ORIGINAL ARTICLE

The Patient-Determined Disease Steps scale is not interchangeable with the Expanded Disease Status Scale in mild to moderate multiple sclerosis

Yi Chao Foong[1,2,3,4] | Daniel Merlo[1,3] | Melissa Gresle[1,2,5] | Chao Zhu[1] |
Katherine Buzzard[3,5] | Jeannette Lechner-Scott[6,7] | Michael Barnett[8,9] | Bruce Taylor[10] |
Tomas Kalincik[11,12] | Trevor Kilpatrick[12,13] | David Darby[1,2,3] | Pamela Dobay[14] |
Johan van Beek[14] | Robert Hyde[14] | Helmut Butzkueven[1,2] | Anneke van der Walt[1,2]

[1]Department of Neuroscience, Central Clinical School, Monash University, Melbourne, Victoria, Australia

[2]Alfred Health, Melbourne, Victoria, Australia

[3]Eastern Health, Melbourne, Victoria, Australia

[4]Royal Hobart Hospital, Hobart, Tasmania, Australia

[5]Melbourne Health, Melbourne, Victoria, Australia

[6]The University of Newcastle, Newcastle, New South Wales, Australia

[7]Hunter New England Health, Newcastle, New South Wales, Australia

[8]Brain and Mind Centre, The University of Sydney, Sydney, New South Wales, Australia

[9]Sydney Neuroimaging Analysis Centre, Camperdown, New South Wales, Australia

[10]Menzies Institute for Medical Research, University of Tasmania, Hobart, Tasmania, Australia

[11]Department of Medicine, CORe, University of Melbourne, Melbourne, Victoria, Australia

[12]Department of Neurology, Neuroimmunology Centre, Royal Melbourne Hospital, Melbourne, Victoria, Australia

[13]Florey Department of Neuroscience and Mental Health, University of Melbourne, Melbourne, Victoria, Australia

[14]Biogen International GmbH, Zug, Switzerland

Correspondence
Yi Chao Foong, Department of Neuroscience, Central Clinical School, Monash University, Melbourne, VIC, Australia.
Email: yi.foong@monash.edu

Abstract

Background and purpose: The validity, reliability, and longitudinal performance of the Patient-Determined Disease Steps (PDDS) scale is unknown in people with multiple sclerosis (MS) with mild to moderate disability. We aimed to examine the psychometric properties and longitudinal performance of the PDDS.

Methods: We included relapsing–remitting MS patients with an Expanded Disability Status Scale (EDSS) score of less than 4. Validity and test–retest reliability was examined. Longitudinal data were analysed with mixed-effect modelling and Cohen's kappa for concordance in confirmed disability progression (CDP).

Results: We recruited a total of 1093 participants, of whom 904 had complete baseline data. The baseline correlation between PDDS and EDSS was weak ($\rho = 0.45$, $p < 0.001$). PDDS had stronger correlations with patient-reported outcomes (PROs). Conversely,

EDSS had stronger correlations with age, disease duration, Kurtzke's functional systems and processing speed test. PDDS test–retest reliability was good to excellent (concordance correlation coefficient=0.73–0.89). Longitudinally, PDDS was associated with EDSS, age and depression. A higher EDSS score was associated with greater PDSS progression. The magnitude of these associations was small. There was no concordance in CDP as assessed by PDDS and EDSS.

Conclusion: The PDDS has greater correlation with other PROs but less correlation with other MS-related outcome measures compared to the EDSS. There was little correlation between PDDS and EDSS longitudinally. Our findings suggest that the PDDS scale is not interchangeable with the EDSS.

INTRODUCTION

The Expanded Disability Status Scale (EDSS) is the most commonly used measure of disability in people with multiple sclerosis (pwMS) [1]. However, the EDSS requires a trained neurologist, is time-consuming and does not reflect the subjective perspective of pwMS. Increasingly, patient-reported outcomes (PROs) are recognized as an important aspect of understanding patient needs and perspectives, particularly for a disease as heterogenous as multiple sclerosis (MS) [2]. The available PROs for disability are not specific to MS, lack validation and are not widely recognized by regulatory bodies [3]. In response to this, there has been a drive to develop and validate existing PROs that meet the needs of patients, researchers, healthcare practitioners and industry, reflected in the PROs for MS initiative that was launched in 2019 [2].

The Patient-Determined Disease Steps (PDDS) scale was adapted from the Physician-Reported Disease Steps scale as a self-reported outcome measure of disability [4]. Previous cross-sectional validation studies have shown a moderate to strong correlation between PDDS and EDSS across a range of patient populations [5–10]. This has led to many physicians replacing the EDSS with the PDDS scale, particularly in the context of increased telehealth during the COVID-19 pandemic [11]. The PDDS scale has also been introduced as a surrogate to the EDSS to assess confirmed disability worsening in some studies [12]. However, there has been concern that this correlation is driven by agreement at higher ends of the scale, where significant emphasis is placed on ambulation. There is also a lack of data on the longitudinal relationship between PDDS and EDSS and the impact of psychological comorbidities on PDDS reporting. Therefore, we aimed to (i) validate the PDDS scale in a cohort of people with mild to moderate MS; (ii) explore the longitudinal trajectory of the PDDS scale; and (iii) examine the relationship between psychological well-being and PDDS.

METHODS

Study design

We recruited adult MS participants between December 2017 and November 2021 from six tertiary MS clinics in Australia. All participants were enrolled in the Australian arm of the IMPROVE-MS study and the MSBase registry. Inclusion criteria for IMPROVE-MS were: a confirmed diagnosis of relapsing–remitting MS or diagnosis of clinically isolated syndrome with evidence of lesions on magnetic resonance imaging that meet the Paty A or Paty B criteria [13]; age >18 years; and an EDSS score of less than 4. There were 28 minor deviations from the inclusion criteria where EDSS score was 4 or above, mostly due to delays in baseline EDSS assessment because of restrictions in attending clinics in person due to COVID-19. These participants were allowed to continue in this study. For MSBase, all patients with MS who were attending a participating site and had provided informed consent were eligible. All assessments (EDSS, PROs, the multiple sclerosis performance test [MSPT] and PDDS) were performed during routine clinic visits (approximately 6-monthly). Clinical data were extracted from the MSBase registry. Kurtzke's functional system (FS) scores and EDSS scores were recorded in MSBase by Neurostatus-certified investigators. This study was approved by the Melbourne Health Human Research Ethics Committee and all participants provided informed consent prior to any data collection.

