Review

Measuring Treatment Response in Progressive Multiple Sclerosis—Considerations for Adapting to an Era of Multiple Treatment Options

Nik Krajnc [1,2], **Thomas Berger** [1] **and Gabriel Bsteh** [1,*]

1   Department of Neurology, Medical University of Vienna, 1090 Vienna, Austria;
   nik.krajnc@meduniwien.ac.at (N.K.); thomas.berger@medunwien.ac.at (T.B.)
2   Department of Neurology, University Medical Centre Ljubljana, 1000 Ljubljana, Slovenia
*   Correspondence: gabriel.bsteh@meduniwien.ac.at

Abstract: Disability in multiple sclerosis accrues predominantly in the progressive forms of the disease. While disease-modifying treatment of relapsing MS has drastically evolved over the last quarter-century, the development of efficient drugs for preventing or at least delaying disability in progressive MS has proven more challenging. In that way, many drugs (especially disease-modifying treatments) have been researched in the aspect of delaying disability progression in patients with a progressive course of the disease. While there are some disease-modifying treatments approved for progressive multiple sclerosis, their effect is moderate and limited mostly to patients with clinical and/or radiological signs of disease activity. Several phase III trials have used different primary outcomes with different time frames to define disease progression and to evaluate the efficacy of a disease-modifying treatment. The lack of sufficiently sensitive outcome measures could be a possible explanation for the negative clinical trials in progressive multiple sclerosis. On the other hand, even with a potential outcome measure that would be sensitive enough to determine disease progression and, thus, the efficacy or failure of a disease-modifying treatment, the question of clinical relevance remains unanswered. In this systematic review, we analyzed outcome measures and definitions of disease progression in phase III clinical trials in primary and secondary progressive multiple sclerosis. We discuss advantages and disadvantages of clinical and paraclinical outcome measures aiming for practical ways of combining them to detect disability progression more sensitively both in future clinical trials and current clinical routine.

Keywords: multiple sclerosis; treatment response; clinical trial; outcome measure; disability progression

1. Introduction

Multiple sclerosis (MS) is a chronic immune-mediated disease of the central nervous system, pathophysiologically characterized by inflammation, as well as neurodegeneration occurring in even early stages of the disease [1]. While 10–15% of patients display a progressive disease course from onset, called primary progressive MS (PPMS), natural history studies have indicated that about two-thirds of untreated patients with an initially relapsing course develop a progressive disease course, called secondary progressive MS (SPMS), after a mean of 15–20 years with broad interindividual variance [2–4]. Although advances in the treatment of relapsing MS have resulted in both a decrease and a delay of progressive MS (PMS), the degree of disability accrual is defined by PMS and the conversion to SPMS remains the single most important predictor of long-term outcome [5–10].

Encouragingly, efforts in recent years have yielded the first effective treatment options for PMS, albeit currently not as effective as in RMS [11,12]. Both the continuing quest for finding effective treatments for PMS and the growing challenge for navigating and sequencing treatment options require reliable methods for assessing response or failure of a given treatment.

As we currently still lack treatments with the potential to repair or restore neuroaxonal damage, the basic principle of treatment response remains as prevention of acute or chronic inflammation-driven deterioration, i.e., disease or disability progression dependent or independent from clinical relapses. To establish the prevention of deterioration in a valid way requires testing instruments sensitive enough to detect deterioration but also specific and reliable enough to distinguish true deterioration from performance fluctuations and measurement-inherent errors. Equally important, outcome measures need to be able to answer whether the detected changes are clinically relevant, both at the moment of assessment and in terms of predicting future changes.

Anticipating the availability of multiple treatment options for PMS in the not-too-distant future, outcome measures will also be required for evidence-based decision making to determine when to change or add DMT in PMS. In this regard, it is also to be defined what is the minimal or maximal time frame required to measure the efficacy or failure of DMT.

Both successful and unsuccessful phase III clinical trials using different outcome measures with different time frames to define disease progression when evaluating the efficacy of DMT agents in PMS provide instructive data to compare and assess outcome measures. In that way, we analyzed the phase III clinical trials dealing with PMS with a special focus on the applied outcome measures and the definition of disability progression. Further on, we open the question of the frequency of patients' evaluation in order to confirm the progression's sustainability. Finally, we present our view on what should be taken into account when measuring DMT response both when planning future trials and treating patients with PMS.

2. Methods

2.1. Search Methods

A review of the literature concerning phase III clinical trials in PMS was performed using PubMed. Search terms included the following: progressive, multiple sclerosis, disease-modifying treatment AND phase III trial. The last search was performed on the 30th of June 2021. This search yielded 63 results.

2.2. Selection Criteria

Our selection criteria were language (English) and article type (clinical trial). After selection criteria were applied, 48 articles were excluded. The results were hand-searched to exclude additional nonrelevant publications. We have chosen to include only clinical trials published after 1993. The systematic review was prepared according to the latest PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) guidelines (see Figure 1).
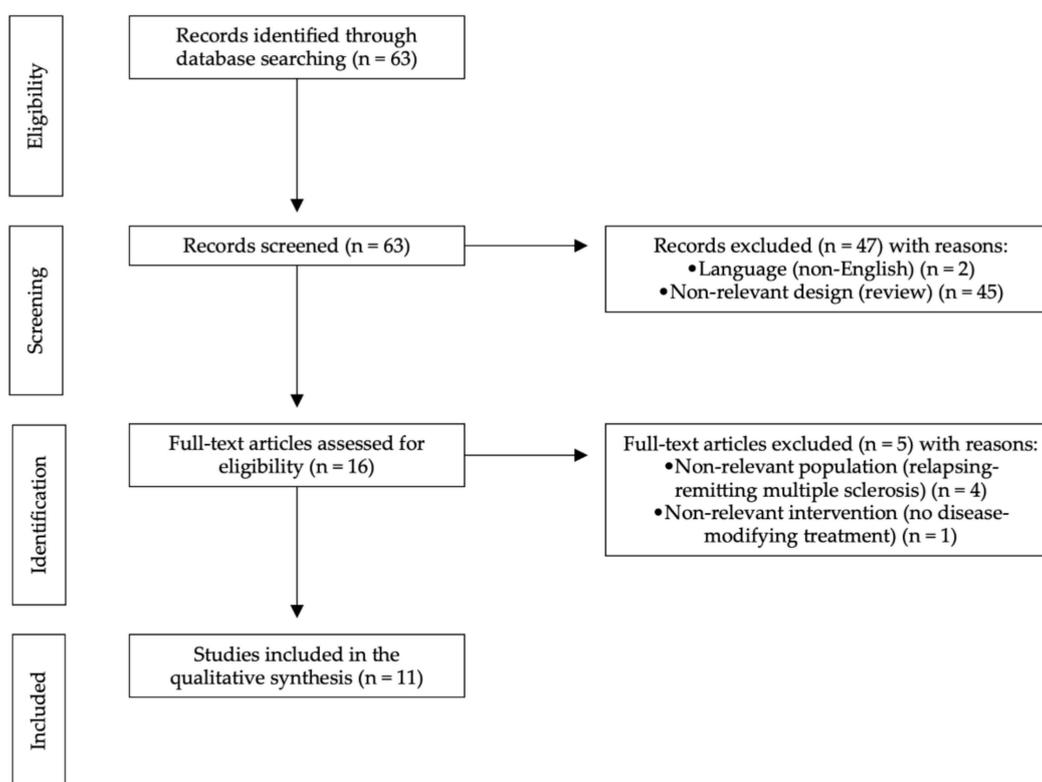
Figure 1. PRISMA flow diagram demonstrating included and excluded papers, and the reasons for exclusion in this systematic review.

3. Results

We identified 12 outcome measures from 11 trials dealing with the efficacy of DMT in both PPMS and SPMS. We present their characteristics and designs in Table 1.

Table 1. Phase III clinical trials and their designs in progressive multiple sclerosis.

Study Name	Study Design	Inclusion Criteria	Intervention and Comparator	Study Duration	Primary Outcome Measure	Definition of Progression	Sustained Progression	Secondary Outcome Measures
						Primary progressive multiple sclerosis		
ORATORIO	Randomized, double-blind, placebo-controlled	Age 18–55 years EDSS 3.0–6.5 with pyramidal FS ≥ 2 Duration of MS symptoms of <15 years (EDSS > 5.0) or <10 years (EDSS ≤ 5.0) OCB in the CSF	Ocrelizumab (n = 488) vs. placebo (n = 244)	120 weeks	EDSS	Increase in EDSS score for ≥1.0 point (EDSS ≤ 5.5) or ≥0.5 point (EDSS ≥ 6.0)	12 weeks	24 week CDP, change in T25FW from baseline to week 120 MRI: change in the total volume of brain lesions (T2-weighted) at week 120, change in whole brain volume from week 24 to week 120 PROMs: change in the SF-36 Physical component
INFORMS	Randomized, double-blind, placebo-controlled	Age 25–65 years EDSS 3.5–6.0 with pyramidal FS ≥ 2 Increase in EDSS for ≥0.5 point in the past 2 years T25FW < 30 s Disease duration 2–10 years	Fingolimod (n = 336) vs. placebo (n = 487)	36 months	EDSS-Plus	Increase in EDSS score for ≥1.0 point (EDSS ≤ 5.0) or ≥0.5 point (EDSS ≥ 5.5) or increase of ≥20% in the T25FW or 9HPT	3 months	Time to 3 month CDP MRI: brain volume change, number of Gd-positive lesions, number of T2 lesions PROMs: PRIMUS, EQ-5D, UFIS, MSWS-12
OLYMPUS	Randomized, double-blind, placebo-controlled	Age 18–65 years EDSS 2.0–6.5 with pyramidal FS ≥ 2 OCB in CSF Disease duration ≥1 year	Rituximab (n = 292) vs. placebo (n = 147)	96 weeks	EDSS	Increase in EDSS score for ≥1.0 point (EDSS ≤ 5.5) or ≥0.5 point (EDSS ≥ 6.0)	12 weeks	Time to 24 week CDP, change in MSFC Pharmacodynamic assessments of B-cells counts and immunoglobulin levels MRI: number of Gd-enhancing lesions, change in volume of T2 lesions, change in brain volume

Table 1. *Cont.*

Study Name	Study Design	Inclusion Criteria	Intervention and Comparator	Study Duration	Primary Outcome Measure	Definition of Progression	Sustained Progression	Secondary Outcome Measures
PROMiSE	Randomized, double-blind, placebo-controlled	Age 30–65 years EDSS 3.0–6.5 with pyramidal FS $\geq$ 2 Evidence of myelopathy for at least 6 months before the screening	Glatiramer acetate (n = 627) vs. placebo (n = 316)	36 months	EDSS	Increase in EDSS score for $\geq$1.0 point (EDSS $\leq$ 5.0) or $\geq$0.5 point (EDSS $\geq$ 5.5)	3 months	Proportion of progression-free patients, change in EDSS and MSFC MRI: number and volume of T2 lesions, number of Gd-enhancing lesions, brain volume loss
Secondary progressive multiple sclerosis								
SPECTRIMS	Randomized, double-blind, placebo-controlled	Age 18–55 years EDSS 3.0–6.5 $\geq$2 relapses or increase in EDSS for $\geq$1.0 point in the previous 2 years	Interferon β 1a (n = 360) vs. placebo (n = 358)	36 months	EDSS	Increase in EDSS score for $\geq$1.0 point (EDSS $\leq$ 5.5) or $\geq$0.5 point (EDSS $\geq$ 6.0)	3 months	Time to becoming wheelchair-bound (EDSS $\geq$ 7.0), ARR, time to first relapse, proportion of patients with moderate/severe relapses MRI: volume of T2 lesions, number of newly active lesions, change in brain volume
Interferon β 1a study	Randomized, double-blind, placebo-controlled	Age 18–55 years EDSS 1.0–3.5 $\geq$2 relapses in the prior 3 years No relapse up to 2 months before randomization Disease duration $\geq$1 year	Interferon β 1a (n = 158) vs. placebo (n = 143)	104 weeks	EDSS	Increase in EDSS score for $\geq$1.0 point	6 months	Number of relapses MRI: number and volume of Gd-enhancing lesions, T2 lesion volume
IMPACT	Randomized, double-blind, placebo-controlled	Age 18–60 years EDSS 3.5–6.5	Interferon β 1a (n = 217) vs. placebo (n = 219)	24 months	MSFC	Decrease in the mean and median MSFC Z-scores	3 months	Change in EDSS, ARR MRI: number of new or enlarging T2 lesions and number of Gd-enhancing lesions PROMs: BDI, MSQLI

