



## Original article

## Assessing life balance of European people with multiple sclerosis: A multicenter clinimetric study within the RIMS network



Daphne Kos<sup>a,b,1,\*</sup>, Sofie Ferdinand<sup>b,1</sup>, Marijke Duportail<sup>b,1</sup>, Isaline Eijssen<sup>c,1</sup>, Sofie Schouteden<sup>d,1</sup>, Lore Kerkhofs<sup>d,1</sup>, Jelka Jansa<sup>e,1</sup>, Núria Fillo<sup>f,1</sup>, Kathleen Matuska<sup>g</sup>, Heleen Beckerman<sup>c,1</sup>

<sup>a</sup> Department of Rehabilitation Sciences, KU Leuven, Tervuursevest 101, PO Box 1501, Leuven 3001, Belgium

<sup>b</sup> National MS Center, Vanheylenstraat 16, Melsbroek 1820, Belgium

<sup>c</sup> Department of Rehabilitation Medicine, Amsterdam UMC, Vrije Universiteit Amsterdam, MS Center Amsterdam, PO Box 7057, Amsterdam, MB 1007, the Netherlands

<sup>d</sup> Rehabilitation- and MS Center Overpelt, Boemerangstraat 2, Overpelt 3900, Belgium

<sup>e</sup> University Medical Centre Ljubljana (SI), Zaloska 2, Ljubljana SI-1000, Slovenia

<sup>f</sup> Cemcat (Multiple Sclerosis Center of Catalonia), PG. Vall d'Hebron 119-129, Barcelona, ES, Spain

<sup>g</sup> St Catherine University Minneapolis/Minnesota, 123 A Fontbonne Hall, Mail Stop #F-25, 2004 Randolph Ave, St. Paul, MN 55105, USA

## ARTICLE INFO

## Keywords:

Quality of life  
Rehabilitation  
Cross-cultural validation  
Patient reported outcome measure (PROM)  
Meaningful activities

## ABSTRACT

**Background:** Life balance is defined as “a satisfying pattern of daily activity that is healthful, meaningful, and sustainable to an individual within the context of his or her current life circumstances”. To assess life balance, the self-report instrument Life Balance Inventory (LBI) has been developed in the US. The aim of this study was to evaluate cross-cultural, construct validity and test-retest reliability of translated versions of the LBI in people with multiple sclerosis (MS) within different European cultures (Dutch, Flemish, Slovenian, and Spanish).

**Method:** The LBI was translated according to the principles of forward/backward translation and the cultural adaption process of patient-reported outcomes and evaluated in people with MS in each country/language area. LBI (score range 1–3; higher scores refer to better balance) was registered twice with an interval of 7 days to evaluate test-retest reliability using Intraclass Correlation Coefficients (ICCs) and Bland Altman analyses. To evaluate construct validity, Pearson correlations of the LBI with quality of life, fatigue, depression and self-efficacy were explored.

**Results:** The total sample ( $n = 313$ ,  $50 \pm 11$  years of age, MS duration  $13 \pm 8$  years) consisted of five subsamples: Dutch ( $n = 81$ , 74% women,  $54 \pm 9.6$  years of age), Flemish 1 ( $n = 42$ , 57% women,  $49 \pm 12$  years), Flemish 2 ( $n = 105$ , 63% women,  $50 \pm 10.6$  years), Slovenian ( $n = 48$ , 79% women,  $44 \pm 11.2$  years) and Spanish ( $n = 37$ , 62% women,  $47 \pm 9.0$  years). Baseline total LBI scores differed between subsamples ( $F(4, 312) = 7.19$ ,  $p < 0.001$ ). ICC [95% CI] of total LBI was 0.88 [0.83–0.92] (Flemish 2), 0.65 [0.39–0.82] (Flemish 1), 0.55 [0.37–0.69] (Dutch), 0.45 [0.15–0.67] (Spanish) and 0.35 [0.07–0.59] (Slovenian). Systematic error was present in one sample; no proportional bias occurred. Correlations ranged from 0.05 to 0.55 for quality of life and self-efficacy, from  $-0.50$  to  $0.05$  for fatigue and from  $-0.44$  to  $-0.28$  for depression, not fully supporting the hypotheses.

**Conclusion:** The study results provide limited support for test-retest reliability, cross-cultural and construct validity of the LBI in different European subsamples. Although LBI may serve as a supportive tool in goalsetting in rehabilitation, the current version of LBI is not recommended for (international) research purposes.