Multiple sclerosis performance test

The MSPT is an iPad®-based disability assessment tool that includes the processing speed test (PST), manual dexterity test (MDT) and walking speed test (WST), the contrast sensitivity test (not used in this study), Neuro-Qol (not used in this study) and a questionnaire called My Health that gathers patient demographics, MS history, treatment information and PDDS score [14].

In the PST, the iPad displays a symbol key which contains nine symbols with corresponding numbers. After a practice test, participants were presented with a series of rows of 15 symbols and instructed to select corresponding numbers. Once a row was completed, a new row of symbols was presented. The total duration of the test was 2 min. The total number of correct responses was recorded.

In the MDT, participants transferred nine pegs from the starting row into a grid of nine holes as quickly as possible and then, without pausing, removed the pegs one at a time and returned them to the starting row using their dominant hand before touching a tablet screen upon task completion. The mean time is recorded in seconds. This was repeated for the non-dominant hand.

The WST was used to assess walking speed. Participants were asked to walk 25 feet as quickly and safely as possible. The main outcome measure was the mean time in seconds to complete two trials of the WST [15].

Patient-Determined Disease Steps

Participants selected one of nine items presented on the MSPT application that best described their current level of disability. These items are: 0 – normal; 1 – mild disability; 2 – moderate disability; 3 – gait disability; 4 – early cane; 5 – late cane; 6 – bilateral support; 7 – wheelchair/scooter; and 8 – bedridden.

Patient-reported outcomes

We used three PROs, namely: the Multiple Sclerosis International Quality of Life Questionnaire (MusiQoL), the Penn State Worry Questionnaire (PSWQ) and the Patient Health Questionnaire (PHQ-9).

The MusiQoL measures health-related quality of life in pwMS. It consists of 31 items related to nine domains of quality of life rated on a scale of 1 to 5 [16]. The nine dimensions are: activities of daily living (ADL); coping (COP); psychological well-being (PWB); rejection (REJ); relationship with family (RFA); relationship with friends (RFR); relationship with healthcare system (RHCS); symptoms (SPT); and sentimental and sexual life (SSL).

The PSWQ has 16 items that measure tendency to worry on a scale of 1 to 5 [17]. The PHQ-9 is a nine-item measure of depressive symptoms in the past 2 weeks, with a scale of 0 to 3 [18].

Higher scores on the PSWQ and PHQ-9 indicate a worse tendency to worry and greater depression, respectively, whilst higher scores on the MusiQoL indicate better quality of life. All three have been validated in pwMS [19–21]. Electronic versions of these PROs were administered electronically using an iPad, as previously described [22].

Statistical analysis

The baseline characteristics of the population were reported as median (interquartile range [IQR]) or mean (standard deviation) for continuous variables and number (percentile) for discrete variables, respectively.

To examine concurrent validity of the PDDS scale, we examined correlations between the PDDS and MS-related outcomes. The correlation between PDDS and EDSS was examined with the Spearman rank correlation (ρ) test, with ρ values < 0.3 indicating very weak, $0.3 \leq \rho < 0.5$ indicating weak, $0.5 \leq \rho < 0.7$ indicating moderate, $0.7 \leq \rho < 0.9$ indicating strong, and $\rho \geq 0.9$ indicating very strong correlations [23]. In situations where the EDSS was not administered on the same day as the PDDS scale, the closest EDSS score within the 3 months before or after the PDDS reporting date was used as the corresponding EDSS score. The strength of the agreement between PDDS and EDSS was measured by weighted Cohen's kappa coefficient, with Cohen's kappa values < 0.2 indicating no agreement, ≥ 0.2 to < 0.4 minimal agreement, ≥ 0.4 to < 0.6 weak agreement, ≥ 0.6 to < 0.8 moderate agreement, and ≥ 0.8 strong agreement, respectively [24].

Lin's concordance correlation coefficient (CCC) was calculated to examine test–retest reliability. CCC values of < 0.5, ≥ 0.5 to < 0.75, ≥ 0.75 to < 0.9, and ≥ 0.9 were interpreted as poor, moderate, good and excellent reliability, respectively [25]. Given the long intra-test interval, we performed further sensitivity analysis restricting the CCC analysis to those without EDSS change between visits [26]. Fisher's z-test was used to compare the correlation coefficients.

We used a linear mixed-effects model with PDDS as the dependent variable and EDSS as the independent variable to test their relationship over time. A spaghetti plot with a fitted line was used to visualize the relationship. The covariates of baseline age, baseline disease duration, gender, EDSS and the interaction term between EDSS and visit number were evaluated in the multivariable models. A patient-level random effect with random intercepts was used.

The above process was repeated with participants that had questionnaire data. Not all participants had questionnaire data due to COVID-19 visit restrictions and contact time restrictions. Participating centres were asked to prioritize the MSPT over questionnaires where time restrictions were present. In addition to the aforementioned covariates, MusiQoL, PSWQ and PHQ total scores were added to the multivariable model. Interaction terms between MusiQoL, PSWQ, PHQ and visit number were also included.

In our study, confirmed disability progression (CDP) was defined as 6-month sustained increase in PDDS of greater than or equal to 1 [12], or 6-month sustained increase in EDSS score of greater or equal to 1.5 if baseline EDSS score was 0, greater than or equal to 1 if baseline EDSS score was 1 to 5.5, and greater than or equal to 0.5 if baseline EDSS score was 6 or more [27]. Weighted Cohen's kappa was used to examine the concordance between CDP as assessed by the EDSS and the PDDS scale. Further sensitivity analyses with different cut-off points for CDP were performed, including 3-month sustained EDSS/PDDS change, 1-point 6-month sustained EDSS change and unconfirmed disability progression without a minimum intra-test time period. Confirmed disability improvement (CDI) was defined as a 6-month sustained decrease in PDDS score of greater

than or equal to 1, or a 6-month sustained decrease in EDSS score of greater than or equal to 1 if baseline EDSS score was between 2 and 5.5, and greater than or equal to 0.5 if baseline EDSS was 6 or more.