Table 1. *Cont.*

Study Name	Study Design	Inclusion Criteria	Intervention and Comparator	Study Duration	Primary Outcome Measure	Definition of Progression	Sustained Progression	Secondary Outcome Measures
Cladribine study	Randomized, double-blind, placebo-controlled	Clinically definite secondary progressive MS for ≥2 years	Cladribine (n = 24) vs. placebo (n = 24)	2 years	EDSS, SNRS	The change in both measurements was calculated every month[1]		MRI: number and volume of T2-lesions CSF: total protein and immunoglobulin concentration, oligoclonal bands
MIMS	Randomized, double-blind, placebo-controlled	Age 18–55 years EDSS 3.0–6.0 Increase in EDSS for ≥1.0 point during 18 months before enrolment (progressive relapsing-remitting or secondary progressive MS) No relapse up to 8 weeks before randomization	Mitoxantrone (n = 124) vs. placebo (n = 64)	24 months	Composite score[2]	Only a change between baseline and 24 months in each measurement of the composite score was calculated		Proportion of patients with deterioration of ≥1.0 EDSS point, proportion of patients with such deterioration confirmed after 3 and 6 months, time to first EDSS deterioration, time to first relapse, ARR, proportion of patients without relapse, number of days in hospital, use of wheelchair assistance MRI: number and volume of Gd-enhancing, and nonenhancing T1 and T2 lesions PROMs: quality of life (Stanford health assessment questionnaire)

Table 1. *Cont.*

Study Name	Study Design	Inclusion Criteria	Intervention and Comparator	Study Duration	Primary Outcome Measure	Definition of Progression	Sustained Progression	Secondary Outcome Measures
ASCEND	Randomized, double-blind, placebo-controlled	Age 18–58 years EDSS 3.0–6.5 with disability progression during the year before enrolment MSSS ≥ 4 No relapse up to 3 months before randomization	Natalizumab (n = 439) vs. placebo (n = 448)	96 weeks	EDSS-Plus	Increase in EDSS score for ≥1.0 point (EDSS ≤ 5.5) or ≥0.5 point (EDSS ≥ 6.0) or increase of ≥20% in the T25FW or 9HPT	3 months	Proportion of patients with improvement in T25FW, proportion of patients with disability progression, measured by individual physical EDSS FS scores MRI: change in whole brain volume between week 24 and 96 PROMs: change in patient-reported ambulatory status (MSWS-12), manual ability (ABILHAND questionnaire), quality of life (MSIS-29 physical score)[3]
EXPAND	Randomized, double-blind, placebo-controlled	Age 18–60 years EDSS 3.0–6.5 progression during 2 years before enrolment No relapse up to 3 months before randomization	Siponimod (n = 1105) vs. placebo (n = 546)	Protocol amended after occurrence of 374 CDP events and after ≥95% of patients had been assigned to treatment for ≥12 months	EDSS	Increase in EDSS score for ≥1.0 points (EDSS ≤ 5.0) or increase of ≥0.5 (EDSS ≥ 5.5)	3 months	Time to 3 and 6 month confirmed worsening of T25FW, ARR, time to first relapse, proportion of relapse-free patients MRI: number of new or enlarging T2 lesions, change in T2 lesion volume, number of Gd-enhancing lesions, brain volume change PROMs: change in score on the MSWS-12

[1] Meaningful change was defined as an increase in EDSS for ≥1.0 point or decrease in SNRS for ≥10 points after one year. [2] Composite score of EDSS, ambulation index, number of treated relapses, time to first treated relapse, and standardized neurological status. [3] Extended trial included additional secondary endpoints: proportion of patients with disability progression on the multicomponent endpoint during additional follow-up time, change in EDSS, T25FW, and 9HPT from baseline to week 156, change in whole brain and grey matter volume, change in number of T2 lesions, change in 6-Minute Walk Test, MSIS-29 Physical, and SDMT, and change in Work Productivity and Activity Impairment Questionnaire. ARR: annualized relapse rate, BDI: Beck Depression Inventory, CDP: confirmed disease progression, CSF: cerebrospinal fluid, EDSS: Expanded Disability Status Scale, EQ-5D: EuroQoL, MRI: magnetic resonance imaging, MS: multiple sclerosis, MSFC: Multiple Sclerosis Functional Composite, MSIS-29: Multiple Sclerosis Impact Scale-29, MSWS-12: 12-item Multiple Sclerosis Walking Scale, OCB: oligoclonal bands, PRIMUS: Patient-Reported Indices in Multiple Sclerosis, PROMs: patient-reported outcome measures, QoL: quality of life, SDMT: Symbol Digit Modalities Test, SF-36: Medical Outcomes Study 36-Item Short-Form Health Survey, SNRS: Scripps neurologic rating score, T25FW: Timed 25-Foot Walk, UFIS: Unidimensional Fatigue Impact Scale, 9HPT: 9-Hole Peg Test.

3.1. Expanded Disability Status Scale

The Expanded Disability Status Scale (EDSS) is the mainstay primary outcome measure of disability progression in MS. Developed in 1983, it is widely used in both studies and clinical routine and is, thus, widely familiar to MS clinicians [13]. Based on a standard neurological examination, MS-related impairment is quantified on a 0–10 scale emphasizing walking ability (EDSS score 4.0–6.0), or other functional impairments (EDSS score ≥ 6.5) [13].

There is some variability in the definition of disability progression: ORATORIO, OLYMPUS, and SPECTRIMS adopted the now-widely recommended definition of an increase of ≥ 1.0 points from baseline scores ≤ 5.5 or ≥ 0.5 points (EDSS ≥ 6.0), whereas the PROMiSE and EXPAND trials used the EDSS ≥ 5.5 as the threshold for a 0.5-point step change [11,12,14–16]. As EDSS 6.0 defines the need for an assistive device, it seems to be a more reasonable cut-off value for a 0.5-point change to detect disability progression. The interferon β 1b trial defined disability progression as an increase of ≥ 1.0 point regardless of the baseline EDSS, but the trial only enrolled patients with a lower level of disability (EDSS 1.0–3.5) [17].

The minimum time period of sustainment required in order to meet the criteria for disability progression also varies considerably from 12 weeks (ORATORIO and OLYMPUS), to 3 months (INFORMS, PROMiSE, and EXPAND) or 6 months (interferon β 1b study) [11,12,14,16–18]. Of note, worsening needs to be confirmed in the same functional score in which it was detected at baseline in order to maintain the validity of the EDSS [19]. In that way, the necessary time frame to determine disease progression with EDSS can be inferred as ≥ 3 months.

In the MIMS study, however, only the change in EDSS between baseline and month 24 was evaluated [20]. The use of a roving EDSS reference value, in which the increase or decrease in EDSS has to be separated from the last EDSS assessment by at least 24 or 48 weeks, allows the more sensitive measurement of disability progression in comparison to the conventional fixed study baseline EDSS but decreases specificity [21,22].

Although EDSS is easy to administer, well-known, and includes most functional systems that might be affected in MS, it also has some limitations [23]. Due to its nonlinear nature, an increase of 1.0 point has a different significance for each level [24,25]. EDSS also lacks sensitivity for detecting a change in clinical status in patients with a progressive course, especially upper-extremity and cognitive dysfunction, as it is heavily weighted toward ambulation at higher degrees of disability [26–29]. Moreover, it also lacks inter- and intra-rater reliability, the latter being more variable with lower EDSS scores [30–33].

3.2. Timed 25-Foot Walk Test

The Timed 25-Foot Walk Test (T25FW) evaluates the time required to walk the distance of 25 feet (7.62 m) in the patient's usual manner, including potential use of assistive devices [34]. The test is performed twice, and the average of both trials is scored. T25FW is reliable over short and long periods of time, has a high degree of validity, and responds to a clinically meaningful change in disability progression [34–38].

The INFORMS and ASCEND clinical trials used the T25FW as one of the primary outcome measures [18,39]. An increase of $\geq 20\%$ in the T25FW was defined as clinically significant [35,40–44]. The necessary time frame to determine disease progression with T25FW is ≥ 3 months.

T25FW has advantages as it is more sensitive to detect disease progression compared to EDSS, can be easily administered, and is inexpensive and applicable among a wide range of patients with different disability levels [45,46]. However, with patients reaching wheelchair dependence (EDSS ≥ 7.0), it becomes less useful in determining disease progression (ceiling effect), and it also shows some flooring effects, which, however, seems to be less important in patients with PMS [47].

3.3. 9-Hole Peg Test

The 9-Hole Peg Test (9HPT) evaluates manual dexterity by measuring the time in which a patient moves 9 pegs one by one into 9 holes on a board and back into the box [48]. The test is performed twice with each hand, and the time is averaged for each side separately [49]. An alternative scoring with the number of pegs placed in 50 or 100 s can be recorded; in this case, results are expressed as the number of pegs placed per second [50,51].

The INFORMS and ASCEND clinical trials used the 9HPT as one of the primary outcome measures [18,39]. An increase of ≥20% in the 9HPT was defined as clinically significant as described before [42,43]. However, the minimum 9HPT score at baseline should be at least 21 s in order to detect such a change [52]. The necessary time frame to determine disease progression with 9HPT is ≥3 months.

The advantages of 9HPT are that it is easily accessible, requires no training, and evaluates measures that are not included in EDSS. Besides, it shows excellent test–retest reliability, as well as intra- and interrater reliability [53]. However, 9HPT demonstrates a practice effect with each repetition, as well as floor and ceiling effect [52]. The former can be avoided with the calculation of pegs per second in people with severe upper limb dysfunction [54].

3.4. Paced Auditory Serial Addition Test

The Paced Auditory Serial Addition Test (PASAT) has been used since 1974 as a common research tool for different neurological disorders, including MS [55]. It assesses the auditory information processing capacity by the presentation of 60 numbers every 3 s to the patient who is asked to add each new digit to the one immediately prior to it [56]. The scores most commonly reported for the PASAT are the number of correct responses for each trial (maximum 60 points) or the total number of correct responses over all trials [55].

Only the IMPACT study used PASAT as one of the primary outcome measures, where a decrease of ≥20% in the PASAT score was defined as clinically significant [57]. The necessary time to determine disease progression with PASAT is at least 3 months.

PASAT has some advantages as it evaluates measures not included in EDSS and shows a high level of test–retest reliability [55]. However, it demonstrates a practice effect, which seems to be independent of the duration of the test–retest interval. Besides, it is perceived as a relatively difficult task even for individuals with high intellectual capability and can be interpreted differently (number of correct responses, the average correct response time, etc.) [58].

3.5. Symbol Digit Modalities Test

The Symbol Digit Modalities Test (SDMT) is a sentinel test for cognitive impairment in patients with MS (60% specificity and 91% sensitivity) [59,60]. Using a reference key, the test taker has 90 s to pair specific numbers with given geometric figures. The task evaluates sustained attention, capacity of concentration, and visuomotor speed [61–63].

SDMT was used in the EXPAND clinical trial in which a ≥4-point or ≥10% change in SDMT was considered clinically relevant [64,65]. The necessary time frame to determine disease progression with SDMT is ≥3 months.

SDMT is time-efficient, easy to administer, change-sensitive, and independent of language. It also correlates less strongly with EDSS and the other performance measures (T25FW and 9HPT), providing additional information by assessing functions not captured by the other measures [47]. It shows superior reliability, sensitivity, and greater patient acceptability compared to PASAT and, hence, SDMT is recommended for use instead of PASAT [64–66]. However, its normative values are affected by several factors such as age, sex, education, and cultural background [67]. Besides, it might demonstrate a slight practice effect, which is why a change in the key sequence is recommended for each repetition [68].

3.6. Low-Contrast Letter Acuity

The Low-Contract Letter Acuity (LCLA) chart is based on Early Treatment Diabetic Retinopathy Study High-Contrast Letter Acuity (HCLA) charts, but uses grey letters with 2.5% and 1.25% contrast levels as opposed to black letters (100% contrast level) [69]. Patients are asked to read letters from a 2 m distance proceeding top to bottom and from left to right. The score for each chart is quantified as the number of letters identified correctly with a maximum score of 70 letters [70]. LCLA shows good structure-function-correlation both with retinal atrophy and lesions in the posterior visual pathway [71–73].