## List of abbreviations

ICC intraclass correlation

LBI Life Balance Inventory  
MFIS Modified Fatigue Impact Scale  
MS multiple sclerosis

\* Corresponding author at: Department of Rehabilitation Sciences, KU Leuven, Tervuursevest 101, PO Box 1501, Leuven 3001, Belgium.

E-mail addresses: [daphne.kos@kuleuven.be](mailto:daphne.kos@kuleuven.be) (D. Kos), [sofie.ferdinand@mscenter.be](mailto:sofie.ferdinand@mscenter.be) (S. Ferdinand), [marijke.duportail@mscenter.be](mailto:marijke.duportail@mscenter.be) (M. Duportail), [i.eijssen@vumc.nl](mailto:i.eijssen@vumc.nl) (I. Eijssen), [lore.kerkhofs@msreva.be](mailto:lore.kerkhofs@msreva.be) (L. Kerkhofs), [jelka.jansa@kclj.si](mailto:jelka.jansa@kclj.si) (J. Jansa), [nfillo@cem-cat.org](mailto:nfillo@cem-cat.org) (N. Fillo), [kmatuska@stkate.edu](mailto:kmatuska@stkate.edu) (K. Matuska), [h.beckerman@vumc.nl](mailto:h.beckerman@vumc.nl) (H. Beckerman).

<sup>1</sup> European Network for best practice and research - Rehabilitation in MS (RIMS; [www.eurims.org](http://www.eurims.org)), SIG Occupation group, <https://www.eurims.org/content/special-interest-groups-sigs.html#>

<https://doi.org/10.1016/j.msard.2019.101879>

Received 27 August 2019; Received in revised form 20 November 2019; Accepted 28 November 2019

2211-0348/© 2019 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

MSNQ	Multiple Sclerosis Neuropsychological Screening Questionnaire
QIDS-SR16	Quick Inventory of Depressive Symptomatology Self-Rated
SDC	smallest detectable change
SDMT	Symbol Digit Modalities Test
SEM	standard error of measurement
SF36	Short Form-36 health survey
UW-SES	University of Washington Self-Efficacy Scale

## 1. Background

Life balance is a concept concerning all human beings. A balanced person is able to spread time (and energy) optimally across the different life domains (work, leisure, self-care and rest). A satisfying pattern of daily activity is healthy, meaningful and sustainable (Matuska and Christiansen, 2008).

Multiple sclerosis (MS) is a chronic disease of inflammation, demyelination and axonal loss in the central nervous system with a prevalence of over 100 cases per 100,000 citizens in the Northern countries in Europe and US (Glaser et al., 2019). Due to several consequences of the disease, people with multiple sclerosis (MS) may encounter reduced quality of life and the ability to self-steer an optimal balance in daily activity may be compromised (Matuska and Erickson, 2008; Barin et al., 2018).

To provide a tailor-made coaching of people with MS in finding the most optimal balance, it is advisable to assess (disturbances in) life balance. The Life Balance Inventory (LBI) is one of the few instruments specifically assessing the construct of life balance and assesses the perceived congruence between desired and actual time use in various activity categories (Matuska, 2012a, 2012b). Besides, the LBI is an inventory of life balance domains, referring to the activity capital, i.e. how many activities one is or would like to be involved in.

The LBI is a self-report instrument, consisting of a list of 53 activities within different daily life domains (instrumental activities of daily living, work, rest, play, education, leisure and social participation). LBI includes four subscales: health, identity, relationship and challenge/interest.

The original (American) version of the LBI showed good reliability and acceptable validity in a general Western population (Matuska, 2012a, 2012b). To date, limited clinimetric evidence of the LBI is available in samples of people with MS or in other cultures or language areas (Larivière and Levasseur, 2016; Wagman et al., 2015).

The aim of this study is to evaluate test-retest reliability, cross-cultural and construct validity of the translated versions of the Life Balance Inventory (LBI) in people with MS within different European cultures. Research questions include: 1) How valid and reliable are the Dutch, Flemish, Slovenian and Spanish versions of the LBI to assess the construct of life balance in people with multiple sclerosis; 2) Do the Dutch, Flemish, Slovenian and Spanish versions of the LBI have adequate cross-cultural validity in people with MS to assess life balance in these cultures?

## 2. Methods

### 2.1. Design and sample

The cross-sectional multicenter clinimetric study was performed within the network of RIMS (Rehabilitation In Multiple Sclerosis, [www.eurims.org](http://www.eurims.org)), including five MS rehabilitation centers in Belgium, the Netherlands, Slovenia and Spain. Data collection took place from February 2017 to July 2018.