A p value < 0.05 was taken to indicate statistical significance. All statistical analyses were performed using R 4.1.2.

RESULTS

Participant characteristics

Participant characteristics are presented in Table 1. In total, at the time of this analysis 1093 participants were consented. Of these, 983 had baseline PDDS data, but only 904 participants had corresponding EDSS data at baseline (Figure S1). Of the 904, 688 (76.1%) were female and 804 (88.9%) were right-handed.

The mean number of visits where PDDS was recorded was 2.7, with 763 participants having two visits, 506 having three, 300 having four, 156 having five and 59 having six where PDDS was recorded. The median time between visits is presented in Table 4, with an increase in time between Visit 3 and Visit 4 due to the COVID-19 pandemic. The median (IQR) PDDS score was 0 (0–1) and the median (IQR) EDSS score was 1.5 (1–2). The full distribution of PDDS and EDSS scores is shown in Tables S1 and S2. At baseline, the median

TABLE 1 Participant characteristics at baseline ($n = 904$ unless otherwise specified).

	Median (IQR)
Age, years	40.9 (33.7–48.9)
Disease duration, years	7.6 (3.8–11.7)
Visit number	2 (1–4)
Female, n (%)	688 (76.1)
EDSS score	1.5 (1–2)
PDDS score	0 (0–1)
MDT (dominant hand; $n = 870$), s	22.5 (20.2–25.5)
MDT (non-dominant hand; $n = 870$), s	23.6 (21.2–26.6)
WST ($n = 837$), s	5.8 (5.0–6.8)
PST ($n = 879$)	54 (47–60)
MusiQoL index (range 0–100; $n = 736$)	78.2 (69.4–86.4)
MusiQoL psychological score (range 0–100; $n = 736$)	80 (65–95)
PSWQ score (range 16–80; $n = 725$)	40 (27–52)
PHQ score (range 0–27; $n = 732$)	5 (2–9)

Abbreviations: EDSS, Expanded Disability Status Scale; IQR, interquartile range; MDT, manual dexterity test; MusiQoL, Multiple Sclerosis International Quality of Life questionnaire; PDDS, Patient-Determined Disease Steps; PHQ, Patient Health Questionnaire; PST, processing speed test; PSWQ, Penn State Worry Questionnaire; WST, walking speed test.
Note: Higher scores on the PSWQ and PHQ are indicative of greater worry and depression, respectively. Lower scores on the MusiQoL index or subscores are indicative of worse quality of life.

(IQR) age was 40.9 (33.7–48.9) years and the median (IQR) disease duration was 7.6 (3.8–11.7) years.

Baseline correlation between EDSS and PDDS

Baseline correlation between EDSS and PDDS was weak (Spearman's $\rho = 0.45$, $p < 0.001$) (Figure 1). As shown in Figure 1, there was little correlation between EDSS and corresponding PDDS scores.

Concurrent validity with functional system subscores

The correlation coefficients and 95% confidence intervals (CIs) between PDDS and EDSS with FS scores from the EDSS are provided in Table 2. PDDS scores were weakly correlated with FS scores except for bowel/bladder subscores, for which there was no significant association. Correlations between FS scores and EDSS were weak to moderate, and were very weak to weak for PDDS. Correlation coefficients were significantly stronger for the EDSS compared to the PDDS for all FS subscores.

Concurrent validity with other outcome variables

The correlation coefficients and 95% CIs between PDDS and EDSS with age, disease duration, WST, MDT, PST, PHQ, MusiQoL and PSWQ are presented in Table 3. PDDS and EDSS were very weakly correlated with age and disease duration, although correlations with the EDSS were significantly stronger. For WST and MDT, correlations were very weak for EDSS and PDDS, with no significant difference. For PST, correlation with the PDDS was very weak and was weak with the EDSS.

For PROs, PDDS was correlated weakly with the PHQ, and very weakly with PSWQ and MusiQoL. EDSS correlated very weakly with all three questionnaires. For all three PROs, correlations with the PDDS were significantly stronger than with the EDSS.

Figure 2 shows the correlation coefficients for EDSS and PDSS with MusiQoL subscores. For the EDSS, correlations were minimal to weak, with ADL and RFR having the strongest association. For PDDS, correlations ranged from minimal to moderate, with ADL and RFR again having the strongest association. Correlations were significantly stronger with the PDDS for the overall MusiQoL index and ADL, PWB, RFR and SSL subscores.

Test–retest reliability

The test–retest reproducibility of PDDS scores was examined between each visit (Table 4). Overall, test–retest reliability was good to excellent, with an increasing trend with each visit. We performed a sensitivity analysis restricting the analysis to those without EDSS change between visits and found similar results (Table 4).

