among patients with severe disability (EDSS>6). Ideally, the increment should be continuous, but that would require an electronically braked cycle ergometer with that option. These protocols would reasonably allow the tests to last between 8 and 12 minutes, as recommended by Wasserman *et al.*[78]

Regarding treadmill tests, various protocols have been developed in healthy subjects and diseased individuals.[67] Our review identified 3 studies using a treadmill protocol: 2 using the Naughton protocol and one using the modified Balke protocol.[27, 53, 54] These protocols are based on an increment of, respectively, the slope and the speed, at an individualize rate (Naughton), and the slope at a rate of 2%/2 min. Both tests are stopped at volitional exhaustion or for safety reasons. These protocols are suitable for persons with impaired exercise tolerance.[67] In patients with peripheral arterial disease, VO_{2peak} did not differ between tests performed on a treadmill or on a cycle ergometer.[79] However, limitations in walking capacity are common among persons with MS and could bias this type of test.[80]

Whole-body maximal exercise testing without gas exchange analysis

These protocols aim to assess exercise tolerance functions, but their validity is uncertain, considering the fact that the psychometric properties of these tests have never been thoroughly studied.

In 1990, Storer *et al.* computed an equation predicting VO_{2max} during a cycle ergometer test based on weight, age and WR_{peak} using a multiple linear regression of data collected from 231 healthy subjects. The equation for men was:

$$VO_{2max} = 10.51 \times WR_{peak} + 6.35 \times \text{weight} - 10.49 \times \text{age} + 519.3 \text{ mL/min}$$

and the equation for women was:

$$VO_{2max} = 9.39 \times WR_{peak} + 7.70 \times \text{weight} - 5.88 \times \text{age} + 136.7 \text{ mL/min.}$$ [81]

In 2012, Motl and Fernhall validated this equation for women with MS in a cross-sectional study.[43] In that study, 32 women with MS (EDSS 0-6) and 16 healthy controls performed a test that started at a power of 0 W and increased continuously at a rate of 15 W/min until exhaustion. VO_{2max} is defined as the maximum recorded VO_2 value when 2 of the 3 aforementioned criteria are met. The linear regression analysis between the predicted VO_{2max} and the measured VO_{2max} showed a strong correlation between these two parameters, regardless of subgroup (for patients with MS: r=0.94). The main limitation of this study was that it included only women. The authors concluded that the VO_{2max} prediction from the WR_{peak} could be useful for monitoring patient's fitness over time in clinical practice and for assessing the improvement associated with exercise and other types of rehabilitation programs in research.[43] However, no prospective studies regarding exercise tolerance functions in patients with MS have used this method, which is still disadvantaged by its requirement of maximal exercise effort.

Whole-body submaximal exercise testing with gas exchange analysis

MRT and $OUES_{submax}$ are two types of indices that can be derived from this category of test.

MRT reflects the response of the cardiorespiratory system to an exercise stress. Hence, MRT is an index of aerobic capacity (b4451 in the ICF classification). The psychometric properties of this tool have not been studied in healthy subjects nor among patients with MS. One study, assessing MRT during a step test, reported that the within-subject coefficient of variability for two tests performed in the same week was 4.5%.[82]

The $OUES_{submax}$ is another index that requires gas exchange analysis without the need to attain maximal exercise effort. However, the best protocol and duration of the test needed to ensure optimal validity and its psychometric properties remain unknown.

Many other tools are available to assess exercise tolerance functions based on submaximal cycle ergometer tests without gas exchange analysis (*e.g.*, Astrand-Rhyming Test,[83] $pWC_{75\%}$/kg,[77] and $WCI_{65\%HRreserve}$),[84] but to our knowledge, none have specifically been studied among patients with MS. These tools have the advantage of not requiring a gas exchange analyzer and thus allow for an accurate assessment at a lower cost. They are also technically easier to administer and do not require the patient to wear a mask.

Six-minute Walk Test

Although frequently used in clinical rehabilitation, the purpose of this test occasionally seems unclear among patients with MS.[85] Thus, to clarify the utility of the 6MWT and use it appropriately, determining the domains for which this tool is valid seems important.

Examining more closely the different purposes of the walking tests, it is worth noting that the ICF domain that this tool is meant to target is not clear. Authors sometimes use terms that appear to pertain more to a matter of impairment (*e.g.*, aerobic capacity, walking endurance, specific endurance component of mobility) and sometimes use phrases that address activity limitations (*e.g.*, walking ability, mobility). American Thoracic Society guidelines use the expression "submaximal functional capacity," but none of the articles cited a specifically targeted domain. Savci *et al.*, in a cross-sectional study comparing the 6MWT in 30 healthy subjects and 30 patients with MS with mild to moderate disability (EDSS 1.5-6), concluded that the 6MWT better reflects the ability to achieve activities of daily living (*i.e.,* activity) than maximum oxygen consumption (*i.e.,* body function).[86] In a recent study, Sandroff *et al.* examined the correlates of 6MWD and concluded that 6MWD assessed walking performance, which is an activity (d450 in the ICF classification), rather than aerobic and muscular fitness.[87] Despite the lack of evidence in this field, it can plausibly be hypothesized that 6MWD targets an activity and its limitation, rather than a body function, in patients with MS.