The LCLA chart was first used as an exploratory outcome measure in the IMPACT clinical trial, where the clinically meaningful change was determined at a ≥ 7 letters loss, which is beyond the threshold of test–retest variability [74–77]. The necessary time frame to determine disease progression with the LCLA chart seems to be ≥ 3 months.

LCLA is easy to administer and interpret, provides a linear scale for statistical analysis, and shows excellent interrater reliability [70]. It correlates less strongly with EDSS and other performance measures, supporting its additive value to assess functions that are not captured by the latter [47]. Still, some correlation between LCLA and cognitive performance, measured by SDMT, has been described lately [78]. LCLA charts have some limitations such as a potential floor effect with lower contrast levels and a ceiling effect with higher contrast levels as most letters on the chart are easily readable with no potential for improvement over time [79]. Besides, the conditions in which the test is performed should be well-controlled (no surface reflections on the charts, even backlighting in a retro-illuminated cabinet) in order to receive comparable data.

3.7. Scripps Neurologic Rating Scale

The Scripps Neurologic Rating Scale (SNRS) is a summary measure of individual components comprising a neurological examination, designed for use in MS [80]. The assignment of points in the SNRS reflects the examiner's clinical assessment of each component in the neurological examination; points are then summed to arrive at the overall SNRS score. The responsiveness of most of the components seems to be in the small to moderate range; however, some components, such as the bladder, bowel, or sexual dysfunction, are insensitive to change [81]. Still, SNRS may be more sensitive than EDSS alone.

Only the Cladribine Study included SNRS as one of two primary outcome measures (besides EDSS) [82]. Although a change in SNRS was noted every month, additional analysis showed that a decrease in SNRS for ≥ 10 points after one year was meaningful in the sense of progression [82]. In that way, the necessary time frame to determine disease progression with SNRS seems to be ≥ 12 months.

Although SNRS seems to be more sensitive than EDSS, it is seldomly used and may not bring much additional value when determining disease progression. Besides, it seems that the necessary time frame to determine disease progression is considerably longer than in other measures.

3.8. Magnetic Resonance Imaging

In all mentioned clinical trials, imaging outcome measures have been used as secondary outcome measures, mostly the number of new/enlarging and/or gadolinium-enhancing lesions, with only a few of them including brain volume change as well. Grey matter atrophy occurs in all phenotypes of MS and is associated with disability accumulation [83]. Recently, cut-offs to distinguish pathological brain atrophy related to MS from the physiological change have been established, with 0.40% per year performing best for detecting physical disability progression (65% sensitivity and 80% specificity) [84]. Cortical atrophy seems to accelerate in PMS compared to relapsing-remitting MS (−0.87% vs. −0.48%, respectively) [85]. Although measurement techniques allow good reproducibility and sensitivity to changes, there are several limitations that need to be taken into account (e.g., technical limitations and dependence on confounding factors) [86–90]. Besides, pa-

tients treated with DMT have a slight decrease in the brain volume in the first 6–12 months (pseudoatrophy) due to the resolution of inflammation and its associated edema [91]. A possible solution would be to perform a re-baseline MRI 3–6 months after DMT initiation, although this would lengthen the time of observation, as a necessary time frame between two MRIs seems to be ≥ 1 year.

On the other hand, none of the clinical trials used spinal cord atrophy as an outcome measure, although it is considered a major driver of disability accumulation and is increasingly recognized as an additional imaging outcome measure besides brain atrophy [92]. The measurement of the cross-sectional area of the spinal cord presents a sensitive biomarker, especially as the estimated annual rate of spinal cord atrophy is greater when compared to the rate of brain atrophy in patients with MS (-1.78% vs. -0.5%) [93,94]. However, it remains to be elucidated what constitutes a relevant change in spinal cord atrophy as opposed to background noise or natural change. There are, however, some limitations that need to be considered. The spinal cord is a relatively small structure, which makes it difficult to obtain high reproducibility and responsiveness to changes, especially in a multi-center setting with variability in imaging protocols and scanners. As for brain MRI, a necessary time frame to determine disease progression is ≥ 1 year.

3.9. Composite Outcome Measures

The Multiple Sclerosis Functional Composite (MSFC) was developed in 1999 as an alternative approach to address the perceived shortcomings of the EDSS [48,95]. It contains three of the most important domains of the disease, including lower (T25FW) and upper extremity function (9HPT), and cognition (PASAT) [96]. To combine the results, Z-scores are calculated from each test and averaged to form a composite score [48].

The validity of MSFC as a clinical outcome measure is supported by several independent studies, including the IMPACT study. In this study, decreases in mean and median MSFC Z-scores were marked as progression [74]. The threshold for each of the measurements to detect clinically meaningful disability progression was reported at $\geq 20\%$ [42,43,97]. The necessary time frame to determine disease progression with MSFC is ≥ 3 months.

MSFC is a more reliable tool than EDSS with the ability to be conducted quickly by a trained technician. Besides, the Z-score is a continuous and unitless variable, facilitating longitudinal analysis over time [98]. However, the use of Z-scores opens the question of its relation to the clinically meaningful change in function. The lack of a vision measure seems to be a weakness of the original MSFC, and its components may suffer from ceiling and floor effects depending on the level of disability. Still, MSFC seems to be more sensitive than EDSS in detecting disability progression [99].

As PASAT has not been shown to sensitively detect cognitive progression in PMS, a composite measure of disability progression, incorporating the EDSS, T25FW, and 9HPT, has been considered (EDSS-Plus) [37,97,100]. It should be differentiated from MSFC, which uses analytical methods to create a single score from different scales and is, in that way, difficult to be appropriately weighted, whereas a multicomponent endpoint combines different outcomes and analyzes them on their own.

The EDSS-Plus was used in the ASCEND and INFORMS clinical trial [18,39]. In both, disability progression was defined as an increase of ≥ 1.0 points (EDSS ≤ 5.5) or ≥ 0.5 points (EDSS ≥ 6.0), or an increase of $\geq 20\%$ in the T25FW or 9HPT [18,39]. The necessary time frame to determine disease progression with EDSS-Plus is ≥ 3 months.

The EDSS-Plus is a more sensitive tool than EDSS alone, identifying 60% of patients as progressors, compared with 25%, 42%, and 34%, when using the individual tests alone (EDSS, T25FW, and 9HPT, respectively) [97]. This was confirmed in the INFORMS study, where the 3 month CDP for EDSS, T25FW, and 9HPT was 54%, 63%, and 34%, respectively, with a composite score confirming progression in 77% of patients [18]. Similar conclusions were drawn from the post hoc analysis of the ORATORIO clinical trial [101]. Apart from greater sensitivity, and assessing short-distance ambulation and upper-extremity

function, both T25FW and 9HPT are continuous measures compared to EDSS, which has an ordinal scale.

Composite scores have several strengths by including the combination of different assessments of MS disease activity and describing treatment outcomes more comprehensively than its single parameters [102]. The use of composite endpoints broadens measures of the therapeutic effect and can decrease the required sample size or increase the trial's ability to detect smaller efficacy levels [103]. It has to be kept in mind that the application of composite measures may increase the sensitivity to detect disability progression, but does not even out the intrinsic limitations of the single measures within the composite.

4. Discussion

We have identified several outcome measures dealing with different aspects of disease pathology, as well as with different definitions of and the sensibility to predict disease progression. Groups of patients involved in the clinical trials were relatively homogenous as mostly patients with EDSS 3.5–6.0 were included in the studies. Although the most commonly used primary outcome measure was EDSS, we have seen that others, especially composite outcome measures, show a higher level of sensitivity to detect disability progression.

Although the aim of this paper was not to address the definition of SPMS, it remains an important question that should not remain unanswered. Most of the clinical trials have used the definition of SPMS as previously diagnosed RRMS with signs of disability progression (EDSS change ≥ 1.0) before enrolment. In the previous years, a new definition of SPMS has been proposed with an increase of ≥ 1.0 points from baseline scores ≤ 5.5 or ≥ 0.5 points (EDSS ≥ 6.0) in the absence of a relapse, a minimum EDSS score of 4, and a pyramidal FS score of 2, confirmed over ≥ 3 months [104]. An extended discussion about its definition exceeds the aim of this paper but was recently published in a systematic review about the recognition of conversion to SPMS [105].

4.1. Detection of Disability Progression Depends on the Level of Disability

While natural history studies have suggested a relatively uniform pace of progression across patients, it is now well established that the velocity of disability progression—and thus the probability within a certain time frame—depends not only on the sensitivity of the measure but also on the level of disability at the starting point of a given observation period. Obviously, most data are available for EDSS in this regard, where the probability of disease progression significantly decreases for higher EDSS scores, particularly ≥ 7 [106,107].

Patients included in PMS clinical trials shared a comparable level of disability at scores 3.0–6.0; a question that should be put forward is how to evaluate progression in patients with a lower (EDSS ≤ 4.0) or higher (EDSS ≥ 7.0) level of disability. In those patient groups, the additional use of other measures to determine disability progression, such as 9HPT and SDMT, seems reasonable [108,109].

As the baseline measures appear to be predictive of future disability progression, the proposed outcome measure should be carefully chosen according to patients' baseline data. Patients with a lower baseline T25FW (<6 s) show a lower rate of T25FW worsening; a cut-off of baseline T25FW ≥ 8 s predicts 40% of disability progression [45,110,111]. A cut-off of baseline 9HPT ≥ 21 s has been suggested, too, in order to detect a clinically meaningful change [52]. Moreover, patients with a lower baseline SDMT score (< 43) appear to have a higher chance for disability progression, measured by SDMT [64,66].

Among clinical outcome measures, the 9HPT could play an important role in those patients who maintain some degree of manual dexterity but are already wheelchair-dependent (EDSS ≥ 7.0). On the other hand, SDMT should be used in every patient regardless of the prior cognitive impairment or other level of disability, but it should be kept in mind that the sensitivity to change grows with SDMT lower than 43.

Moreover, with the reaching of higher degrees of disability, we should aim to take more paraclinical outcome measures into consideration. Lately, serum neurofilament light

chain (sNfL) levels have been proven to correlate well with disability in both SPMS and PPMS patients and increase over time, even in those whose EDSS remained stable in the observation period [112].

4.2. The Absence of Worsening vs. Improvement of Disability

With potentially neuroprotective drugs being developed in the following years, it is important to distinguish between measuring the treatment effect in worsening vs. the absence of worsening as opposed to improvement vs. stability vs. worsening from a methodological point of view. In the current landscape of DMT for PMS, there is no treatment available achieving sustainable improvement. Thus, an ideal outcome measure would still only show the worsening of disability or stability over time, whereas the improvement of disability would be considered noise due to random variation or measurement error [113]. In that way, Koch et al. analyzed the reliability of the outcome measure and showed that the T25FW and the 9HPT have lower "improvement" rates compared to EDSS, consistent with less measurement error, suggesting them to be reliable outcomes [114]. Kalincik et al. suggested the comparison of the screening and baseline measurements in order to estimate their measurement errors, which could then inform definitions of disability progression and improvement for different baseline values [115].

Although disability improvement was not measured as a distinct outcome in most trials, there are some trials investigating potential neuroprotective drug effects aiming at improvement in outcome measures. A trial investigating high-dose biotin treatment in SPMS defined improvement as a decrease of ≥ 0.5 point (EDSS ≥ 6.0) or ≥ 1.0 point (EDSS ≤ 5.5) or a decrease of $\geq 20\%$ for T25FW, and reported that 13% of patients in the biotin arm (compared to 0% treated with placebo) improved in disability [116]. However, according to a recent analysis, we could expect that approximately 15% of patients (14.6% in the ASCEND and 15.8% in the IMPACT study) would experience such an "improvement," based only on the imprecision of the outcome measure [114]. Therefore, if trials are to use disability improvement as an outcome measure, the anticipated rates of the latter should be corrected accordingly. Importantly, the confounding influence of relapses and subsequent improvement have to be considered, which are often challenging to distinguish from progression, particularly in SPMS with superimposed relapses. To reduce this source of bias, definitions of disability progression should include sustainment for at least 6 months.