FIGURE 1 Weighted scatter plot of Patient-Determined Disease Steps (PDDS) and Expanded Disability Status Scale (EDSS) at baseline ($n = 904$). Size of dots and shading correspond with number of participants at a corresponding PDDS and EDSS score, with larger dots and lighter blue representing higher numbers.
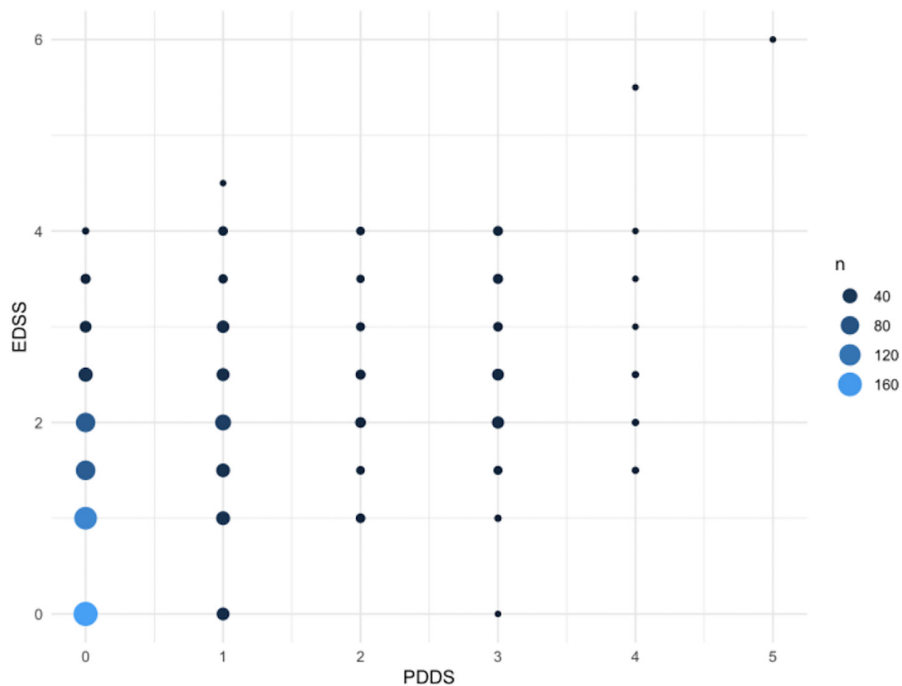

TABLE 2 Spearman's correlations between Patient-Determined Disease Steps, Expanded Disability Status Scale and functional system subscores.

FS	PDDS	PDDS 95% CI	p value	EDSS	EDSS 95% CI	p-Value	p value of Fisher's z-test	Total n
Visual	0.34	0.29–0.42	<0.001	0.62	0.60–0.68	<0.001	<0.001	811
Brainstem	0.35	0.30–0.42	<0.001	0.57	0.54–0.63	<0.001	<0.001	811
Pyramidal	0.23	0.16–0.29	<0.001	0.34	0.29–0.41	<0.001	<0.001	811
Cerebellar	0.32	0.26–0.38	<0.001	0.57	0.51–0.61	<0.001	<0.001	810
Sensory	0.32	0.25–0.38	<0.001	0.53	0.48–0.58	<0.001	<0.001	809
Bladder/Bowel	0.05	−0.01 to 0.12	0.16	0.31	0.26–0.38	<0.001	<0.001	806
Mental	0.30	0.24–0.37	<0.001	0.42	0.38–0.49	<0.001	0.0009	811

Abbreviations: CI, confidence interval; EDSS, Expanded Disability Status Scale; FS, functional system; PDDS, Patient-Determined Disease Steps.

Longitudinal analysis

Initial univariable analyses showed that EDSS score increased by 0.05 (95% CI 0.03–0.07; $p < 0.001$) and PDDS score increased by 0.06 (95% CI 0.03–0.09; $p < 0.001$) per visit. Figures 3 and 4 shows the respective spaghetti plots for EDSS and PDDS trajectories in this study cohort. In the final model, EDSS ($\beta = 0.26$, 95% CI 0.20–0.32; $p < 0.001$) and age ($\beta = 0.01$, 95% CI 0.00–0.01, $p < 0.001$) were independently associated with PDDS (Table S3). The interaction term between EDSS and visit number was significant ($\beta = 0.03$, 95% CI 0.01–0.06; $p < 0.01$), signifying that, with increasing EDSS score, there was greater increase in the PDDS score over time. The amount of variance in the model explained by individual variation was 64.8% (0.57/0.88).

We hypothesized that MusiQoL scores, depression (as assessed by PHQ) and anxiety (as assessed by PSWQ) would also affect PDDS longitudinally. The total number of observations were smaller in this analysis given not all had questionnaire data (1702 observations vs. 2440 in the previous analysis). In the final model (Table S4), in addition to age and EDSS, PHQ total scores ($\beta = 0.04$, 95% CI 0.02–0.06, $p < 0.001$) were significantly associated with PDDS. However, when an interaction term between visit number and MusiQoL/PSWQ/PHQ score was added, none of them was significantly associated with PDDS. This signifies that MusiQoL, PSWQ and PHQ scores do not affect the trajectory of PDDS over time, unlike EDSS.

Longitudinal concordance between confirmed disability progression as assessed by EDSS and PDDS

There were 33 CDP events as assessed by EDSS (4.1%, 33/797) and 69 as assessed by PDDS (8.7%, 69/797), out of a possible 797 (Table S5). Concordance between 6-month CDP as assessed by PDDS and EDSS was negligible (weighted Cohen's kappa = 0.09, 95% CI −0.01 to 0.18). Further sensitivity analyses were performed with 3-month CDP (weighted Cohen's kappa = 0.10, 95% CI 0.00 to 0.19), unconfirmed disability progression without a minimum intra-test

TABLE 3 Spearman's correlations between Patient-Determined Disease Steps, Expanded Disability Status Scale and age, disease duration, and walking speed test, processing speed test, manual dexterity test, Patient Health Questionnaire, Multiple Sclerosis International Quality of Life questionnaire and Penn State Worry Questionnaire scores.

Variable	PDDS	PDDS 95% CI	p value	EDSS	EDSS 95% CI	p value	p value of Fisher's z-test	n
Age	0.18	0.12–0.25	<0.001	0.29	0.23–0.35	<0.001	<0.001	904
Disease duration	0.10	0.05–0.18	<0.001	0.18	0.13–0.25	<0.001	0.002	904
WST	0.20	0.16–0.29	<0.001	0.20	0.16–0.28	<0.001	0.83	837
MDT (right)	0.20	0.14–0.27	<0.001	0.25	0.19–0.32	<0.001	0.27	870
MDT (left)	0.20	0.14–0.27	<0.001	0.28	0.22–0.34	<0.001	0.06	870
MDT (dominant)	0.20	0.14–0.27	<0.001	0.26	0.20–0.32	<0.001	0.17	870
MDT (non-dominant)	0.20	0.14–0.27	<0.001	0.28	0.22–0.34	<0.001	0.10	870
PST	−0.18	−0.12 to −0.25	<0.001	−0.30	−0.25 to −0.37	<0.001	<0.001	879
PHQ	0.36	0.30–0.42	<0.001	0.29	0.23–0.36	<0.001	0.009	732
MusiQoL	−0.29	−0.21 to −0.34	<0.001	−0.20	−0.14 to −0.28	<0.001	<0.001	736
PSWQ	0.17	0.10–0.25	<0.001	0.07	0.01–0.15	0.06	0.01	725