In a review designed to establish the core outcome measures of exercise studies in patients with MS, Paul *et al.* recommended systematically using the 6MWT as an "exercise tolerance" assessment tool.[63] Although often used for this purpose, no studies have demonstrated the validity of this test in measuring aerobic capacity in patients with MS. Sandroff *et al.*, in a cross-sectional study involving 31 patients with MS and 31 healthy subjects, used methods similar to those of Kuspinar [31] and observed that the correlation between relative VO_{2max} and 6MWD was moderate in patients with MS (r=0.62) but strong in healthy subjects (r=0.86).[40] In 44 middle-aged healthy subjects, Burr *et al.* observed only a weak correlation between 6MWD and relative VO_{2max} measured by gas exchange analysis during a 6MWT (r=0.49).[88] Savci *et al.* showed a strong correlation between 6MWD and the ability to perform activities of daily living (assessed by the Barthel Index (BI); r=-0.81, EDSS r=-0.76).[86] The good correlation (>0.70) between the 6MWD and EDSS score seems logical, since this scale is partially based on walking ability. According to a statistical analysis using stepwise multiple regression, the BI, resting heart rate and Fatigue Severity Scale (FSS) score are independent predictors of the 6MWD and propose a predictive equation with a good correlation coefficient (r=0.90). Furthermore, Goldman *et al.* noted a good correlation between the 6MWD and subjective walking ability (Multiple Sclerosis Walking Scale, MSWS-12, r=-0.81) and a moderate correlation with the physical dimension of fatigue (Modified Fatigue Impact Scale, physical subscale, MFISphy, r=-0.66).[73] In a recent study, Sandroff *et al.* compared measures of walking performance (timed 25-foot walk and spontaneous gait velocity) and physical fitness (VO_{2peak} and lower limb muscular strength) as head-to-head predictors of 6MWD performance in patients with MS (EDSS 1.5-6.5).[87] They concluded that 6MWD assesses walking performance rather than aerobic and muscular fitness, regardless of the EDSS score. Their study was interesting but had two limitations: first, the authors did not measure the oxygen cost of walking during the test, and second, they did not stratify their analysis for the EDSS score, precluding any interpretations regarding disability-based subgroups.

Goldman *et al.* observed no learning effect among patients with MS, contrary to what was observed in other populations (healthy subjects and fibromyalgia), in whom the authors reported a learning effect between 4% and 33% capping the second repetition, likely because the patients adopted a more regular pace.[89, 90]

The responsiveness of this tool has not been specifically studied among patients with MS, but it has been studied in a geriatric population (N.=592, mean age 74.1 years) and in a population of patients with stroke (N.=100).[91] They estimated that the MDC was approximately 20 m and the MCID approximately 50 m. These results seem compatible with the standard deviation of the 6MWD found in different studies (SD=~100 m). However, the MDC calculated according to the aforementioned formula, assuming an SD of 100 m and an ICC of 0.96, was 55 m. The included studies reported significant improvements in 6MWD at the group level that ranged between 15 m and 34 m on average.

The 6MWT has an acceptable administrator and patient burden, takes less than ten minutes to complete, and does not require any expensive materials. The test is well tolerated among patients with MS, even those presenting with severe gait impairments.[73] Furthermore, it is a safe test regarding locomotor and cardiovascular issues. Finally, its instructions are quite simple for the patient to understand, and the results are easy to interpret, which is one advantage of the test.

Modified Canadian Aerobic Fitness Test

The Canadian Home Fitness Test (CHFT), father of the mCAFT, is a submaximal step test that was developed in 1976 by Shephard and Bailey. These authors suggested that this test could provide an "*approximate measure of an individual's physical fitness in order to stimulate an increase in physical activity*."[92] The test was modified by Weller *et al.* in 1993: stages were added for subjects to reach 85% of their age-predicted maximal heart rate. They proposed a prediction equation estimating actual VO_{2max} based on the attained level, age and body weight (R^2=0.77).[75] In 2010, Kuspinar *et al.* studied whether VO_{2max} could be predicted by submaximal tests, including mCAFT, among patients suffering from MS. They observed that the mCAFT alone predicted 64% of the variance in absolute VO_{2max} and that 74% of VO_{2max} could be predicted when adding grip strength and body weight to the equation.

Unfortunately, the psychometric and clinical properties of this test remain unexplored among patients with MS.

Another concern is that this test could be influenced by the strength of the lower limb and the general mobility capacity and, therefore, reflect these components more than the aerobic capacity. This should be explored by further studies.