The study protocol was approved by the ethical review boards of all partners. All participants provided informed consent. Participants were recruited by (digital and paper) leaflets and personal communication of staff in the participating in- and outpatient clinics. Study participation was allowed when the following criteria were met: aged over 18 years,

definite diagnoses of MS, native speaking local language, adequate cognitive functioning (Symbol Digit Modalities Test, SDMT > 49) (Parmenter et al., 2007; López-góngora et al., 2015) or Multiple Sclerosis Neuropsychological Screening Questionnaire (MSNQ) (patient version)  $\leq 23$  (Benedict et al., 2003,2004)), stable clinical situation (no relapse up till 1 month prior to baseline assessment) and no initiation or changes in treatment during the time interval between test and retest. Every center included a minimum of 30 eligible subjects (de Vet et al., 2011).

### 2.2. Procedure

Participants completed all assessment instruments in an online or paper questionnaire set. To explore test-retest reliability, LBI was assessed twice with an interval of one week including an additional question about important changes during this short interval. An invitation to the second administration was sent by the online system one week following the first assessment.

### 2.3. Assessment

The questionnaire set included demographic and clinical information, like employment, living situation and disease duration. Mobility status was assessed by self-reporting the ability to walk indoor and outdoor with or without using an aid or a (manual versus powered) wheelchair.

The Life Balance Inventory (LBI) assesses the extent to which an individual perceives congruence between desired and actual time spent on 53 activity categories (Matuska, 2010, 2012a). The administration involves two steps: first, participants select whether or not they do or want to do each of the 53 activities (step 1). Then, for each of the selected activities, the satisfaction with the amount of real time spending on that activity compared to the amount of desired time in the past month is rated on a three-point scale. The highest score 3 is given in a balanced situation (“about right for me”); a slight imbalance (“sometimes less/more than I want”) is scored 2 and severe imbalance (“always less/more than I want”) is scored 1 (step 2) (Additional file 1). Scoring of the LBI generates a total average score across all performed or desired activities, with the total LBI score ranging between 1.00 and 3.00 with higher scores reflecting better life balance. Additionally, the LBI generates four subscale scores using 51 items (2 items are not used for subscale calculation): 1) Health subscale (6 items) (e.g. relaxing, getting regular exercise), 2) Identity subscale (15 items) (e.g. taking care of your appearance, participating in religious events), 3) Relationship subscale (10 items) (e.g. doing things with friends, partner), and 4) Challenge/interest subscale (20 items) (e.g. working for pay, making music). Internal consistency of the original LBI in a general population is high (Cronbach's alpha = 0.97) and convergent validity of the original LBI with the Depression Anxiety Stress Scale, the Personal Wellbeing Index-Adult and the Basic Psychological Needs Scale is demonstrated (Matuska, 2012a, 2012b).

Translation of the LBI was performed according to guidance on best practice for translation of patient reported outcomes (Wild et al., 2005). The original American version of the LBI was translated to the four local languages by each participating center. Independent back-translation was performed by professional interpreters coordinated by one person (DK). Consistency check of the original LBI versus backward translations and finding consensus of inconsistencies was performed by the research team in collaboration with the original developer of the LBI (KM). Some small changes were made based on these consensus meetings, e.g. the item “taking the bus” is changed into “taking public transport”.

To assess construct validity, and more specifically the convergent validity, several questionnaires were used based on clinimetric properties and availability in participating languages.