4.3. Time Frame for Objectivation of Treatment Response

Equally important to what defines treatment response is the time frame required to define treatment response, i.e., how long and how often do we need to measure. In most trials, disability worsening had to be confirmed after 3–6 months, and/or at the last trial visit in order to be labeled "sustained." When comparing the results for 12 weeks and 24 weeks confirmed disease progression in the ORATORIO study, the magnitude of the effect of ocrelizumab on clinical endpoints was similar [11]. A recent analysis of Koch et al. showed that patients at the 3 month confirmed progression translated to sustained progression at 12 months in 66–88% of patients with SPMS [114,117]. However, the 3 to 6 month confirmed disability progression may still overestimate the accumulation of permanent disability by up to 30% [118]. In that way, disability progression, confirmed at the end of the study, seems to be more accurate, eliminating the effect of a potential relapse and truly determining progression independent of a relapse-associated disability, although possibly leading to fewer worsening events and increasing the required study sample size [119,120]. Hence, future clinical trials should focus more on relapse-independent disability progression, especially as there are some DMTs already available for clinically and/or radiologically active PMS, who are, however, lacking a sufficient effect on the prevention of disability progression. A recent study suggested a better responsiveness using a roving EDSS model, comparing the possible progression with the last EDSS before the progression instead of the study baseline. Whilst this increases the number of events and reduces the required sample size, it inevitably results in increased "false-positives" [22].

Clinical trials traditionally use observation periods of two years. In the SPECTRIMS study, the estimated probabilities of remaining progression free were calculated for each 3 month period throughout the study period. Treatment effects became visible after 9 months of treatment and were significant after 12 months, maintaining significance for each 3 month period throughout the remainder of the study [15]. In that way, the minimal necessary time frame to assess DMT response or failure seems to be at least one year. However, as disability progression remains a rare event, the observation period needs to be extended to at least two years in clinical trials even with a combination of multiple outcome measures. In clinical routine for assessment in individual patients, the required observation period will probably be considerably longer.

4.4. Other Outcome Measures

There are several other potential biomarkers of disease progression that have not been incorporated into phase III clinical trials yet. Optical coherence tomography (OCT) is a noninvasive tool that allows high-resolution measurement of retinal layer thickness, and has already been used in some phase 2 clinical trials in PMS. Peripapillary retinal nerve fiber layer (pRNFL) and ganglion cell-inner plexiform layer (GCIPL) thicknesses are robust indicators of neuroaxonal degeneration [121,122]. They are both significantly reduced in patients with MS regardless of prior optic neuritis, and are associated with both present and future physical and cognitive disabilities, as well as brain atrophy [121–124]. With a specificity of 90% and a sensitivity of 76%, an annual pRNFL thinning rate of more than 1.5 μm is able to distinguish between stable and progressive MS [125]. A recent study found that an annual loss in GCIPL above a cut-off ≥1 μm accurately identifies clinically progressing patients (87% sensitivity and 90% specificity) and presents an even stronger predictor of clinical progression than pRNFL does [126]. As it is a fast, noninvasive imaging method that reliably produces quantitative measures, it has great potential to test neuroprotective strategies over a short time frame (≥12 months).

Still, OCT has some limitations, too, especially in patients with PMS. With the progression of the retinal atrophy, at some point, the floor effect is reached, after which no more thinning is observable. However, the flooring effect is less prominent in GCIPL than in pRNFL [127]. Besides, OCT is not available in every center and requires a lot of expertise and data control in order to be valid.

sNfLs are an emerging biomarker garnering considerable interest over the past few years. NfLs are a major component of the neuronal cytoskeleton, and neuroaxonal damage causes their release into the extracellular space and further into the cerebrospinal fluid and the blood. NfL levels in the cerebrospinal fluid are associated with clinical and radiological activity and show good correlation with future worsening on the EDSS and brain and cervical spinal volume loss [128–131]. The advent of the Single Molecule Array (Simoa®) enabled highly sensitive quantitation of the NfL in the peripheral blood [128]. sNfL levels correlate with disability and increase over time, even in the absence of prior/subsequent disability progression, and are associated with various MRI parameters of neuroaxonal degeneration (T1 black holes, brain and spinal cord atrophy) [112,131–133]. sNfL can be easily obtained and should be therefore considered to be included as an outcome measure in newer progressive MS trials with a potential to detect disease progression more sensitively and especially in a shorter time frame (≥3 months).

sNfL still has some limitations that have to be considered. There are no established reference values that could be interpreted pathologic nor are they specific for MS, as they only imply neuroaxonal damage of any cause. Until now, no definition of an sNfL response to either treatment or disease progression has been established. Furthermore, only a limited set of data is available considering sNfL in patients with PMS [134].

4.5. Recommendations for Future Study Trials and Everyday Clinical Practice

Clinical trials in PMS require study populations with sufficient pre-test probability of disability progression in order to provide sufficient statistical power for detection of

treatment effects and avoid type II error. For want of better prediction tools, this is accomplished by applying recent disability progression as an inclusion criterion. Currently, this is almost exclusively based on the assessment of progression by EDSS, which is insufficient, and we should strive to use more sensitive and complimentary measures.

An important point concerns the sample size needed for each of the outcome measures in order to reliably and validly distinguish progressing and nonprogressing patients. This aspect has become particularly important with the occurrence of newly designed multi-arm trials for investigating DMT for PMS. Where EDSS was used as the primary outcome measure, approximately 20% of patients showed signs of progression within a period of two years. Based on this anticipated incidence in disease progression and an alpha level of 0.05 and power of 80%, at least 400 patients are required for a trial with 1:1 randomization. The same applies for other clinical measures, in which a 20% change is defined as clinically relevant (T25FW and 9HPT). For SDMT with a relevant change of 10%, the sample should be much larger (n = 870). While a combination of multiple clinical parameters would, in theory, reduce the sample size required to about 150 patients per treatment arm, more focus should be placed on biomarkers of neurodegenerative processes such as brain atrophy and retinal thinning. Paraclinical measures would allow us to reduce the sample size as, with higher sensitivity, the anticipated progression rate would be higher, too. Both brain atrophy and OCT measures (pRFNL and GCIPL) require a much smaller sample size in order to detect a clinically relevant change (n = 60 and 120, respectively) [135]. However, less is known about the relevant change in sNfL, and its levels are also dependent on the type of DMT [136].

In clinical trials in PMS, EDSS should be gradually replaced as a primary endpoint in determining disease progression by clinical composite scores such as EDSS-Plus in order to achieve greater sensitivity. Lately, a combinatorial weight-adjusted disability score (CombiWISE) has been developed, integrating four clinical scales, i.e., EDSS, SNRS, T25FW, and 9HPT [137]. However, these composite scores need to be formally validated as a surrogate marker of treatment response, i.e., long-term outcome, in prospective clinical trials. Clinical composite scores should be complemented as secondary or explorative outcome measures by paraclinical parameters, including cerebral or spinal MRI parameters, OCT markers, and sNfL with the goal of defining paraclinical patterns of subclinical disease progression associated with later clinical response. As MRI, OCT, and especially sNfL could determine the degree of neuroaxonal damage within a one-year interval, these patterns could then—after prospective validation—be used as surrogates enabling an earlier assessment of treatment response in order to both objectify and shorten the process.

However, clinical trials are still far away from the actual clinical routine, as not every clinical center has the capacity to perform all the suggested outcome measurements in order to sensitively detect disease progression in their patients. Patient-reported outcomes (PROs), which have also been used in several clinical trials, might seem a reasonable option to overcome a lack of both staff and equipment and might also reflect often neglected aspects such as patients' quality of life and the ability to carry out activities of daily living. However, currently available patient-reported outcome measures are neither specific nor sensitive enough to determine disease progression to fulfil the required standards of reliability and objectivity [138].

We would strongly advocate to at least adopt a minimal set of assessments of PMS patients in addition to EDSS. In ambulatory patients, the performance of T25FW and SDMT would significantly increase sensitivity with an additional capacity of 3–5 min required. In nonambulatory patients, 9HPT should be used instead of T25FW. Assessments should at least be conducted at an interval of six months, with three months desirable if possible (see Figure 2).
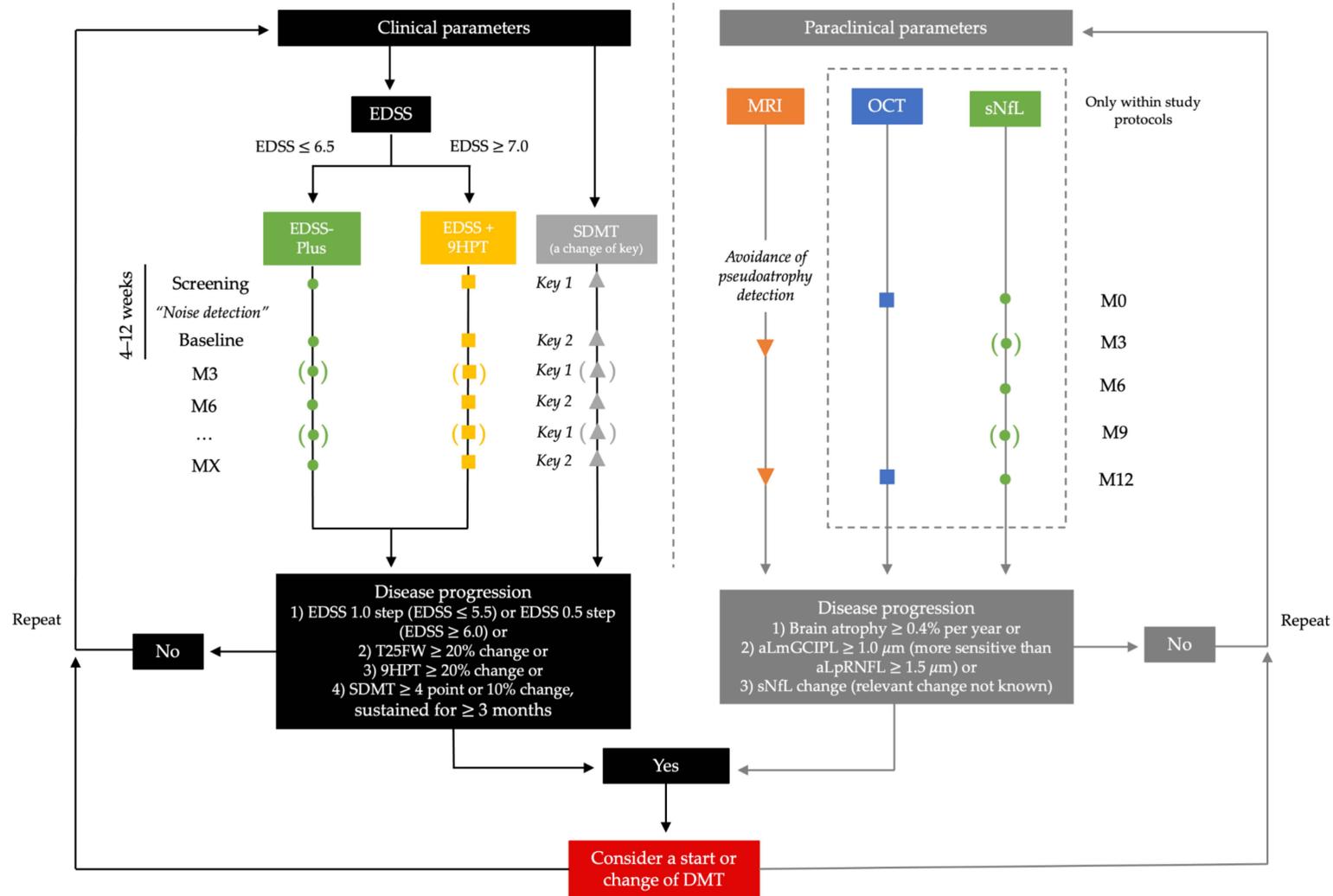
Figure 2. Proposed protocol for assessment of patients with progressive multiple sclerosis in clinical trials and everyday clinical routine. The measurements should be performed every 3 (clinical trials) or 6 months (everyday clinical routine) with 3 months desirable if possible. EDSS—Expanded Disability Status Scale, MRI—magnetic resonance imaging, OCT—optical coherence tomography, SDMT—Symbol Digit Modalities Test, sNfL—serum neurofilament light chain, T25FW—Times 25-Foot Walk Test, 9HPT—9-Hole Peg Test.

Regarding paraclinical assessments, brain and cervical spine MRI should be conducted at an interval of 1–2 years in order to detect signs of inflammatory disease activity and—where feasible—measure atrophy over time. Apart from that, we see a great potential in the bringing of the more available and reproducible biomarkers of disease progression, such as OCT and sNfL, into everyday clinical practice. However, while these cut-offs show good reliability on a group level, they are less reliable when assessing individual patients in everyday clinical practice. That should be taken into account when determining progression independently of study protocols, and needs to be further elucidated.