Ruffier-Dickson Test

This test was developed in 1950 as an "index of resistance of the body to effort" [93] and is now widely used as a screening test in primary care and sports medicine. Given the nature of the test and its purpose, it could reasonably be classified as a tool evaluating an exercise tolerance function (b455 in the ICF classification). Unfortunately, its psychometric properties have not been studied in either healthy subjects or patients suffering from MS. The advantages of this test include its simple administration, calculation and interpretation procedures, its submaximal nature, as well as the low time burden and lack of needed materials. As for the mCAFT, lower limbs impairments could bias the capacity of the test to effectively measure aerobic capacity. However, its administration should not be recommended until its psychometric properties are explored among patients suffering from MS.

General discussion

Based on the evidence found, we proposed recommendations for the selection of a tool in various contexts. Figure 2 shows an algorithm for the selection of a tool to assess aerobic capacity among persons with MS.

One of the most important criteria to consider when assessing the quality of a measuring tool is its valid-

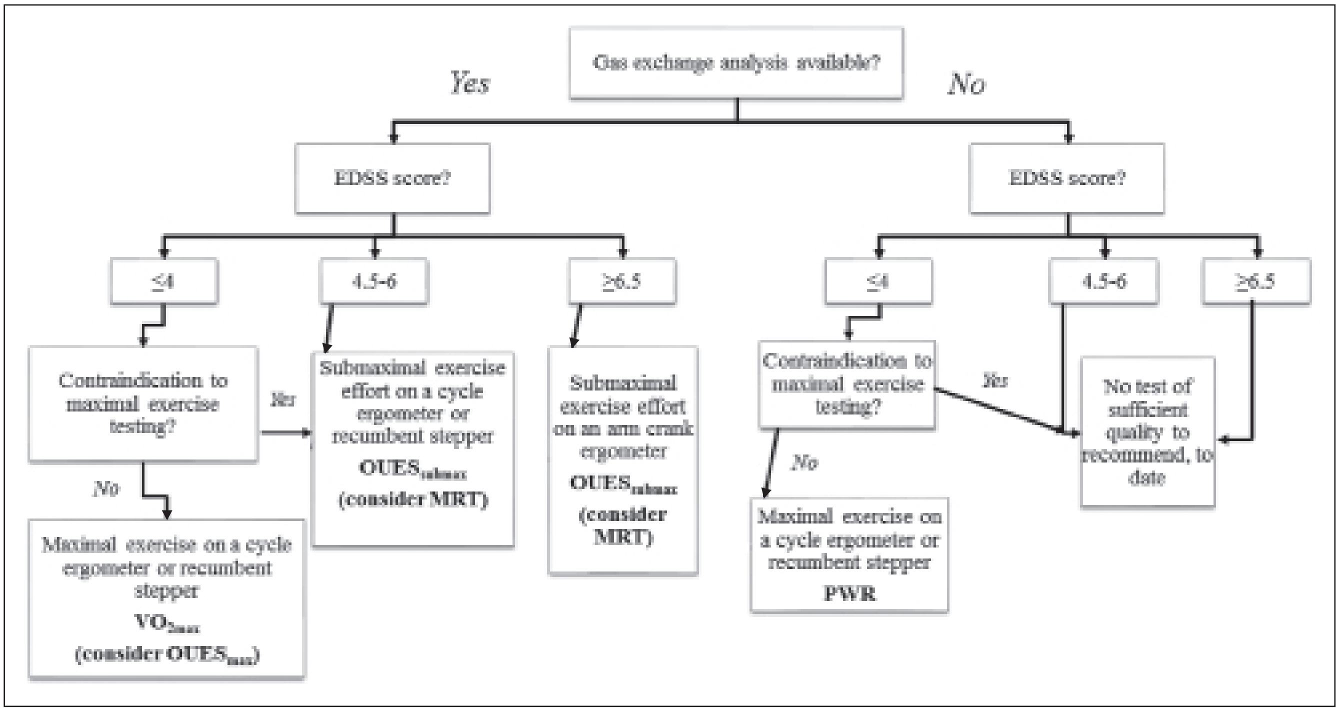

Figure 2.—Selecting a tool to assess exercise tolerance functions in persons with multiple sclerosis: a proposed algorithm.

ity. Direct measurement of VO_{2peak} by gas exchange analysis during progressive maximal exercise effort is considered the gold standard for assessing cardiorespiratory fitness.[67] Accordingly, this method has the highest content and criterion validity when the criteria for maximal exercise are attained. Alternative evaluation methods using a whole-body ergometer — maximum indirect (estimation by WR_{peak}) and direct submaximal ($OUES_{50}$) — show good correlations with direct assessment of VO_{2max}, and could be used when gas exchange analyzer are not available or when maximal effort is not reachable.[42, 43] Moreover, these tools seem to have the best content validity, as there seems to be a consensus among authors about the purpose of these tests (*i.e.*, to assess aerobic capacity). When available and reasonably practicable, tests including gas exchange analysis should always be preferred has they present the best psychometrics properties. Their reliability is excellent and a floor or ceiling effect is unlikely, allowing to precisely assess patients in a high range of capacity. Their responsiveness allows to detect any improvement or worsening of about 10%.