**Table 1**  
Demographic and clinical characteristics.

Variable	Dutch (n = 81)	Slovenian (n = 48)	Flemish 1 (n = 42)	Flemish 2 (n = 105)	Spanish (n = 37)	Difference between samples (p-value)*
Gender (% female)	74	79	57	63	62	0.09
Age (mean ± standard deviation, yr)	54 ± 9.6	44 ± 11.2	49 ± 12	50 ± 10.6	47 ± 9.0	<0.01
Educational level (% >12 years)	58	63	38	60	81	<0.01
Payed Job (% yes)	32	52	17	31	16	<0.01
Duration MS (years, min–max)	17 (4–35)	7 (0–29)	12 (0–30)	13 (0–32)	14 (1–29)	<0.01
Current rehab program (% inpatient/outpatient)	0/27	4/35	21/67	34/49	0/100	<0.01
Mobility (% no walking aids/walking aids/wheelchair)	51/19/30	56/29/15	14/19/67	39/18/43	24/14/62	<0.01
MS Neuropsychological Screening Questionnaire (MSNQ) (median, min–max)	25 (5–56)	20 (0–46)	30 (3–51)	23 (5–59)	–	0.11
Symbol Digit Modalities Test (SDMT) (median, min–max)	–	–	–	–	48 (37–66)	NA
Time between two measurements (median, SD, days)	9.1 (8.9)	11.8 (15.3)	17.5 (20.8)	6.1 (4.7)	3.2 (2.3)	<0.01
Life Balance Inventory (LBI) baseline (median, min–max)						
LBI total	2.50 (1.56–3.00)	2.47 (1.22–3.00)	2.36 (1.44–2.92)	2.32 (1.27–3.00)	2.13 (1.15–2.91)	<0.01
LBI health	2.50 (1.00–3.00)	2.50 (1.17–3.00)	2.37 (1.67–3.00)	2.40 (1.00–3.00)	2.50 (1.40–3.00)	0.72
LBI challenges	2.45 (1.50–3.00)	2.41 (1.50–3.00)	2.37 (1.43–3.00)	2.31 (1.13–3.00)	2.29 (1.11–3.00)	0.01
LBI relationships	2.63 (1.25–3.00)	2.57 (1.00–3.00)	2.39 (1.17–3.00)	2.29 (1.13–3.00)	2.00 (1.00–3.00)	<0.01
LBI identity	2.56 (1.33–3.00)	2.45 (1.17–3.00)	2.32 (1.25–3.00)	2.27 (1.00–3.00)	2.00 (1.00–3.00)	<0.01
Modified Fatigue Impact Scale (MFIS) (median, min–max)						
MFIS total	45 (4–68)	44 (3–73)	49 (17–77)	42 (0–67)	43 (19–73)	0.05
MFIS physical	21 (0–36)	23 (1–31)	24 (10–35)	21 (0–35)	24 (3–34)	0.01
MFIS cognitive	20 (0–34)	12.5 (1–36)	21 (0–38)	18 (0–32)	17 (0–35)	0.22
MFIS psychosocial	4 (0–8)	4 (0–8)	5 (0–8)	4 (0–8)	5 (1–8)	0.06
Quick Inventory of Depressive Symptomatology Self-Rated (QIDS-SR16) total score (median, min–max)	7 (0–17)	–	8 (0–18)	8 (0–19)	9 (3–23)	0.04
No depression (n, %)	28 (35)	–	14 (33)	29 (28)	12 (32)	0.21
Mild depression (n, %)	38 (48)	–	14 (33)	45 (43)	11 (30)	
Moderate depression (n, %)	13 (16)	–	9 (21)	25 (24)	7 (19)	
Severe depression (n, %)	1 (1)	–	5 (12)	6 (6)	19 (16)	
Very severe depression (n, %)	–	–	–	–	1 (3)	
University of Washington Self-Efficacy Scale (UW-SES), t-score (mean (SD))	46 (9.9)	44.9 (9.9)	40.7 (8.1)	41.8 (7.9)	43.3 (8.1)	<0.01
Quality of Life, Short Form-36 health survey (SF36) (median, min–max)						
SF36 physical functioning	55 (0–100)	55 (0–100)	17.5 (0–95)	35 (0–100)	35 (5–80)	<0.01
SF36 role limitation (physical)	25 (0–100)	25 (0–100)	0 (0–100)	0 (0–100)	0 (0–100)	<0.01
SF36 role limitation (emotional)	100 (0–100)	66.7 (0–100)	16.7 (0–100)	66.7 (0–100)	100 (0–100)	<0.01
SF36 social functioning	62.5 (12.5–100)	50 (12.5–100)	50 (12.5–100)	50 (0–100)	62.5 (12.5–100)	<0.01
SF36 bodily pain	77.6 (10.2–100)	55.1 (0–100)	65.3 (10.2–100)	75.5 (10.2–100)	67.3 (0–100)	<0.01
SF36 mental health	76 (32–100)	68 (20–96)	66 (12–96)	68 (12–100)	68 (16–96)	0.18
SF36 vitality	50 (10–90)	45 (0–85)	45 (15–80)	50 (10–100)	35 (0–80)	0.03
SF36 general health	50 (15–90)	52.5 (10–90)	35 (5–95)	40 (5–95)	35 (15–80)	<0.01

\* based on one-way analysis of variance (ANOVA) or Kruskal-Wallis test.