Anticipating the availability of multiple treatment options for PMS in the not-too-distant future, clinical outcome measures will have to provide the basis for the determination of treatment response or failure and, consequently, when to change DMT in PMS. Considering efficacy profiles of currently available DMTs, as well as substances in the pipeline, the prevention of further disease progression rather than sustainable improvement will represent a realistic treatment goal for the foreseeable future. However, we should not be content with establishing a stable EDSS, but rather strive toward adopting a multivariate definition of "no evidence of progression" incorporating complementary measures. The occurrence of worsening within one or more components should then trigger reevaluation of the benefit–risk ratio and consideration of DMT change (see Figure 2).

5. Conclusions

Several outcome measures have been used in the phase III clinical trials in PMS, with EDSS being most commonly accepted by regulatory agencies, which is unlikely to change in the near future. However, combining it with other measures to create composite outcome measures may provide better sensitivity to detect disease progression. Therefore, a novel trial methodology and development of sensitive and clinically meaningful outcome measures and biomarkers is urgently needed.

In the hopefully not-too-distant future, more and more potentially neuroprotective drugs are going to be studied. In order to identify effective drugs within reasonable time periods and costs, we need more precise and objective outcome parameters. Paraclinical biomarkers have the potential to detect disease progression earlier and more sensitively, and also reflect the change over time even in patients with stable EDSS. Among those, especially MRI, OCT and sNfL should be included in future phase III clinical trials in order to gain knowledge and facilitate the transfer of these biomarkers to clinical routine.

With the goal not only to delay, but even to prevent, disease progression, it is time to move beyond clinical outcome measures toward measuring disease progression that is invisible to the naked eye.

Conflicts of Interest: NK: has participated in meetings sponsored by, and received speaker honoraria or travel funding from, Roche, Novartis, and Merck, and has received a grant for a Multiple Sclerosis Clinical Training Fellowship Programme from ECTRIMS. TB: has participated in meetings sponsored by and received honoraria (lectures, advisory boards, and consultations) from pharmaceutical companies marketing treatments for MS: Allergan, Almirall, Bayer, Biogen, Biologix, Bionorica,

References

1. Oh, J.; Vidal-Jordana, A.; Montalban, X. Multiple sclerosis: Clinical aspects. *Curr. Opin. Neurol.* **2018**, *31*, 752–759. [CrossRef]
2. Miller, D.H.; Leary, S.M. Primary-progressive multiple sclerosis. *Lancet Neurol.* **2007**, *6*, 903–912. [CrossRef]
3. Kister, I.; Bacon, T.E.; Chamot, E.; Salter, A.R.; Cutter, G.R.; Kalina, J.T.; Herbert, J. Natural History of Multiple Sclerosis Symptoms. *Int. J. MS Care* **2013**, *15*, 146–156. [CrossRef]
4. Tutuncu, M.; Tang, J.; Zeid, N.A.; Kale, N.; Crusan, D.J.; Atkinson, E.J.; Siva, A.; Pittock, S.J.; Pirko, I.; Keegan, B.M.; et al. Onset of progressive phase is an age-dependent clinical milestone in multiple sclerosis. *Mult. Scler. J.* **2012**, *19*, 188–198. [CrossRef]
5. Weinshenker, B.G.; Bass, B.; Rice, G.P.; Noseworthy, J.; Carriere, W.; Baskerville, J.; Ebers, G.C. The natural history of multiple sclerosis: A geographically based study. I. Clinical course and disability. *Brain* **1989**, *112*, 133–146. [CrossRef] [PubMed]
6. Confavreux, C.; Vukusic, S.; Adeleine, P. Early clinical predictors and progression of irreversible disability in multiple sclerosis: An amnesic process. *Brain* **2003**, *126*, 770–782. [CrossRef] [PubMed]
7. Runmarker, B.; Andersen, O. Prognostic factors in a multiple sclerosis incidence cohort with twenty-five years of follow-up. *Brain* **1993**, *116*, 117–134. [CrossRef]
8. Bsteh, G.; Ehling, R.; Lutterotti, A.; Hegen, H.; Di Pauli, F.; Auer, M.; Deisenhammer, F.; Reindl, M.; Berger, T. Long Term Clinical Prognostic Factors in Relapsing-Remitting Multiple Sclerosis: Insights from a 10-Year Observational Study. *PLoS ONE* **2016**, *11*, e0158978. [CrossRef]
9. Scott, T.F.; Hackett, C.T.; Quigley, M.R.; Schramke, C.J. Relapsing multiple sclerosis patients treated with disease modifying therapy exhibit highly variable disease progression: A predictive model. *Clin. Neurol. Neurosurg.* **2014**, *127*, 86–92. [CrossRef]
10. Brown, W.; Coles, A.; Horakova, D.; Havrdova, E.; Izquierdo, G.; Prat, A.; Girard, M.; Duquette, P.; Trojano, M.; Lugaresi, A.; et al. Association of Initial Disease-Modifying Therapy with Later Conversion to Secondary Progressive Multiple Sclerosis. *JAMA* **2019**, *321*, 175–187. [CrossRef] [PubMed]
11. Montalban, X.; Hauser, S.L.; Kappos, L.; Arnold, D.L.; Bar-Or, A.; Comi, G.; De Seze, J.; Giovannoni, G.; Hartung, H.-P.; Hemmer, B.; et al. Ocrelizumab versus Placebo in Primary Progressive Multiple Sclerosis. *N. Engl. J. Med.* **2017**, *376*, 209–220. [CrossRef] [PubMed]
12. Kappos, L.; Bar-Or, A.; Cree, B.; Fox, R.J.; Giovannoni, G.; Gold, R.; Vermersch, P.; Arnold, D.L.; Arnould, S.; Scherz, T.; et al. Siponimod versus placebo in secondary progressive multiple sclerosis (EXPAND): A double-blind, randomised, phase 3 study. *Lancet* **2018**, *391*, 1263–1273. [CrossRef]
13. Kurtzke, J.F. Rating neurologic impairment in multiple sclerosis: An expanded disability status scale (EDSS). *Neurology* **1983**, *33*, 1444. [CrossRef] [PubMed]
14. Hawker, K.; O'Connor, P.; Freedman, M.S.; Calabresi, P.A.; Antel, J.; Simon, J.; Hauser, S.; Waubant, E.; Vollmer, T.; Panitch, H.; et al. Rituximab in patients with primary progressive multiple sclerosis: Results of a randomized double-blind placebo-controlled multicenter trial. *Ann. Neurol.* **2009**, *66*, 460–471. [CrossRef] [PubMed]
15. Kappos, L. Placebo-controlled multicentre randomised trial of interferon beta-1b in treatment of secondary progressive multiple sclerosis. European Study Group on interferon beta-1b in secondary progressive MS. *Lancet* **1998**, *352*, 1491–1497. [CrossRef]
16. Wolinsky, J.S.; Narayana, P.A.; O'Connor, P.; Coyle, P.K.; Ford, C.; Johnson, K.; Miller, A.; Pardo, L.; Kadosh, S.; Ladkani, D.; et al. Glatiramer acetate in primary progressive multiple sclerosis: Results of a multinational, multicenter, double-blind, placebo-controlled trial. *Ann. Neurol.* **2007**, *61*, 14–24. [CrossRef]
17. SPECTRIMS Study Group. Secondary Progressive Efficacy Clinical Trial of Recombinant Interferon-Beta-1a in MSSG. Randomized controlled trial of interferon- beta-1a in secondary progressive MS: Clinical results. *Neurology* **2001**, *56*, 1496–1504. [CrossRef]
18. Lublin, F.; Miller, D.H.; Freedman, M.S.; Cree, B.; Wolinsky, J.S.; Weiner, H.; Lubetzki, C.; Hartung, H.-P.; Montalban, X.; Uitdehaag, B.M.J.; et al. Oral fingolimod in primary progressive multiple sclerosis (INFORMS): A phase 3, randomised, double-blind, placebo-controlled trial. *Lancet* **2016**, *387*, 1075–1084. [CrossRef]
19. Lublin, F.D.; Reingold, S.C.; Cohen, J.A.; Cutter, G.R.; Sørensen, P.S.; Thompson, A.J.; Wolinsky, J.S.; Balcer, L.J.; Banwell, B.; Barkhof, F.; et al. Defining the clinical course of multiple sclerosis: The 2013 revisions. *Neurology* **2014**, *83*, 278–286. [CrossRef]
20. Hartung, H.-P.; Gonsette, R.; Konig, N.; Kwiecinski, H.; Guseo, A.; Morrissey, S.P.; Krapf, H.; Zwingers, T. Mitoxantrone in progressive multiple sclerosis: A placebo-controlled, double-blind, randomised, multicentre trial. *Lancet* **2002**, *360*, 2018–2025. [CrossRef]
21. Lepore, V.; Bosetti, C.; Santucci, C.; Iaffaldano, P.; Trojano, M.; Mosconi, P.; Totaro, R.; Coniglio, M.G.; Bossio, R.B.; Valentino, P.; et al. Detection of disability worsening in relapsing-remitting multiple sclerosis patients: A real-world roving Expanded Disability Status Scale reference analysis from the Italian Multiple Sclerosis Register. *Eur. J. Neurol.* **2020**, *28*, 567–578. [CrossRef] [PubMed]

22. Kappos, L.; Butzkueven, H.; Wiendl, H.; Spelman, T.; Pellegrini, F.; Chen, Y.; Dong, Q.; Koendgen, H.; Belachew, S.; Trojano, M.; et al. Greater sensitivity to multiple sclerosis disability worsening and progression events using a roving versus a fixed reference value in a prospective cohort study. *Mult. Scler. J.* **2017**, *24*, 963–973. [CrossRef] [PubMed]

23. Kremenchutzky, M.; Rice, G.P.A.; Baskerville, J.; Wingerchuk, D.M.; Ebers, G.C. The natural history of multiple sclerosis: A geographically based study 9: Observations on the progressive phase of the disease. *Brain* **2006**, *129*, 584–594. [CrossRef]

24. Cinar, B.P.; Yorgun, Y.G. What We Learned from The History of Multiple Sclerosis Measurement: Expanded Disability Status Scale. *Noro. Psikiyatr. Ars.* **2018**, *55*, S69–S75.

25. Sharrack, B.; Hughes, R.A.C.; Soudain, S.; Dunn, G. The psychometric properties of clinical rating scales used in multiple sclerosis. *Brain* **1999**, *122*, 141–159. [CrossRef]

26. Cadavid, D.; Tang, Y.; O'Neill, G. Responsiveness of the Expanded Disability Status Scale (EDSS) to disease progression and therapeutic intervention in progressive forms of multiple sclerosis. *Rev. Neurol.* **2010**, *51*, 321–329.