Maximal exercise effort is not always achievable in patients suffering from MS, especially when their EDSS score is higher than 4, because of a lack of strength and the presence of pain and fatigability. When gas exchange analysis is available, a submaximal test, such as the $OUES_{submax}$, is preferable for these patients. It could assess both patients with mild functional limitations using a maximal modality and those with moderate disability using a submaximal modality. The OUES could be an interesting option, but a standardized protocol is lacking.[42] The MRT could also be considered, but further studies are needed to explore its properties.

When gas exchange analysis is not feasible, the determination of WR_{peak} could be a useful option among patients able to reach maximal exercise effort (*i.e.* generally patients with an EDSS score below 4). An equation based on this metric, the age and the weight of the patients is available to accurately estimate the actual VO_{2peak}.[43] However, when maximal exercise effort is not attainable, no test have demonstrated sufficient psychometric properties to be used with confidence. 6MWT was demonstrated to measure walking capacity rather than aerobic capacity. The psychometric properties of simple submaximal tests, such as Ruffier-Dickson Test or $PWC_{75\%}$, should be determined by further studies in persons with MS.

It is important to note that all the methods described are designed to evaluate patients with MS with mild to moderate disability (EDSS 0-6.5). Only one of the 48 studies reviewed included patients with MS with more severe disability (EDSS 6.5-8) [52] and showed that arm crank ergometer testing with gas exchange analysis is a safe and acceptable modality for assessing exercise tolerance functions in these patients.

Limitations of the study

A limitation of this review is its qualitative nature. Thus, the studies are not stratified by type or quality, which varied widely between the included studies. As we reviewed the tests included in research reports only, other instruments used only in clinical practice could have been missed.

Conclusions

The tools intended to assess aerobic capacity analyzed in this review have different advantages and disadvantages for patients with MS. Walk tests have been well studied in MS, are conducted at submaximal effort (and are therefore acceptable for patients with a significant level of disability) and are easy to perform. However, they seem to assess overall mobility and walking performance rather than aerobic capacity in particular. Tests based on gas exchange analysis seem to be the most valid and reliable tools. Direct maximal exercise effort tools are considered the gold standard, but they can be conducted only under certain conditions. Indeed, while performing this type of test, persons suffering from MS are frequently limited by impaired lower limb strength, pain, motivational issues and/or central fatigue. VO_{2max} estimation by WR_{peak} has the same disadvantages (other than the need to use a gas exchange analyzer) and has not been widely used in patients with MS. OUES seems to be a promising tool, in that a well-designed protocol could directly assess physical condition using a maximal exercise effort test in patients able to reach this level and a submaximal test in others. This test is unfortunately almost as cumbersome to implement as direct measurement of VO_{2max}, and its psychometric properties have been under-studied in patients with MS. Thus, the use of tests with gas exchange analysis at maximal effort for patients suffering from MS with EDSS≤4 and at submaximal effort for patients with EDSS≥4.5 should be recommended to assess aerobic capacity, both in research and in clinical practice.