Abbreviations: CI, confidence interval; EDSS, Expanded Disability Status Scale; MDT, manual dexterity test; MusiQoL, Multiple Sclerosis International Quality of Life questionnaire; PDDS, Patient-Determined Disease Steps; PHQ, Patient Health Questionnaire; PST, processing speed test; PSWQ, Penn State Worry Questionnaire; WST, walking speed test.
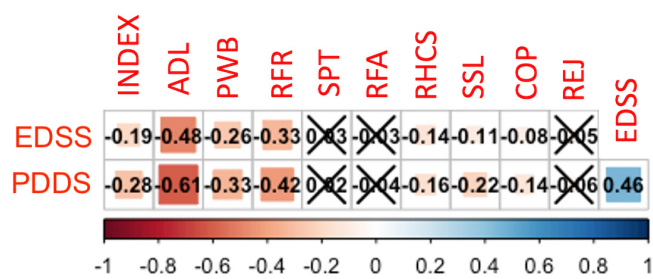

FIGURE 2 Spearman's correlations between Multiple Sclerosis International Quality of Life questionnaire (MusiQoL) subscores and Patient-Determined Disease Steps (PDDS)/ Expanded Disability Status Scale (EDSS) scores. Heat map indicates strength of the correlation (rho). Crosses (for symptoms [SPT], relationship with family [RFA] and rejection [REJ]) indicate nonsignificant results ($p > 0.05$). ADL, activities of daily living; COP, coping; INDEX, MusiQoL index score; PWB, psychological well-being; RFR, relationship with friends; RHCS, relationship with healthcare system; SSL, sentimental and sexual life.

time period (weighted Cohen's kappa = 0.13, 95% CI 0.06 to 0.19), and defining CDP on EDSS as a 6-month sustained 1-point change (weighted Cohen's kappa = 0.07, 95% CI −0.02 to 0.15). Concordance between the two measures remained negligible. There were 24 CDI events as assessed by EDSS (3.0%, 24/797) and 54 as assessed by PDDS (6.8%, 54/797), out of a possible 797. Concordance for CDI was negligible (weighted Cohen's kappa = 0.01, 95% CI −0.06 to 0.08).

DISCUSSION

This large longitudinal study of people with mild to moderate MS demonstrates that the PDDS should not be used interchangeably with the EDSS as a measure of disability. Compared to the EDSS, correlations between the PDDS and other PROs were stronger, but correlations with assessor-based and performance-based outcomes were weaker. Consistent with previous reports, our findings support the test–retest reliability of the PDDS. Finally, we have shown a longitudinal association between PDDS and EDSS and psychological well-being, and that those with greater disability as assessed by EDSS have greater progression on the PDDS. However, the magnitude of the association was small, and there was little concordance in CDP as defined by PDDS and EDSS. This has implications for the use of the PDDS in clinical settings and as a trial outcome. We suggest that patient-reported disability as assessed by the PDDS is still a valuable outcome measure but should be considered complementary to and not replace clinician-reported or performance-based disability measures.

In our study, the correlation between PDDS and EDSS was weaker compared to previously published data [8–10]. This is likely explained by the lower level of neurological impairment in our cohort, with few participants reporting gait impairment. Given the greater weight placed on ambulation in both the PDDS and EDSS at the higher ends of the scales, there is a greater agreement in pwMS with gait disability [8].

Correlations with FS subscores were weaker across all FSs for the PDDS compared to the EDSS. In the original validation study by Learmonth et al. [10], there was no significant difference in correlations between PDDS, EDSS and most FS subscores. However, that study was smaller in size, and the cohort participants had a greater disability level. Given that FS subscores determine the final EDSS step in mild to moderate MS, this was not unexpected. We found that bowel/bladder FS was not correlated with PDDS, probably because of the lack of specific assessment in PDDS. Given reported prevalence rates of 40%–70% and 30%–50% for bladder and bowel dysfunction, respectively, this is likely to represent significant

TABLE 4 Lin's concordance correlation coefficients for the Patient-Determined Disease Steps scale in between visits.

Visit	Lin's CCC (95% CI)	N	Time between visits, median, (IQR) days	Lin's CCC for those without EDSS score change (95% CI)	N without EDSS score change
1–2	0.73 (0.70–0.76)	763	223 (182–399)	0.74 (0.69–0.78)	382
2–3	0.78 (0.74–0.81)	506	231 (182–378)	0.82 (0.78–0.86)	216
3–4	0.79 (0.74–0.83)	300	272 (182–392)	0.76 (0.68–0.83)	126
4–5	0.85 (0.79–0.88)	156	228 (182–365)	0.89 (0.83–0.93)	64
5–6	0.89 (0.82–0.93)	59	217 (182–364)	0.91 (0.80–0.96)	22