27. Hyland, M.; Rudick, R.A. Challenges to clinical trials in multiple sclerosis: Outcome measures in the era of disease-modifying drugs. *Curr. Opin. Neurol.* **2011**, *24*, 255–261. [CrossRef]

28. Kragt, J.J.; Thompson, A.; Montalban, X.; Tintore, M.; Rio, J.; Polman, C.H.; Uitdehaag, B. Responsiveness and predictive value of EDSS and MSFC in primary progressive MS. *Neurology* **2008**, *70*, 1084–1091. [CrossRef]

29. Brissart, H.; Sauvée, M.; Latarche, C.; Dillier, C.; Debouverie, M. Integration of Cognitive Impairment in the Expanded Disability Status Scale of 215 Patients with Multiple Sclerosis. *Eur. Neurol.* **2010**, *64*, 345–350. [CrossRef] [PubMed]

30. Amato, M.P.; Fratiglioni, L.; Groppi, C.; Siracusa, G.; Amaducci, L. Interrater Reliability in Assessing Functional Systems and Disability on the Kurtzke Scale in Multiple Sclerosis. *Arch. Neurol.* **1988**, *45*, 746–748. [CrossRef]

31. Goodkin, D.E.; Cookfair, D.; Wende, K.; Bourdette, D.; Pullicino, P.; Scherokman, B.; Whitham, R. Inter- and intrarater scoring agreement using grades 1.0 to 3.5 of the Kurtzke Expanded Disability Status Scale (EDSS). *Neurology* **1992**, *42*, 859. [CrossRef]

32. Noseworthy, J.H.; Vandervoort, M.K.; Wong, C.J.; Ebers, G. Interrater variability with the Expanded Disability Status Scale (EDSS) and Functional Systems (FS) in a multiple sclerosis clinical trial. The Canadian Cooperation MS Study Group. *Neurology* **1990**, *40*, 971. [CrossRef]

33. Meyer-Moock, S.; Feng, Y.-S.; Maeurer, M.; Dippel, F.-W.; Kohlmann, T. Systematic literature review and validity evaluation of the Expanded Disability Status Scale (EDSS) and the Multiple Sclerosis Functional Composite (MSFC) in patients with multiple sclerosis. *BMC Neurol.* **2014**, *14*, 58. [CrossRef] [PubMed]

34. Kieseier, B.C.; Pozzilli, C. Assessing walking disability in multiple sclerosis. *Mult. Scler. J.* **2012**, *18*, 914–924. [CrossRef] [PubMed]

35. Motl, R.W.; Cohen, J.A.; Benedict, R.; Phillips, G.; LaRocca, N.; Hudson, L.D.; Rudick, R.; Multiple Sclerosis Outcome Assessments Consortium. Validity of the timed 25-foot walk as an ambulatory performance outcome measure for multiple sclerosis. *Mult. Scler. J.* **2017**, *23*, 704–710. [CrossRef] [PubMed]

36. Bosma, L.V.; Kragt, J.J.; Knol, D.L.; Polman, C.H.; Uitdehaag, B.M. Clinical scales in progressive MS: Predicting long-term disability. *Mult. Scler. J.* **2011**, *18*, 345–350. [CrossRef]

37. Orbach, R.; Zhao, Z.; Wang, Y.-C.; O'Neill, G.; Cadavid, D. Comparison of Disease Activity in SPMS and PPMS in the Context of Multicenter Clinical Trials. *PLoS ONE* **2012**, *7*, e45409. [CrossRef]

38. Phan-Ba, R.; Pace, A.; Calay, P.; Grodent, P.; Douchamps, F.; Hyde, R.; Hotermans, C.; Delvaux, V.; Hansen, I.; Moonen, G.; et al. Comparison of the Timed 25-Foot and the 100-Meter Walk as Performance Measures in Multiple Sclerosis. *Neurorehabilit. Neural Repair* **2011**, *25*, 672–679. [CrossRef]

39. Kapoor, R.; Ho, P.-R.; Campbell, N.; Chang, I.; Deykin, A.; Forrestal, F.; Lucas, N.; Yu, B.; Arnold, D.L.; Freedman, M.S.; et al. Effect of natalizumab on disease progression in secondary progressive multiple sclerosis (ASCEND): A phase 3, randomised, double-blind, placebo-controlled trial with an open-label extension. *Lancet Neurol.* **2018**, *17*, 405–415. [CrossRef]

40. Kaufman, M.; Moyer, D.; Norton, J. The significant change for the Timed 25-foot Walk in the multiple sclerosis functional composite. *Mult. Scler. J.* **2000**, *6*, 286–290. [CrossRef]

41. Coleman, C.I.; Sobieraj, D.; Marinucci, L.N. Minimally important clinical difference of the Timed 25-Foot Walk Test: Results from a randomized controlled trial in patients with multiple sclerosis. *Curr. Med. Res. Opin.* **2011**, *28*, 49–56. [CrossRef] [PubMed]

42. Kragt, J.J.; Linden, F.A.H.V.D.; Nielsen, J.M.; Uitdehaag, B.M.J.; Polman, C.H. Clinical impact of 20% worsening on Timed 25-foot Walk and 9-hole Peg Test in multiple sclerosis. *Mult. Scler. J.* **2006**, *12*, 594–598. [CrossRef] [PubMed]

43. Bosma, L.V.; Kragt, J.J.; Brieva, L.; Khaleeli, Z.; Montalban, X.; Polman, C.H.; Thompson, A.J.; Tintoré, M.; Uitdehaag, B.M.J. Progression on the Multiple Sclerosis Functional Composite in multiple sclerosis: What is the optimal cut-off for the three components? *Mult. Scler. J.* **2010**, *16*, 862–867. [CrossRef] [PubMed]

44. Hobart, J.; Blight, A.R.; Goodman, A.; Lynn, F.; Putzki, N. Timed 25-Foot Walk: Direct evidence that improving 20% or greater is clinically meaningful in MS. *Neurology* **2013**, *80*, 1509–1517. [CrossRef]

45. Koch, M.; Mostert, J.; Uitdehaag, B.; Cutter, G. Clinical outcome measures in SPMS trials: An analysis of the IMPACT and ASCEND original trial data sets. *Mult. Scler. J.* **2019**, *26*, 1540–1549. [CrossRef] [PubMed]

46. Sikes, E.M.; Cederberg, K.L.; Sandroff, B.M.; Bartolucci, A.; Motl, R.W. Quantitative Synthesis of Timed 25-Foot Walk Performance in Multiple Sclerosis. *Arch. Phys. Med. Rehabil.* **2019**, *101*, 524–534. [CrossRef]

47. Goldman, M.D.; LaRocca, N.G.; Rudick, R.A.; Hudson, L.D.; Chin, P.S.; Francis, G.S.; Jacobs, A.; Kapoor, R.; Matthews, P.M.; Mowry, E.M.; et al. Evaluation of multiple sclerosis disability outcome measures using pooled clinical trial data. *Neurology* **2019**, *93*, e1921–e1931. [CrossRef]

48. Cutter, G.R.; Baier, M.L.; Rudick, R.A.; Cookfair, D.L.; Fischer, J.S.; Petkau, J.; Syndulko, K.; Weinshenker, B.G.; Antel, J.; Confavreux, C.; et al. Development of a multiple sclerosis functional composite as a clinical trial outcome measure. *Brain* **1999**, *122*, 871–882. [CrossRef]

49. Goodkin, D.E.; Hertsgaard, D.; Seminary, J. Upper extremity function in multiple sclerosis: Improving assessment sensitivity with box-and-block and nine-hole peg tests. *Arch. Phys. Med. Rehabil.* **1988**, *69*, 850–854. [PubMed]

50. Bertoni, R.; Lamers, I.; Chen, C.C.; Feys, P.; Cattaneo, D. Unilateral and bilateral upper limb dysfunction at body functions, activity and participation levels in people with multiple sclerosis. *Mult. Scler. J.* **2015**, *21*, 1566–1574. [CrossRef] [PubMed]

51. Lamers, I.; Cattaneo, D.; Chen, C.C.; Bertoni, R.; Van Wijmeersch, B.; Feys, P. Associations of Upper Limb Disability Measures on Different Levels of the International Classification of Functioning, Disability and Health in People with Multiple Sclerosis. *Phys. Ther.* **2015**, *95*, 65–75. [CrossRef]

52. Solaro, C.; Cattaneo, D.; Brichetto, G.; Castelli, L.; Tacchino, A.; Gervasoni, E.; Prosperini, L. Clinical correlates of 9-hole peg test in a large population of people with multiple sclerosis. *Mult. Scler. Relat. Disord.* **2019**, *30*, 1–8. [CrossRef] [PubMed]

53. Rosti-Otajärvi, E.; Hämäläinen, P.; Koivisto, K.; Hokkanen, L. The reliability of the MSFC and its components. *Acta Neurol. Scand.* **2007**, *117*, 421–427. [CrossRef] [PubMed]

54. Feys, P.; Lamers, I.; Francis, G.; Benedict, R.; Phillips, G.; LaRocca, N.; Hudson, L.D.; Rudick, R.; Multiple Sclerosis Outcome Assessments Consortium. The Nine-Hole Peg Test as a manual dexterity performance measure for multiple sclerosis. *Mult. Scler. J.* **2017**, *23*, 711–720. [CrossRef] [PubMed]

55. Tombaugh, T.N. A comprehensive review of the Paced Auditory Serial Addition Test (PASAT). *Arch. Clin. Neuropsychol.* **2006**, *21*, 53–76. [CrossRef]

56. Gronwall, D.M.A. Paced Auditory Serial-Addition Task: A Measure of Recovery from Concussion. *Percept. Mot. Ski.* **1977**, *44*, 367–373. [CrossRef]

57. Rudick, R.A.; Cutter, G.; Reingold, S. The multiple sclerosis functional composite: A new clinical outcome measure for multiple sderosis trials. *Mult. Scler. J.* **2002**, *8*, 359–365. [CrossRef] [PubMed]

58. Brooks, J.B.B.; Giraud, V.O.; Saleh, Y.J.; Rodrigues, S.J.; Daia, L.A.; Fragoso, Y.D. Paced auditory serial addition test (PASAT): A very difficult test even for individuals with high intellectual capability. *Arq. Neuro-Psiquiatria* **2011**, *69*, 482–484. [CrossRef]

59. Van Schependom, J.; D'Hooghe, M.B.; Cleynhens, K.; D'hooge, M.; Haelewyck, M.C.; De Keyser, J.; Nagels, G. The Symbol Digit Modalities Test as sentinel test for cognitive impairment in multiple sclerosis. *Eur. J. Neurol.* **2014**, *21*, 1219-e72. [CrossRef]

60. Kalb, R.; Beier, M.; Benedict, R.H.; Charvet, L.; Costello, K.; Feinstein, A.; Gingold, J.; Goverover, Y.; Halper, J.; Harris, C.; et al. Recommendations for cognitive screening and management in multiple sclerosis care. *Mult. Scler. J.* **2018**, *24*, 1665–1680. [CrossRef]

61. Strober, L.; DeLuca, J.; Benedict, R.H.; Jacobs, A.; Cohen, J.A.; Chiaravalloti, N.; Hudson, L.D.; Rudick, R.A.; LaRocca, N.G.; Multiple Sclerosis Outcome Assessments Consortium (MSOAC). Symbol Digit Modalities Test: A valid clinical trial endpoint for measuring cognition in multiple sclerosis. *Mult. Scler. J.* **2018**, *25*, 1781–1790. [CrossRef]

62. Ontaneda, D.; LaRocca, N.; Coetzee, T.; Rudick, R. Revisiting the Multiple Sclerosis Functional Composite: Proceedings from the National Multiple Sclerosis Society (NMSS) Task Force on Clinical Disability Measures. *Mult. Scler. J.* **2012**, *18*, 1074–1080. [CrossRef] [PubMed]

63. López-Góngora, M.; Querol, L.; Escartín, A. A one-year follow-up study of the Symbol Digit Modalities Test (SDMT) and the Paced Auditory Serial Addition Test (PASAT) in relapsing-remitting multiple sclerosis: An appraisal of comparative longitudinal sensitivity. *BMC Neurol.* **2015**, *15*, 1–8. [CrossRef]

64. Benedict, R.H.B.; Tomic, D.; Cree, B.A.; Fox, R.; Giovannoni, G.; Bar-Or, A.; Gold, R.; Vermersch, P.; Pohlmann, H.; Wright, I.; et al. Siponimod and cognition in secondary progressive multiple sclerosis: EXPAND secondary analyses. *Neurology* **2021**, *96*, e376–e386. [CrossRef] [PubMed]

65. Benedict, R.H.; DeLuca, J.; Phillips, G.; LaRocca, N.; Hudson, L.D.; Rudick, R.; Multiple Sclerosis Outcome Assessments Consortium. Validity of the Symbol Digit Modalities Test as a cognition performance outcome measure for multiple sclerosis. *Mult. Scler. J.* **2017**, *23*, 721–733. [CrossRef] [PubMed]

66. Drake, A.; Weinstock-Guttman, B.; Morrow, S.A.; Hojnacki, D.; Munschauer, F.E.; Benedict, R.H.B. Psychometrics and normative data for the Multiple Sclerosis Functional Composite: Replacing the PASAT with the Symbol Digit Modalities Test. *Mult. Scler. J.* **2009**, *16*, 228–237. [CrossRef]

67. Kiely, K.M.; Butterworth, P.; Watson, N.; Wooden, M. The Symbol Digit Modalities Test: Normative Data from a Large Nationally Representative Sample of Australians. *Arch. Clin. Neuropsychol.* **2014**, *29*, 767–775. [CrossRef]

68. Roar, M.; Illes, Z.; Sejbaek, T. Practice effect in Symbol Digit Modalities Test in multiple sclerosis patients treated with natalizumab. *Mult. Scler. Relat. Disord.* **2016**, *10*, 116–122. [CrossRef]

69. Galetta, K.M.; Balcer, L.J. Measures of visual pathway structure and function in MS: Clinical usefulness and role for MS trials. *Mult. Scler. Relat. Disord.* **2013**, *2*, 172–182. [CrossRef]