References

1. Compston A, Coles A. Multiple sclerosis. Lancet 2008;372:1502-17.
2. Ramagopalan SV, Sadovnick AD. Epidemiology of multiple sclerosis. Neurol Clin 2011;29:207-17.
3. Caspersen CJ, Powell KE, Christenson GM. Physical activity, exercise, and physical fitness: definitions and distinctions for health-related research. Public Health Rep 1985;100:126-31.
4. Guerra E, di Cagno A, Mancini P, Sperandii F, Quaranta F, Ciminelli E, *et al.* Physical fitness assessment in multiple sclerosis patients: a controlled study. Res Dev Disabil 2014;35:2527-33.
5. Langeskov-Christensen M, Heine M, Kwakkel G, Dalgas U. Aerobic capacity in persons with multiple sclerosis: a systematic review and meta-analysis. Sports Med 2015;45:905-23.
6. Motl RW, Goldman M. Physical inactivity, neurological disability, and cardiorespiratory fitness in multiple sclerosis. Acta Neurol Scand 2011;123:98-104.
7. Sandoval AE. Exercise in multiple sclerosis. Phys Med Rehabil Clin N Am 2013;24:605-18.
8. World Health Organization. International Classification of Functioning, Disability and Health: ICF. World Health Organization; 2001.
9. Conrad A, Coenen M, Kesselring J, Cieza A. What explains functioning from the perspective of people with multiple sclerosis? J Neurol 2014;261:2283-95.
10. Kluger BM, Krupp LB, Enoka RM. Fatigue and fatigability in neurologic illnesses: proposal for a unified taxonomy. Neurology 2013;80:409-16.
11. Beer S, Khan F, Kesselring J. Rehabilitation interventions in multiple sclerosis: an overview. J Neurol 2012;259:1994-2008.
12. Moher D. Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement. Annals of Internal Medicine 2009;151:264.
13. Andresen EM. Criteria for assessing the tools of disability outcomes research. Arch Phys Med Rehabil 2000;81:S15-20.
14. World Health Organization. International Classification of Functioning, Disability and Health (ICF); 2001 [Internet]. Available from: www.who.int/classifications/icf/en/ [cited 2017, Jun 15].
15. Terwee CB, Bot SD, de Boer MR, van der Windt DA, Knol DL, Dekker J, *et al.* Quality criteria were proposed for measurement properties of health status questionnaires. J Clin Epidemiol 2007;60:34-42.
16. McDowell I. Measuring Health: A Guide to Rating Scales and Questionnaires, Third Edition. New York: Oxford University Press; 2006.
17. Mukaka MM. Statistics corner: A guide to appropriate use of correlation coefficient in medical research. Malawi Med J 2012;24:69-71.
18. Terwee CB, Dekker FW, Wiersinga WM, Prummel MF, Bossuyt PM. On assessing responsiveness of health-related quality of life instruments: guidelines for instrument evaluation. Qual Life Res 2003;12:349-62.
19. Rehabilitaiton Measures Database. Basic review of statistical terms used on the Rehabilitation Measures website; 2015 [Internet]. Available from: http://www.rehabmeasures.org/rehabweb/rhstats.aspx [cited 2017, Jun 15].
20. Bansi J, Bloch W, Gamper U, Kesselring J. Training in MS: influence of two different endurance training protocols (aquatic versus overland) on cytokine and neurotrophin concentrations during three week randomized controlled trial. Mult Scler 2013;19:613-21.
21. Bjarnadottir OH, Konradsdottir AD, Reynisdottir K, Olafsson E.