Abbreviations: CCC, concordance correlation coefficient; CI, confidence interval; EDSS, Expanded Disability Status Scale; IQR, interquartile range.
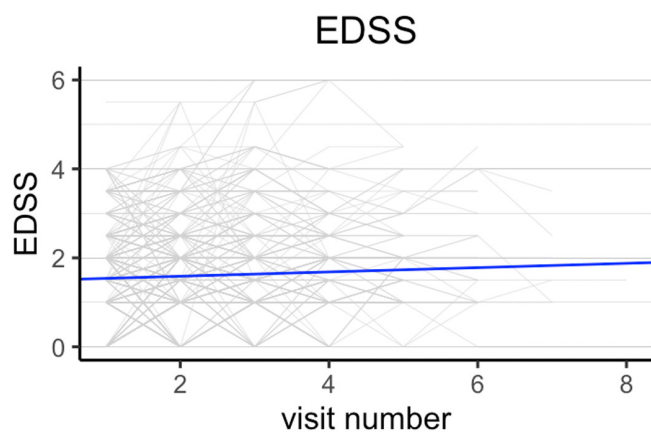
FIGURE 3 Spaghetti plot of Expanded Disability Status Scale (EDSS) trajectory over time with line of best fit. Line of best fit was obtained from the linear mixed-effect model.
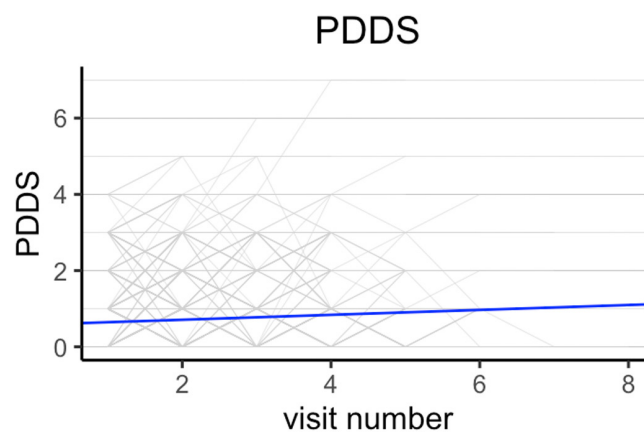
FIGURE 4 Spaghetti plot of Patient-Determined Disease Steps (PDDS) trajectory over time with line of best fit. Line of best fit was obtained from the linear mixed-effect model.

under-reporting [28–30]. Gustavsen et al. reported a prevalence of mild to moderate bladder and bowel dysfunction in up to 40% of those who had a score of 0 on the PDDS scale [31]. The lack of detection of this and other commonly under-reported symptoms may partially explain the discordance between the PDDS and EDSS in people with mild levels of disability. This may also signify that PDDS is less sensitive for the detection of progression independent of relapse activity, which has recently been shown often to manifest as bladder/bowel dysfunction in the Italian MS registry [32]. Future studies may examine whether adding specific bowel/bladder questions may improve correlations between EDSS and PDDS. The PDDS scale could also be compared against other PROs that address this issue specifically, such as the MS Symptom Scores [33].

For demographic, ambulatory, upper limb and cognitive measures (age, disease duration, WST, MDT, PST) the overall correlation with both EDSS and PDDS was weak. The degree of correlation was weaker in our study compared to previous validation studies [7, 9, 10, 34]. Again, this is likely due to the lower levels of disability in our cohort. Age, disease duration and PST were more strongly correlated to EDSS. De David et al. [8] and Kahraman et al. [9] had found a similar pattern with regards to the symbol digit modalities test (SDMT), however in their study, the Fisher's z-test did not reach significance due to their smaller sample size.

Correlations between the PDDS scale and PROs (PHQ, MusiQoL and PSWQ) were stronger compared to the EDSS. For MusiQoL,

correlation with the PDDS scale was stronger for four out of the nine subscores compared to the EDSS. ADL subscores had the strongest correlation with PDDS. This is consistent with other studies showing stronger correlations between PROs compared to performance-based or assessor-based outcome measures in pwMS [22, 35].

Test–retest reliability data for the PDDS scale are relatively scarce. In our study, test–retest reliability was lower compared to previous reports, which have generally reported test–retest reliability in the excellent range [9, 34]. This is likely due to the longer duration in between test and retest, where true change in neurological status may have occurred. As a result, we performed a sensitivity analysis restricting the analysis to those with unchanged EDSS score, which found similar results. Overall, our results support the test–retest reliability of the PDDS scale. It should be noted that durations between visits were longer for the first three visits due to the COVID-19 pandemic.

There has been little literature to date on the longitudinal trajectories of the PDDS. Our longitudinal data analysis showed that age and EDSS were associated longitudinally with PDDS. Higher levels of disability (as signified by EDSS) were associated with more rapid progression in the PDDS. This lends further weight to the validity of PDDS as a measure of disability. However, the magnitude of the association between the PDDS and EDSS was small. This supports the notion that patient-reported disability and clinician-assessed disability are not interchangeable and may be explained by

the underestimation of various MS-related symptoms by the PDDS. Overall, depression (as assessed by the PHQ), but not anxiety (assessed by the PSWQ) or MusiQoL, had a longitudinal association with PDDS, although it did not affect the trajectory of the PDDS. This is consistent with emerging literature that suggests that depressive symptoms have a greater impact on functional outcomes compared to anxiety [36].

Furthermore, we found that there was little concordance between CDP and CDI as assessed by PDDS and EDSS. This has significant implications for the use of PDDS as a marker of disability in clinical trials [12]. Our study suggests that, whilst the PDDS is a valid reflection of patient-reported disability, this is a different construct compared to the EDSS, which measures clinician-reported disability. Whilst they are both important outcome measures to consider, our findings suggest that they should not be used interchangeably. MS research has historically focused on physician-assessed disability, which inadequately reflects the perspective of pwMS. For example, previous studies have shown that those with low EDSS scores can still experience significant productivity loss, underemployment and a heavy burden of 'invisible' symptoms [37 38]. Incorporating patient-reported disability may facilitate shared decision making, improved symptom control and patient satisfaction [3 39]. Future studies may compare the PDDS against the EDSS with regard to its longitudinal association with other biomarkers of MS progression, ecological validity (such as employment) and symptom burden.

Limitations of this study include longer than usual time between clinic visits (usually 6 months) due to the COVID-19 pandemic. Out of the 2440 visits in total with corresponding PDDS and EDSS scores, 12 were telephone-based EDSS assessments. Omitting these from the analysis made no significant difference to our results (data not shown). Finally, the longer duration between test and retest means that a true change in neurological status may have occurred, leading to a lower concordance coefficient. However, this longer interval mimics routine clinical practice and has been used in recent validation studies of other MS outcome variables. The similar reliability metrics found in our sensitivity analysis also emphasize the reliability of the PDDS.

In conclusion, this study shows that the PDDS scale should not be used interchangeably with the EDSS in mild to moderate MS. Compared to the EDSS, it correlates better with other subjective outcome measures but less with MS-related objective outcome measures. Longitudinal associations between PDDS and EDSS and depression were shown. However, the magnitude of these associations was small, and there was no concordance between CDP as assessed by EDSS and PDDS. Future studies may examine the effect of modifying the PDDS to better detect under-reported symptoms such as bladder and bowel dysfunction.