70. Balcer, L.J.; Raynowska, J.; Nolan, R.; Galetta, S.L.; Kapoor, R.; Benedict, R.; Phillips, G.; LaRocca, N.; Hudson, L.; Rudick, R.; et al. Validity of low-contrast letter acuity as a visual performance outcome measure for multiple sclerosis. *Mult. Scler. J.* **2017**, *23*, 734–747. [CrossRef]

71. Talman, L.S.; Bisker, E.R.; Sackel, D.J.; Long, D.A., Jr.; Galetta, K.M.; Ratchford, J.N.; Lile, D.J.; Farrell, S.K.; Loguidice, M.J.; Remington, G.; et al. Longitudinal study of vision and retinal nerve fiber layer thickness in multiple sclerosis. *Ann. Neurol.* **2010**, *67*, 749–760.
72. Walter, S.D.; Ishikawa, H.; Galetta, K.M.; Sakai, R.E.; Feller, D.; Henderson, S.B.; Wilson, J.A.; Maguire, M.G.; Galetta, S.L.; Frohman, E.; et al. Ganglion Cell Loss in Relation to Visual Disability in Multiple Sclerosis. *Ophthalmology* **2012**, *119*, 1250–1257. [CrossRef]
73. Reich, D.S.; Smith, S.A.; Gordon-Lipkin, E.; Ozturk, A.; Caffo, B.S.; Balcer, L.J.; Calabresi, P.A. Damage to the Optic Radiation in Multiple Sclerosis Is Associated With Retinal Injury and Visual Disability. *Arch. Neurol.* **2009**, *66*, 998–1006. [CrossRef]
74. Cohen, J.A.; Cutter, G.R.; Fischer, J.S.; Goodman, A.D.; Heidenreich, F.R.; Kooijmans, M.F.; Sandrock, A.W.; Rudick, R.A.; Simon, J.H.; Simonian, N.A.; et al. Benefit of interferon beta-1a on MSFC progression in secondary progressive MS. *Neurology* **2002**, *59*, 679–687. [CrossRef]
75. Balcer, L.J.; Baier, M.L.; Pelak, V.S.; Fox, R.J.; Shuwairi, S.; Galetta, S.L.; Cutter, G.R.; Maguire, M.G. New low-contrast vision charts: Reliability and test characteristics in patients with multiple sclerosis. *Mult. Scler.* **2000**, *6*, 163–171. [CrossRef] [PubMed]
76. Balcer, L.J.; Galetta, S.L.; Polman, C.H.; Eggenberger, E.; Calabresi, P.A.; Zhang, A.; Scanlon, J.V.; Hyde, R. Low-contrast acuity measures visual improvement in phase 3 trial of natalizumab in relapsing MS. *J. Neurol. Sci.* **2012**, *318*, 119–124. [CrossRef] [PubMed]
77. Beck, R.W.; Maguire, M.G.; Bressler, N.M.; Glassman, A.R.; Lindblad, A.S.; Ferris, F.L. Visual Acuity as an Outcome Measure in Clinical Trials of Retinal Diseases. *Ophthalmology* **2007**, *114*, 1804–1809. [CrossRef]
78. Jakimovski, D.; Benedict, R.H.B.; Weinstock-Guttman, B.; Ozel, O.; Fuchs, T.A.; Lincoff, N.; Bergsland, N.; Dwyer, M.G.; Zivadinov, R. Visual deficits and cognitive assessment of multiple sclerosis: Confounder, correlate, or both? *J. Neurol.* **2021**, *268*, 2578–2588. [CrossRef] [PubMed]
79. Balcer, L.J.; Miller, D.H.; Reingold, S.C.; Cohen, J.A. Vision and vision-related outcome measures in multiple sclerosis. *Brain* **2014**, *138*, 11–27. [CrossRef]
80. Sipe, J.C.; Knobler, R.L.; Braheny, S.L.; Rice, G.P.; Panitch, H.S.; Oldstone, M.B. A neurologic rating scale (NRS) for use in multiple sclerosis. *Neurology* **1984**, *34*, 1368. [CrossRef]
81. Koziol, J.A.; Lucero, A.; Sipe, J.C.; Romine, J.S.; Beutler, E. Responsiveness of the Scripps neurologic rating scale during a multiple sclerosis clinical trial. *Can. J. Neurol. Sci. J. Can. Sci. Neurol.* **1999**, *26*, 283–289. [CrossRef]
82. Sipe, J.C.; Romine, J.S.; Koziol, J.A.; McMillan, R.; Beutler, E.; Sipe, J.C.; Romine, J.S.; Zyroff, J. Cladribine in treatment of chronic progressive multiple sclerosis. *Lancet* **1994**, *344*, 9–13. [CrossRef]
83. Eshaghi, A.; Prados, F.; Brownlee, W.J.; Altmann, D.R.; Tur, C.; Cardoso, M.J.; De Angelis, F.; van de Pavert, S.H.; Cawley, N.; De Stefano, N.; et al. Deep gray matter volume loss drives disability worsening in multiple sclerosis. *Ann. Neurol.* **2018**, *83*, 210–222. [CrossRef]
84. De Stefano, N.; Stromillo, M.L.; Giorgio, A.; Bartolozzi, M.L.; Battaglini, M.; Baldini, M.; Portaccio, E.; Amato, M.P.; Sormani, M.P. Establishing pathological cut-offs of brain atrophy rates in multiple sclerosis. *J. Neurol. Neurosurg. Psychiatry* **2016**, *87*, 93–99. [CrossRef] [PubMed]
85. Eijlers, A.J.; Dekker, I.; Steenwijk, M.D.; Meijer, K.A.; Hulst, H.E.; Pouwels, P.J.; Uitdehaag, B.M.; Barkhof, F.; Vrenken, H.; Schoonheim, M.M.; et al. Cortical atrophy accelerates as cognitive decline worsens in multiple sclerosis. *Neurology* **2019**, *93*, e1348–e1359. [CrossRef]
86. Duning, T.; Kloska, S.; Steinstrater, O.; Kugel, H.; Heindel, W.; Knecht, S. Dehydration confounds the assessment of brain atrophy. *Neurology* **2005**, *64*, 548–550. [CrossRef] [PubMed]
87. Nakamura, K.; Brown, R.A.; Narayanan, S.; Collins, D.L.; Arnold, D.L. Diurnal fluctuations in brain volume: Statistical analyses of MRI from large populations. *NeuroImage* **2015**, *118*, 126–132. [CrossRef]
88. Enzinger, C.; Fazekas, F.; Matthews, P.M.; Ropele, S.; Schmidt, H.; Smith, S. Risk factors for progression of brain atrophy in aging: Six-year follow-up of normal subjects. *Neurology* **2005**, *64*, 1704–1711. [CrossRef]
89. De Stefano, N.; Airas, L.; Grigoriadis, N.; Mattle, H.P.; O'Riordan, J.; Oreja-Guevara, C.; Sellebjerg, F.; Stankoff, B.; Walczak, A.; Wiendl, H.; et al. Clinical relevance of brain volume measures in multiple sclerosis. *CNS Drugs* **2014**, *28*, 147–156.
90. Zivadinov, R.; Raj, B.; Ramanathan, M.; Teter, B.; Durfee, J.; Dwyer, M.; Bergsland, N.; Kolb, C.; Hojnacki, D.; Benedict, R.; et al. Autoimmune Comorbidities Are Associated with Brain Injury in Multiple Sclerosis. *Am. J. Neuroradiol.* **2016**, *37*, 1010–1016. [CrossRef] [PubMed]
91. Zivadinov, R.; Reder, A.T.; Filippi, M.; Minagar, A.; Stuve, O.; Lassmann, H.; Racke, M.K.; Dwyer, M.; Frohman, E.M.; Khan, O. Mechanisms of action of disease-modifying agents and brain volume changes in multiple sclerosis. *Neurology* **2008**, *71*, 136–144. [CrossRef]
92. Song, X.; Li, D.; Qiu, Z.; Su, S.; Wu, Y.; Wang, J.; Liu, Z.; Dong, H. Correlation between EDSS scores and cervical spinal cord atrophy at 3T MRI in multiple sclerosis: A systematic review and meta-analysis. *Mult. Scler. Relat. Disord.* **2019**, *37*, 101426. [CrossRef]
93. Casserly, C.; Seyman, E.; Alcaide-Leon, P.; Guenette, M.; Lyons, C.; Sankar, S.; Svendrovski, A.; Baral, S.; Oh, J. Spinal Cord Atrophy in Multiple Sclerosis: A Systematic Review and Meta-Analysis. *J. Neuroimaging* **2018**, *28*, 556–586. [CrossRef] [PubMed]
94. Moccia, M.; Ruggieri, S.; Ianniello, A.; Toosy, A.; Pozzilli, C.; Ciccarelli, O. Advances in spinal cord imaging in multiple sclerosis. *Ther. Adv. Neurol. Disord.* **2019**, *12*. [CrossRef] [PubMed]

95. Rudick, R.; Antel, J.; Confavreux, C.; Cutter, G.; Ellison, G.; Fischer, J.; Lublin, F.; Miller, A.; Petkau, J.; Rao, S.; et al. Recommendations from the national multiple sclerosis society clinical outcomes assessment task force. *Ann. Neurol.* **1997**, *42*, 379–382. [CrossRef] [PubMed]

96. Fischer, J.S.; Rudick, R.A.; Cutter, G.R.; Reingold, S.C.; National MS Society Clinical Outcomes Assessment Task Force. The Multiple Sclerosis Functional Composite Measure (MSFC): An integrated approach to MS clinical outcome assessment. National MS Society Clinical Outcomes Assessment Task Force. *Mult. Scler. J.* **1999**, *5*, 244–250. [CrossRef] [PubMed]

97. Cadavid, D.; Cohen, J.A.; Freedman, M.S.; Goldman, M.D.; Hartung, H.-P.; Havrdova, E.K.; Jeffery, D.; Kapoor, R.; Miller, A.; Sellebjerg, F.; et al. The EDSS-Plus, an improved endpoint for disability progression in secondary progressive multiple sclerosis. *Mult. Scler. J.* **2016**, *23*, 94–105. [CrossRef]

98. Inojosa, H.; Schriefer, D.; Ziemssen, T. Clinical outcome measures in multiple sclerosis: A review. *Autoimmun. Rev.* **2020**, *19*, 102512. [CrossRef]

99. Bin Sawad, A.; Seoane-Vazquez, E.; Rodriguez-Monguio, R.; Turkistani, F. Evaluation of the Expanded Disability Status Scale and the Multiple Sclerosis Functional Composite as clinical endpoints in multiple sclerosis clinical trials: Quantitative meta-analyses. *Curr. Med. Res. Opin.* **2016**, *32*, 1969–1974. [CrossRef]

100. Bosma, L.; Kragt, J.; Brieva, L.; Khaleeli, Z.; Montalban, X.; Polman, C.; Thompson, A.; Tintoré, M.; Uitdehaag, B.; Thompson, A. The search for responsive clinical endpoints in primary progressive multiple sclerosis. *Mult. Scler. J.* **2009**, *15*, 715–720. [CrossRef]

101. Wolinsky, J.S.; Montalban, X.; Hauser, S.L.; Giovannoni, G.; Vermersch, P.; Bernasconi, C.; Deol-Bhullar, G.; Garren, H.; Chin, P.; Belachew, S.; et al. Evaluation of no evidence of progression or active disease (NEPAD) in patients with primary progressive multiple sclerosis in the ORATORIO trial. *Ann. Neurol.* **2018**, *84*, 527–536. [CrossRef] [PubMed]

102. Hegen, H.; Bsteh, G.; Berger, T. 'No evidence of disease activity'—Is it an appropriate surrogate in multiple sclerosis? *Eur. J. Neurol.* **2018**, *25*, 1107-e101. [CrossRef]

103. Goldberg, R.; Gore, J.M.; Barton, B.; Gurwitz, J. Individual and Composite Study Endpoints: Separating the Wheat from the Chaff. *Am. J. Med.* **2014**, *127*, 379–384. [CrossRef]

104. Lorscheider, J.; Buzzard, K.; Jokubaitis, V.; Spelman, T.; Havrdova, E.; Horakova, D.; Trojano, M.; Izquierdo, G.; Girard, M.; Duquette, P.; et al. Defining secondary progressive multiple sclerosis. *Brain* **2016**, *139*, 2395–2405. [CrossRef]

105. Krajnc, N.; Bsteh, G.; Berger, T. Clinical and Paraclinical Biomarkers and the Hitches to Assess Conversion to Secondary Progressive Multiple Sclerosis: A Systematic Review. *Front. Neurol.* **2021**, *12*, 666868. [CrossRef]

106. Hohol, M.J.; Orav, E.J.; Weiner, H.L. Disease steps in multiple sclerosis: A longitudinal study comparing disease steps and EDSS to evaluate disease progression. *Mult. Scler. J.* **1999**, *5*, 349–354. [CrossRef] [PubMed]

107. Ravnborg, M.; Blinkenberg, M.; Sellebjerg, F.; Ballegaard, M.; Larsen, S.H.; Sørensen, P.S. Responsiveness of the Multiple Sclerosis Impairment Scale in comparison with the Expanded Disability Status Scale. *Mult. Scler. J.* **2005**, *11*, 81–84. [CrossRef] [PubMed]

108. University of California SFMSET; Cree BAC; Hollenbach, J.A.; Bove, R.; Kirkish, G.; Sacco, S.; Caverzasi, E.; Bischof, A.; Gundel, T.; Zhu, A.H.; et al. Silent progression in disease activity-free relapsing multiple sclerosis. *Ann. Neurol.* **2019**, *85*, 653–666.