Multiple sclerosis and brief moderate exercise. A randomised study. Mult Scler 2007;13:776-82.
22. Castellano V, White LJ. Serum brain-derived neurotrophic factor response to aerobic exercise in multiple sclerosis. J Neurol Sci 2008;269:85-91.
23. Feltham MG, Collett J, Izadi H, Wade DT, Morris MG, Meaney AJ, *et al.* Cardiovascular adaptation in people with multiple sclerosis following a twelve week exercise programme suggest deconditioning rather than autonomic dysfunction caused by the disease. Results from a randomized controlled trial. Eur J Phys Rehabil Med 2013;49:765-74.
24. Golzari Z, Shabkhiz F, Soudi S, Kordi MR, Hashemi SM. Combined exercise training reduces IFN-gamma and IL-17 levels in the plasma and the supernatant of peripheral blood mononuclear cells in women with multiple sclerosis. Int Immunopharmacol 2010;10:1415-9.
25. Tallner A, Streber R, Hentschke C, Morgott M, Geidl W, Maurer M, *et al.* Internet-Supported Physical Exercise Training for Persons with Multiple Sclerosis-A Randomised, Controlled Study. Int J Mol Sci 2016;17.
26. Heine M, Wens I, Langeskov-Christensen M, Verschuren O, Eijnde BO, Kwakkel G, *et al.* Cardiopulmonary fitness is related to disease severity in multiple sclerosis. Mult Scler 2016;22:231-8.
27. Sandroff BM, Balto JM, Klaren RE, Sommer SK, DeLuca J, Motl RW. Systematically developed pilot randomized controlled trial of exercise and cognition in persons with multiple sclerosis. Neurocase 2016;22:443-50.
28. Pariser G, Madras D, Weiss E. Outcomes of an aquatic exercise program including aerobic capacity, lactate threshold, and fatigue in two individuals with multiple sclerosis. J Neurol Phys Ther 2006;30:82-90.
29. Griffith G, Klaren RE, Motl RW, Baynard T, Fernhall B. Experimental protocol of a randomized controlled clinical trial investigating exercise, subclinical atherosclerosis, and walking mobility in persons with multiple sclerosis. Contemp Clin Trials 2015;41:280-6.
30. Briken S, Gold SM, Patra S, Vettorazzi E, Harbs D, Tallner A, *et al.* Effects of exercise on fitness and cognition in progressive MS: a randomized, controlled pilot trial. Mult Scler 2014;20:382-90.
31. Kuspinar A, Andersen RE, Teng SY, Asano M, Mayo NE. Predicting exercise capacity through submaximal fitness tests in persons with multiple sclerosis. Arch Phys Med Rehabil 2010;91:1410-7.
32. Langeskov-Christensen M, Langeskov-Christensen D, Overgaard K, Moller AB, Dalgas U. Validity and reliability of VO(2)-max measurements in persons with multiple sclerosis. J Neurol Sci 2014;342:79-87.
33. Larson RD, McCully KK, Larson DJ, Pryor WM, White LJ. Lower-limb performance disparities: implications for exercise prescription in multiple sclerosis. J Rehabil Res Dev 2014;51:1537-44.
34. Rampello A, Franceschini M, Piepoli M, Antenucci R, Lenti G, Olivieri D, *et al.* Effect of aerobic training on walking capacity and maximal exercise tolerance in patients with multiple sclerosis: a randomized crossover controlled study. Phys Ther 2007;87:545-55.
35. Klaren RE, Sandroff BM, Fernhall B, Motl RW. Comprehensive Profile of Cardiopulmonary Exercise Testing in Ambulatory Persons with Multiple Sclerosis. Sports Med 2016;46:1365-79.
36. Wens I, Dalgas U, Vandenabeele F, Grevendonk L, Verboven K, Hansen D, *et al.* High Intensity Exercise in Multiple Sclerosis: Effects on Muscle Contractile Characteristics and Exercise Capacity, a Randomised Controlled Trial. PLoS One 2015;10:e0133697.
37. Konečný L, Pospíšil P, Dufek M, Drlíková L, Anbais FH, Erajhi AA, *et al.* Functional impairment in multiple sclerosis. Scripta Medica Facultatis Medicae Universitatis Brunensis Masarykianae 2007;80:225-32.
38. Valet M, Stoquart G, Glibert Y, Hakizimana JC, Lejeune T. Is fatigue associated with cardiorespiratory endurance among patients suffering from multiple sclerosis? Ann Phys Rehabil Med 2016;59S:e41.
39. Rasova K, Brandejsky P, Havrdova E, Zalisova M, Rexova P. Spiroergometric and spirometric parameters in patients with multiple sclerosis: are there any links between these parameters and fatigue, depression, neurological impairment, disability, handicap and quality of life in multiple sclerosis? Mult Scler 2005;11:213-21.