AUTHOR CONTRIBUTIONS

Yi Chao Foong: Conceptualization; methodology; writing—review and editing; writing—original draft; formal analysis. **Daniel Merlo:** Writing—review and editing; supervision. **Melissa Gresle:** Writing—review and editing; supervision; project administration; methodology; conceptualization. **Chao Zhu:** Writing—review and editing; supervision; formal analysis. **Katherine Buzzard:** Writing—review and editing; supervision. **Jeannette Lechner-Scott:** Writing—review and editing; methodology; conceptualization; investigation; funding acquisition; project administration; data curation. **Michael Barnett:** Data curation; project administration; writing—review and editing; methodology; conceptualization; investigation; funding acquisition. **Bruce Taylor:** Conceptualization; investigation; funding acquisition; writing—review and editing; methodology; data curation; project administration. **Tomas Kalincik:** Project administration; data curation; methodology; writing—review and editing; conceptualization; investigation; funding acquisition. **Trevor Kilpatrick:** Conceptualization; investigation; funding acquisition; methodology; writing—review and editing; project administration; data curation. **David Darby:** Project administration; writing—review and editing; methodology; conceptualization; investigation; funding acquisition. **Pamela Dobay:** Project administration; writing—review and editing; methodology; conceptualization; investigation; funding acquisition. **Johan van Beek:** Conceptualization; investigation; funding acquisition; methodology; writing—review and editing; project administration. **Robert Hyde:** Project administration; writing—review and editing; methodology; conceptualization; investigation; funding acquisition. **Helmut Butzkueven:** Conceptualization; investigation; funding acquisition; writing—review and editing; formal analysis; supervision; data curation; methodology; project administration. **Anneke van der Walt:** Methodology; writing—review and editing; conceptualization; investigation; funding acquisition; data curation; supervision; formal analysis; project administration.

FUNDING INFORMATION

ACKNOWLEDGEMENTS

CONFLICT OF INTEREST STATEMENT

Yi Chao Foong received travel compensation from Biogen. He also receives research funding support from the NHMRC, Multiple Sclerosis Research Australia and ANZAN. Melissa Gresle is currently working on observational studies funded by Biogen and Roche. Daniel Merlo has received honoraria from Novartis. Katherine Buzzard has received honoraria for presentations and/or educational support from Biogen, Sanofi Genzyme, Merck, Roche, Alexion and Teva. She serves on medical advisory boards for Merck and Biogen. Jeannette Lechner-Scott received travel compensation from Novartis, Biogen, Roche and Merck. Her institution receives the honoraria for talks and advisory board commitment as well as research grants from Biogen, Merck,

DATA AVAILABILITY STATEMENT

ORCID

Yi Chao Foong https://orcid.org/0000-0001-6447-078X
Tomas Kalincik https://orcid.org/0000-0003-3778-1376

REFERENCES

1. Kurtzke JF. Rating neurologic impairment in multiple sclerosis: an expanded disability status scale (EDSS). *Neurology*. 1983;33(11):1444-1452.
2. Zaratin P, Vermersch P, Amato MP, et al. The agenda of the global patient reported outcomes for multiple sclerosis (PROMS) initiative: progresses and open questions. *Mult Scler Relat Disord*. 2022;61:103757.
3. Brichetto G, Zaratin P. Measuring outcomes that matter most to people with multiple sclerosis: the role of patient-reported outcomes. *Curr Opin Neurol*. 2020;33(3):295-299.
4. Hohol M, Orav E, Weiner H. Disease steps in multiple sclerosis: a longitudinal study comparing disease steps and EDSS to evaluate disease progression. *Mult Scler J*. 1999;5(5):349-354.
5. Solà-Valls N, Vicente-Pascual M, Blanco Y, et al. Spanish validation of the telephone assessed expanded disability status scale and patient determined disease steps in people with multiple sclerosis. *Mult Scler Relat Disord*. 2019;27:333-339.
6. Ann Marrie R, McFadyen C, Yaeger L, Salter A. A systematic review of the validity and reliability of the patient-determined disease steps scale. *Int J MS Care*. 2022;25:20-25.
7. Aldughmi M, Al-Shorman A, Khalil H, El-Salem K, Alghwiri A. Translation and validation of the Arabic version of the patient determined disease steps in people with multiple sclerosis. *Physiother Theory Pract*. 2022;38(9):1281-1288.
8. de David AC, Sasaki JE, Ramari C, et al. Validation of the Brazilian version of the patient-determined disease steps scale in persons with multiple sclerosis. *Mult Scler Relat Disord*. 2019;30:208-214.
9. Kahraman T, Özdoğar AT, Özakbaş S. Cross-cultural adaptation, validity and reliability of the Turkish version of the patient determined disease steps scale in persons with multiple sclerosis. *Physiother Theory Pract*. 2021;37(4):527-534.
10. Learmonth YC, Motl RW, Sandroff BM, Pula JH, Cadavid D. Validation of patient determined disease steps (PDDS) scale scores in persons with multiple sclerosis. *BMC Neurol*. 2013;13(1):1-8.
11. Alonso R, Carvajal R, Boaventura M, Galleguillos L. Experience of south American MS and/or NMOSD experts in practice during the COVID-19 pandemic: focus on telemedicine. *Mult Scler Relat Disord*. 2021;48:102702. doi:10.1016/j.msard.2020.102702
12. Salter A, Lancia S, Cutter G, et al. A propensity-matched comparison of long-term disability worsening in patients with multiple sclerosis treated with dimethyl fumarate or fingolimod. *Ther Adv Neurol Disord*. 2021;14:17562864211021177.
13. Paty D, Oger J, Kastrukoff L, et al. MRI in the diagnosis of MS: a prospective study with comparison of clinical evaluation, evoked potentials, oligoclonal banding, and CT. *Neurology*. 1988;38(2):180-185.
14. Rudick RA, Miller D, Bethoux F, et al. The multiple sclerosis performance test (MSPT): an iPad-based disability assessment tool. *J Vis Exp*. 2014;(88):e51318. doi:10.3791/51318
15. Kieseier BC, Pozzilli C. Assessing walking disability in multiple sclerosis. *Mult Scler J*. 2012;18(7):914-924.
16. Fernández O, Baumstarck-Barrau K, Simeoni M-C, Auquier P, Group MS. Patient characteristics and determinants of quality of life in an international population with multiple sclerosis: assessment using the MusiQoL and SF-36 questionnaires. *Mult Scler J*. 2011;17(10):1238-1249.
17. Meyer TJ, Miller ML, Metzger RL, Borkovec TD. Development and validation of the penn state worry questionnaire. *Behav Res Ther*. 1990;28(6):487-495.
18. Marrie RA, Zhang L, Lix LM, et al. The validity and reliability of screening measures for depression and anxiety disorders in multiple sclerosis. *Mult Scler Relat Disord*. 2018;20:9-15.
19. Patrick S, Connick P. Psychometric properties of the PHQ-9 depression scale in people with multiple sclerosis: a systematic review. *PloS One*. 2019;14(2):e0197943. doi:10.1371/journal.pone.0197943