109. Benedict, R.H.; Pol, J.; Yasin, F.; Hojnacki, D.; Kolb, C.; Eckert, S.; Tacca, B.; Drake, A.; Wojcik, C.; Morrow, S.A.; et al. Recovery of cognitive function after relapse in multiple sclerosis. *Mult. Scler. J.* **2020**, *27*, 71–78. [CrossRef]

110. Kalinowski, A.; Cutter, G.; Bozinov, N.; Hinman, J.A.; Hittle, M.; Motl, R.; Odden, M.; Nelson, L.M. The timed 25-foot walk in a large cohort of multiple sclerosis patients. *Mult. Scler. J.* **2021**, 13524585211017013. [CrossRef]

111. Goldman, M.D.; Motl, R.W.; Scagnelli, J.; Pula, J.H.; Sosnoff, J.J.; Cadavid, D. Clinically meaningful performance benchmarks in MS: Timed 25-Foot Walk and the real world. *Neurology* **2013**, *81*, 1856–1863. [CrossRef]

112. Ferraro, D.; Guicciardi, C.; DE Biasi, S.; Pinti, M.; Bedin, R.; Camera, V.; Vitetta, F.; Nasi, M.; Meletti, S.; Sola, P. Plasma neurofilaments correlate with disability in progressive multiple sclerosis patients. *Acta Neurol. Scand.* **2019**, *141*, 16–21. [CrossRef]

113. Ebers, G.C.; Heigenhauser, L.; Daumer, M.; Lederer, C.; Noseworthy, J.H. Disability as an outcome in MS clinical trials. *Neurology* **2008**, *71*, 624–631. [CrossRef] [PubMed]

114. Koch, M.W.; Mostert, J.; Repovic, P.; Bowen, J.D.; Uitdehaag, B.; Cutter, G. Reliability of Outcome Measures in Clinical Trials in Secondary Progressive Multiple Sclerosis. *Neurology* **2020**, *96*, e111–e120. [CrossRef]

115. Kalincik, T.; Sormani, M.P.; Tur, C. Has the Time Come to Revisit Our Standard Measures of Disability Progression in Multiple Sclerosis? *Neurology* **2021**, *96*, 12–13. [CrossRef] [PubMed]

116. Tourbah, A.; Lebrun-Frenay, C.; Edan, G.; Clanet, M.; Papeix, C.; Vukusic, S.; De Sèze, J.; Debouverie, M.; Gout, O.; Clavelou, P.; et al. MD1003 (high-dose biotin) for the treatment of progressive multiple sclerosis: A randomised, double-blind, placebo-controlled study. *Mult. Scler. J.* **2016**, *22*, 1719–1731. [CrossRef]

117. Liu, C.; Blumhardt, L.D. Disability outcome measures in therapeutic trials of relapsing-remitting multiple sclerosis: Effects of heterogeneity of disease course in placebo cohorts. *J. Neurol. Neurosurg. Psychiatry* **2000**, *68*, 450–457. [CrossRef]

118. Kalincik, T.; Cutter, G.; Spelman, T.; Jokubaitis, V.; Havrdova, E.K.; Horáková, D.; Trojano, M.; Izquierdo, G.; Girard, M.; Duquette, P.; et al. Defining reliable disability outcomes in multiple sclerosis. *Brain* **2015**, *138*, 3287–3298. [CrossRef] [PubMed]

119. Ontaneda, D.; Cohen, J.A.; Amato, M.P. Clinical outcome measures for progressive MS trials. *Mult. Scler. J.* **2017**, *23*, 1627–1635. [CrossRef]

120. Kappos, L.; Wolinsky, J.S.; Giovannoni, G.; Arnold, D.L.; Wang, Q.; Bernasconi, C.; Model, F.; Koendgen, H.; Manfrini, M.; Belachew, S.; et al. Contribution of Relapse-Independent Progression vs Relapse-Associated Worsening to Overall Confirmed Disability Accumulation in Typical Relapsing Multiple Sclerosis in a Pooled Analysis of 2 Randomized Clinical Trials. *JAMA Neurol.* **2020**, *77*, 1132–1140. [CrossRef] [PubMed]

121. Saidha, S.; Al-Louzi, O.; Ratchford, J.N.; Bhargava, P.; Oh, J.; Newsome, S.D.; Prince, J.L.; Pham, D.; Roy, S.; Van Zijl, P.; et al. Optical coherence tomography reflects brain atrophy in multiple sclerosis: A four-year study. *Ann. Neurol.* **2015**, *78*, 801–813. [CrossRef] [PubMed]

122. Saidha, S.; Sotirchos, E.S.; Oh, J.; Syc, S.B.; Seigo, M.A.; Shiee, N.; Eckstein, C.; Durbin, M.K.; Oakley, J.D.; Meyer, S.A.; et al. Relationships between retinal axonal and neuronal measures and global central nervous system pathology in multiple sclerosis. *JAMA Neurol.* **2013**, *70*, 34–43. [CrossRef]

123. Bsteh, G.; Hegen, H.; Teuchner, B.; Amprosi, M.; Berek, K.; Ladstätter, F.; Wurth, S.; Auer, M.; Di Pauli, F.; Deisenhammer, F.; et al. Peripapillary retinal nerve fibre layer as measured by optical coherence tomography is a prognostic biomarker not only for physical but also for cognitive disability progression in multiple sclerosis. *Mult. Scler. J.* **2017**, *25*, 196–203. [CrossRef]

124. Martínez-Lapiscina, E.H.; Arnow, S.; Wilson, J.A.; Saidha, S.; Preiningerova, J.L.; Oberwahrenbrock, T.; Brandt, A.U.; Pablo, L.E.; Guerrieri, S.; González-Suárez, I.; et al. Retinal thickness measured with optical coherence tomography and risk of disability worsening in multiple sclerosis: A cohort study. *Lancet Neurol.* **2016**, *15*, 574–584. [CrossRef]

125. Bsteh, G.; Hegen, H.; Teuchner, B.; Berek, K.; Wurth, S.; Auer, M.; Di Pauli, F.; Deisenhammer, F.; Berger, T. Peripapillary retinal nerve fibre layer thinning rate as a biomarker discriminating stable and progressing relapsing–remitting multiple sclerosis. *Eur. J. Neurol.* **2019**, *26*, 865–871. [CrossRef]

126. Bsteh, G.; Berek, K.; Hegen, H.; Altmann, P.; Wurth, S.; Auer, M.; Zinganell, A.; Di Pauli, F.; Rommer, P.; Leutmezer, F.; et al. Macular ganglion cell–inner plexiform layer thinning as a biomarker of disability progression in relapsing multiple sclerosis. *Mult. Scler. J.* **2020**, *27*, 684–694. [CrossRef]

127. Bowd, C.; Zangwill, L.M.; Weinreb, R.N.; Medeiros, F.A.; Belghith, A. Estimating Optical Coherence Tomography Structural Measurement Floors to Improve Detection of Progression in Advanced Glaucoma. *Am. J. Ophthalmol.* **2016**, *175*, 37–44. [CrossRef]

128. Disanto, G.; Barro, C.; Benkert, P.; Naegelin, Y.; Schädelin, S.; Giardiello, A.; Zecca, C.; Blennow, K.; Zetterberg, H.; Leppert, D.; et al. Serum Neurofilament light: A biomarker of neuronal damage in multiple sclerosis. *Ann. Neurol.* **2017**, *81*, 857–870. [CrossRef] [PubMed]

129. Novakova, L.; Zetterberg, H.; Sundström, P.; Axelsson, M.; Khademi, M.; Gunnarsson, M.; Malmeström, C.; Svenningsson, A.; Olsson, T.; Piehl, F.; et al. Monitoring disease activity in multiple sclerosis using serum neurofilament light protein. *Neurology* **2017**, *89*, 2230–2237. [CrossRef]

130. Damasceno, A.; Dias-Carneiro, R.; Moraes, A.; Boldrini, V.O.; Quintiliano, R.P.S.; Silva, V.A.D.P.G.D.; Farias, A.; Brandão, C.O.; Damasceno, B.P.; dos Santos, L.M.B.; et al. Clinical and MRI correlates of CSF neurofilament light chain levels in relapsing and progressive MS. *Mult. Scler. Relat. Disord.* **2019**, *30*, 149–153. [CrossRef]

131. Barro, C.; Benkert, P.; Disanto, G.; Tsagkas, C.; Amann, M.; Naegelin, Y.; Leppert, D.; Gobbi, C.; Granziera, C.; Yaldizli, Ö.; et al. Serum neurofilament as a predictor of disease worsening and brain and spinal cord atrophy in multiple sclerosis. *Brain* **2018**, *141*, 2382–2391. [CrossRef]

132. Kuhle, J.; Kropshofer, H.; Haering, D.A.; Kundu, U.; Meinert, R.; Barro, C.; Dahlke, F.; Tomic, D.; Leppert, D.; Kappos, L. Blood neurofilament light chain as a biomarker of MS disease activity and treatment response. *Neurology* **2019**, *92*, e1007–e1015. [CrossRef]

133. Sellebjerg, F.; Royen, L.; Sørensen, P.S.; Oturai, A.B.; Jensen, P.E.H. Prognostic value of cerebrospinal fluid neurofilament light chain and chitinase-3-like-1 in newly diagnosed patients with multiple sclerosis. *Mult. Scler. J.* **2018**, *25*, 1444–1451. [CrossRef] [PubMed]

134. Kapoor, R.; Smith, K.E.; Allegretta, M.; Arnold, D.L.; Carroll, W.; Comabella, M.; Furlan, R.; Harp, C.; Kuhle, J.; Leppert, D.; et al. Serum neurofilament light as a biomarker in progressive multiple sclerosis. *Neurology* **2020**, *95*, 436–444. [CrossRef] [PubMed]

135. Altmann, D.R.; Jasperse, B.; Barkhof, F.; Beckmann, K.; Filippi, M.; Kappos, L.D.; Molyneux, P.; Polman, C.H.; Pozzilli, C.; Thompson, A.; et al. Sample sizes for brain atrophy outcomes in trials for secondary progressive multiple sclerosis. *Neurology* **2008**, *72*, 595–601. [CrossRef] [PubMed]

136. Delcoigne, B.; Manouchehrinia, A.; Barro, C.; Benkert, P.; Michalak, Z.; Kappos, L.; Leppert, D.; Tsai, J.A.; Plavina, T.; Kieseier, B.C.; et al. Blood neurofilament light levels segregate treatment effects in multiple sclerosis. *Neurology* **2020**, *94*, e1201–e1212. [CrossRef]

137. Kosa, P.; Ghazali, D.; Tanigawa, M.; Barbour, C.; Cortese, I.; Kelley, W.; Snyder, B.; Ohayon, J.; Fenton, K.; Lehky, T.; et al. Development of a Sensitive Outcome for Economical Drug Screening for Progressive Multiple Sclerosis Treatment. *Front. Neurol.* **2016**, *7*, 131. [CrossRef] [PubMed]

138. Khurana, V.; Sharma, H.; Afroz, N.; Callan, A.; Medin, J. Patient-reported outcomes in multiple sclerosis: A systematic comparison of available measures. *Eur. J. Neurol.* **2017**, *24*, 1099–1107. [CrossRef]