40. Sandroff BM, Klaren RE, Motl RW. Relationships among physical inactivity, deconditioning, and walking impairment in persons with multiple sclerosis. J Neurol Phys Ther 2015;39:103-10.
41. Heine M, Hoogervorst EL, Hacking HG, Verschuren O, Kwakkel G. Validity of maximal exercise testing in people with multiple sclerosis and low to moderate levels of disability. Phys Ther 2014;94:1168-75.
42. Heine M, Verschuren O, Kwakkel G. Validity of Oxygen Uptake Efficiency Slope in patients with multiple sclerosis. J Rehabil Med 2014;46:656-61.
43. Motl RW, Fernhall B. Accurate prediction of cardiorespiratory fitness using cycle ergometry in minimally disabled persons with relapsing-remitting multiple sclerosis. Arch Phys Med Rehabil 2012;93:490-5.
44. Motl RW, Pilutti LA, Sandroff BM, Klaren R, Balantrapu S, McAuley E, *et al.* Rationale and design of a randomized controlled, clinical trial investigating a comprehensive exercise stimulus for improving mobility disability outcomes in persons with multiple sclerosis. Contemp Clin Trials 2013;35:151-8.
45. Heine M, van den Akker LE, Verschuren O, Visser-Meily A, Kwakkel G, Group T-AS. Reliability and responsiveness of cardiopulmonary exercise testing in fatigued persons with multiple sclerosis and low to mild disability. PLoS One 2015;10:e0122260.
46. Prakash RS, Snook EM, Motl RW, Kramer AF. Aerobic fitness is associated with gray matter volume and white matter integrity in multiple sclerosis. Brain Res 2010;1341:41-51.
47. Motl RW, Pilutti LA, Hubbard EA, Wetter NC, Sosnoff JJ, Sutton BP. Cardiorespiratory fitness and its association with thalamic, hippocampal, and basal ganglia volumes in multiple sclerosis. Neuroimage Clin 2015;7:661-6.
48. Sandroff BM, Pilutti LA, Benedict RH, Motl RW. Association between physical fitness and cognitive function in multiple sclerosis: does disability status matter? Neurorehabil Neural Repair 2015;29:214-23.
49. Pilutti LA, Motl RW, Edwards TA, Wilund KR. Rationale and design of a randomized controlled clinical trial of functional electrical stimulation cycling in persons with severe multiple sclerosis. Contemporary Clinical Trials Communications 2016;3:147-52.
50. Pilutti LA, Sandroff BM, Klaren RE, Learmonth YC, Platta ME, Hubbard EA, *et al.* Physical Fitness Assessment Across the Disability Spectrum in Persons With Multiple Sclerosis: A Comparison of Testing Modalities. J Neurol Phys Ther 2015;39:241-9.
51. Koseoglu BF, Gokkaya NK, Ergun U, Inan L, Yesiltepe E. Cardiopulmonary and metabolic functions, aerobic capacity, fatigue and quality of life in patients with multiple sclerosis. Acta Neurol Scand 2006;114:261-7.
52. Skjerbaek AG, Naesby M, Lutzen K, Moller AB, Jensen E, Lamers I, *et al.* Endurance training is feasible in severely disabled patients with progressive multiple sclerosis. Mult Scler 2014;20:627-30.
53. Schmidt S, Wonneberger M. Long-term endurance exercise improves aerobic capacity in patients with relapsing-remitting multiple sclerosis: impact of baseline fatigue. J Neurol Sci 2014;336:29-35.
54. Wonneberger M, Schmidt S. Changes of gait parameters following long-term aerobic endurance exercise in mildly disabled multiple sclerosis patients: an exploratory study. Eur J Phys Rehabil Med 2015;51:755-62.
55. Beier M, Bombardier CH, Hartoonian N, Motl RW, Kraft GH. Improved physical fitness correlates with improved cognition in multiple sclerosis. Arch Phys Med Rehabil 2014;95:1328-34.
56. Cakt BD, Nacir B, Genc H, Saracoglu M, Karagoz A, Erdem HR, *et al.* Cycling progressive resistance training for people with multiple sclerosis: a randomized controlled study. Am J Phys Med Rehabil 2010;89:446-57.
57. McCullagh R, Fitzgerald AP, Murphy RP, Cooke G. Long-term benefits of exercising on quality of life and fatigue in multiple sclerosis patients with mild disability: a pilot study. Clin Rehabil 2008;22:206-14.
58. Hansen D, Wens I, Kosten L, Verboven K, Eijnde BO. Slowed exercise-onset Vo2 kinetics during submaximal endurance exercise in subjects with multiple sclerosis. Neurorehabil Neural Repair 2013;27:87-95.