20. Thornton EW, Tedman S, Rigby S, Bashforth H, Young C. Worries and concerns of patients with multiple sclerosis: development of an assessment scale. *Mult Scler J.* 2006;12(2):196-203. doi:10.1191/135248506ms1273oa

21. Simeoni M, Auquier P, Fernandez O, et al. Validation of the multiple sclerosis international quality of life questionnaire. *Mult Scler J.* 2008;14(2):219-230.

22. Merlo D, Kalincik T, Zhu C, et al. Subjective versus objective performance in people with multiple sclerosis using the MSReactor computerised cognitive tests. *Mult Scler Relat Disord.* 2022;58:103393. doi:10.1016/j.msard.2021.103393

23. Mukaka MM. A guide to appropriate use of correlation coefficient in medical research. *Malawi Med J.* 2012;24(3):69-71.

24. McHugh ML. Interrater reliability: the kappa statistic. *Biochem Med (Zagreb).* 2012;22(3):276-282.

25. Koo TK, Li MY. A guideline of selecting and reporting Intraclass correlation coefficients for reliability research. *J Chiropr Med.* 2016;15(2):155-163. doi:10.1016/j.jcm.2016.02.012

26. Goldman MD, LaRocca NG, Rudick RA, et al. Evaluation of multiple sclerosis disability outcome measures using pooled clinical trial data. *Neurology.* 2019;93(21):e1921-e1931.

27. Kalincik T, Cutter G, Spelman T, et al. Defining reliable disability outcomes in multiple sclerosis. *Brain.* 2015;138(11):3287-3298.

28. Abdi F, Kashani ZA, Pakzad R, Alidost F. Urinary disorders and sexual dysfunction in patients with multiple sclerosis: a systematic review and meta-analysis. *Int J Sex Health.* 2020;32(3):312-330.

29. Nazari F, Shaygannejad V, Mohammadi Sichani M, Mansourian M, Hajhashemi V. The prevalence of lower urinary tract symptoms based on individual and clinical parameters in patients with multiple sclerosis. *BMC Neurol.* 2020;20(1):24. doi:10.1186/s12883-019-1582-1

30. Khan F, Pallant JF, Shea TL, Whishaw M. Multiple sclerosis: prevalence and factors impacting bladder and bowel function in an Australian community cohort. *Disabil Rehabil.* 2009;31(19):1567-1576. doi:10.1080/09638280802639566

31. Gustavsen S, Olsson A, Søndergaard HB, et al. The association of selected multiple sclerosis symptoms with disability and quality of life: a large Danish self-report survey. *BMC Neurol.* 2021;21(1):317. doi:10.1186/s12883-021-02344-z

32. Portaccio E, Fonderico M, Aprea M, et al. "Hidden" symptoms drive progression independent of relapse activity in relapsing-onset multiple sclerosis patients. 2022. Presented at: ECTRIMS 2022. Amsterdam, Netherlands.

33. Zhang Y, Taylor BV, Simpson S, Blizzard L, Palmer AJ, van der Mei I. Validation of 0–10 MS symptom scores in the Australian multiple sclerosis longitudinal study. *Mult Scler Relat Disord.* 2020;39:101895. doi:10.1016/j.msard.2019.101895

34. Lavorgna L, Sparaco M, Esposito S, et al. Validity and reproducibility of the Italian version of the patient determined disease steps scale in people with multiple sclerosis. *Mult Scler Relat Disord.* 2017;18:173-176.

35. Schäffler N, Schönberg P, Stephan J, Stellmann JP, Gold SM, Heesen C. Comparison of patient-reported outcome measures in multiple sclerosis. *Acta Neurol Scand.* 2013;128(2):114-121. doi:10.1111/ane.12083

36. Gill S, Santo J, Blair M, Morrow SA. Depressive symptoms are associated with more negative functional outcomes than anxiety symptoms in persons with multiple sclerosis. *J Neuropsychiatry Clin Neurosci.* 2019;31(1):37-42. doi:10.1176/appi.neuropsych.18010011

37. García-Domínguez JM, Maurino J, Martínez-Ginés ML, et al. Economic burden of multiple sclerosis in a population with low physical disability. *BMC Public Health.* 2019;19(1):1-8.

38. Coyne KS, Boscoe AN, Currie BM, Landrian AS, Wandstrat TL. Understanding drivers of employment changes in a multiple sclerosis population. *Int J MS Care.* 2015;17(5):245-252. doi:10.7224/1537-2073.2014-051

39. Fitzgerald KC, Salter A, Tyry T, et al. Validation of the SymptoMScreen with performance-based or clinician-assessed outcomes. *Mult Scler Relat Disord.* 2019;29:86-93.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Foong YC, Merlo D, Gresle M, et al. The Patient-Determined Disease Steps scale is not interchangeable with the Expanded Disease Status Scale in mild to moderate multiple sclerosis. *Eur J Neurol.* 2024;31:e16046. doi:10.1111/ene.16046