59. Hansen D, Wens I, Keytsman C, Eijnde BO, Dendale P. Is long-term exercise intervention effective to improve cardiac autonomic control during exercise in subjects with multiple sclerosis? A randomized controlled trial. Eur J Phys Rehabil Med 2015;51:223-31.
60. Wens I, Hansen D, Verboven K, Deckx N, Kosten L, Stevens AL, *et al.* Impact of 24 Weeks of Resistance and Endurance Exercise on Glucose Tolerance in Persons with Multiple Sclerosis. Am J Phys Med Rehabil 2015;94:838-47.
61. Wens I, Keytsman C, Deckx N, Cools N, Dalgas U, Eijnde BO. Brain derived neurotrophic factor in multiple sclerosis: effect of 24 weeks endurance and resistance training. Eur J Neurol 2016;23:1028-35.
62. Op 't Eijnde B, Keytsman C, Wens I, Hansen D. Whole-body cooling does not compromise muscle oxidative capacity in subjects with multiple sclerosis. NeuroRehabilitation 2014;35:805-11.
63. Paul L, Coote S, Crosbie J, Dixon D, Hale L, Holloway E, *et al.* Core outcome measures for exercise studies in people with multiple sclerosis: recommendations from a multidisciplinary consensus meeting. Mult Scler 2014;20:1641-50.
64. Coote S, Gallagher S, Msetfi R, Larkin A, Newell J, Motl RW, *et al.* A randomised controlled trial of an exercise plus behaviour change intervention in people with multiple sclerosis: the step it up study protocol. BMC Neurol 2014;14:241.
65. Pantea C, Sîrbu E, Pantea S. Effects of physical activity on functional and cardio respiratory capacity in multiple sclerosis. Timisoara Physical Education & Rehabilitation Journal 2011;4:29-34.
66. Borg G. Borg's Perceived Exertion and Pain Scales. Champaign, IL: Human Kinetics; 1998.
67. Medicine ACoS. ACSM's Guidelines for exercise testing and prescription. Baltimore: Lippincott Williams & Wilkins; 2013.
68. Kerling A, Keweloh K, Tegtbur U, Kuck M, Grams L, Horstmann H, *et al.* Effects of a Short Physical Exercise Intervention on Patients with Multiple Sclerosis (MS). Int J Mol Sci 2015;16:15761-75.
69. Buys R, Coeckelberghs E, Vanhees L, Cornelissen VA. The oxygen uptake efficiency slope in 1411 Caucasian healthy men and women aged 20-60 years: reference values. Eur J Prev Cardiol 2015;22:356-63.
70. A. T. S. Committee on Proficiency Standards for Clinical Pulmonary Function Laboratories. ATS statement: guidelines for the six-minute walk test. Am J Respir Crit Care Med 2002;166:111-7.
71. Enright PL, Sherrill DL. Reference equations for the six-minute walk in healthy adults. Am J Respir Crit Care Med 1998;158:1384-7.
72. Geddes EL, Costello E, Raivel K, Wilson R. The effects of a twelve-week home walking program on cardiovascular parameters and fatigue perception of individuals with multiple sclerosis: a pilot study. Cardiopulm Phys Ther J 2009;20:5-12.
73. Goldman MD, Marrie RA, Cohen JA. Evaluation of the six-minute walk in multiple sclerosis subjects and healthy controls. Mult Scler 2008;14:383-90.
74. Paltamaa J, Sjogren T, Peurala SH, Heinonen A. Effects of physiotherapy interventions on balance in multiple sclerosis: a systematic review and meta-analysis of randomized controlled trials. J Rehabil Med 2012;44:811-23.
75. Weller IM, Thomas SG, Corey PN, Cox MH. Prediction of maximal oxygen uptake from a modified Canadian aerobic fitness test. Can J Appl Physiol 1993;18:175-88.
76. Van Laethem C, De Sutter J, Peersman W, Calders P. Intratest reliability and test-retest reproducibility of the oxygen uptake efficiency slope in healthy participants. Eur J Cardiovasc Prev Rehabil 2009;16:493-8.
77. Batcho CS, Thonnard JL, Nielens H. PWC 75%/kg, a fitness index not linked to resting heart rate: testing procedure and reference values. Arch Phys Med Rehabil 2012;93:1196-200.
78. Wasserman K, Hansen JE, Sue DY. Principles of exercise testing and interpretation: including pathophysiology and clinical applications. Wolters Kluwer Health; 2011.
79. Tuner SL, Easton C, Wilson J, Byrne DS, Rogers P, Kilduff LP, *et al.* Cardiopulmonary responses to treadmill and cycle ergometry exercise in patients with peripheral vascular disease. J Vasc Surg 2008;47:123-30.
80. Kieseier BC, Pozzilli C. Assessing walking disability in multiple sclerosis. Mult Scler 2012;18:914-24.
81. Storer TW, Davis JA, Caiozzo VJ. Accurate prediction of VO2max in cycle ergometry. Med Sci Sports Exerc 1990;22:704-12.
82. Muller Pde T, Christofoletti G, Zagatto AM, Paulin FV, Neder JA. Reliability of peak O2 uptake and O2 uptake kinetics in step exercise tests in healthy subjects. Respir Physiol Neurobiol 2015;207:7-13.
83. Astrand PO, Ryhming I. A nomogram for calculation of aerobic capacity (physical fitness) from pulse rate during sub-maximal work. J Appl Physiol 1954;7:218-21.
84. Nielens H, Plaghki L. Perception of pain and exertion during exercise on a cycle ergometer in chronic pain patients. Clin J Pain 1994;10:204-9.
85. Bohannon RW, Crouch R. Minimal clinically important difference for change in 6-minute walk test distance of adults with pathology: a systematic review. J Eval Clin Pract 2017;23:377-81.
86. Savci S, Inal-Ince D, Arikan H, Guclu-Gunduz A, Cetisli-Korkmaz N, Armutlu K, *et al.* Six-minute walk distance as a measure of functional exercise capacity in multiple sclerosis. Disability and Rehabilitation 2009;27:1365-71.
87. Sandroff BM, Pilutti LA, Motl RW. Does the six-minute walk test measure walking performance or physical fitness in persons with multiple sclerosis? NeuroRehabilitation 2015;37:149-55.
88. Burr JF, Bredin SS, Faktor MD, Warburton DE. The 6-minute walk test as a predictor of objectively measured aerobic fitness in healthy working-aged adults. Phys Sportsmed 2011;39:133-9.
89. Wu G, Sanderson B, Bittner V. The 6-minute walk test: how important is the learning effect? Am Heart J 2003;146:129-33.
90. Pankoff BA, Overend TJ, Lucy SD, White KP. Reliability of the six-minute walk test in people with fibromyalgia. Arthritis Care Res 2000;13:291-5.
91. Perera S, Mody SH, Woodman RC, Studenski SA. Meaningful change and responsiveness in common physical performance measures in older adults. J Am Geriatr Soc 2006;54:743-9.
92. Shephard RJ, Bailey DA, Mirwald RL. Development of the Canadian Home Fitness Test. Can Med Assoc J 1976;114:675-9.
93. Dickson J. L'utilisation de l'indice cardiaque de Ruffier dans le contrôle médic sportif. Med Educ Phys Sport 1950;2:65.

Funding.—This review is included in a clinical project funded by the Fondation pour l'Aide à la Recherche sur la Sclérose en Plaques (ARSEP) and by the Fondation Saint-Luc.

Conflicts of interest.—The authors certify that there is no conflict of interest with any financial organization regarding the material discussed in the manuscript.

Article first published online: May 31, 2017. - Manuscript accepted: May 24, 2017. - Manuscript revised: April 13, 2017. - Manuscript received: January 27, 2016.