- 1) Short Form-36 health survey (SF-36v1), a frequently used instrument to assess quality of life (Ware et al., 1993). SF-36v1 includes 36 items and generates eight subscales (physical functioning, role limitations due to physical problems, bodily pain, general health perceptions, vitality, social functioning, role limitations due to emotional problems, and mental health). In each health domain raw scores are transformed to a range from 0 to 100 with higher scores indicating better health. There is considerable evidence for the validity of the SF-36 in a variety of populations including MS (Vickrey et al., 1995).
- 2) University of Washington Self-Efficacy Scale (UW-SES, 6-item version), a reliable and valid short questionnaire, scoring self-efficacy in managing the consequences of the disease in six items on a 5-point Likert scale, resulting in a summary score range of 6–30, transformed into t-scores, with higher scores indicating better self-efficacy (Amtmann et al., 2012).
- 3) Modified Fatigue Impact Scale (MFIS) (Kos et al., 2005; Multiple Sclerosis Council for Clinical Practice Guidelines, 1998), a 21-item questionnaire assessing the impact of fatigue on daily functioning on a 4-item Likert scale, with adequate reliability and validity in MS (Elbers et al., 2012). Total MFIS score of 0–84 and three subscale score of physical (0–36), cognitive (0–40) and psychosocial (0–8) fatigue impact can be calculated, although the latter

subscale is not adequate for separate use (Kos et al., 2005). Higher scores indicate increased fatigue impact.

- 4) Quick Inventory of Depressive Symptomatology Self-Rated (QIDS-SR16) (Rush et al., 1996; Fischer et al., 2015), a 16-item self-report list to assess the severity of the nine diagnostic symptom criteria used in DSM for the prior seven days. Items are scored on a 4-point scale and recoded in a total score from 0 to 27, creating five severity categories: no depression (0–5), mild (6–10), moderate (11–15), severe (16–20) and very severe (21–27) (<http://www.ids-qids.org/administration.html>).

#### 2.4. Data analyses

All statistical analyses were performed using SPSS version 25 (IBM Corporation, 2017) and MedCalc Statistical Software version 16.2.1 (MedCalc Software bvba; Ostend, Belgium, 2016). Alpha level was set at 0.05. Using the Kolmogorov–Smirnov tests, LBI total scores were normally distributed and therefore, parametric statistics were used.

To evaluate test-retest reliability of every LBI version, intraclass correlation coefficients (ICCs) were calculated in a two-way mixed model (absolute agreement). To reach sufficient reliability, an ICC should have a minimum level of 0.70 (de Vet et al., 2011).

Additionally, to explore systematic differences between first and

second LBI administration (test-retest), Bland Altman plots were created for each LBI version, including mean differences (systematic error) and 95% limits of agreement (random error) (de Vet et al., 2011). To evaluate whether the mean differences between measurements were statistically different from zero (i.e. systematic bias), a one-sample *t*-test was used. The 95% limits of agreement were set at  $-1.0$  and  $1.0$  (i.e. a category change in LBI total score) to appraise random error. For LBI total scores, the standard error of measurement (SEM) was calculated for test-retest agreement [ $SEM = SD \times \sqrt{1 - ICC}$ ] and smallest detectable change (SDC) [ $SDC = 1.96 \times \sqrt{2} \times SEM$ ] (de Vet et al., 2011; Beckerman et al., 2001). Proportional bias was investigated by a least-square regression analysis with difference between measurements as dependent variable and mean of two measurements as independent variable. Proportional bias is present when the 95% confidence interval (CI) of the slope  $\beta$  includes the value 1 (Altman and Bland, 1983; Ludbrook, 1997).

Convergent validity was considered with Pearson correlation coefficients between LBI total scores and related constructs, based on the following a priori defined hypotheses: LBI is moderately to highly correlated to quality of life and self-efficacy (correlation coefficients  $\geq 0.5$ ), LBI is negatively correlated to the impact of fatigue and depression (correlation coefficients between 0.3 and 0.7).

### 3. Results

The total sample consisted of 313 people with MS, with a mean age of 50 years, ranging from 22 to 75, with an MS duration of  $13 (\pm 8)$  years. There were five cultural subsamples: Dutch ( $n = 81$ , 74% women,  $54 \pm 9.6$  years of age), Flemish 1 ( $n = 42$ , 57% women,  $49 \pm 12$  years), Flemish 2 ( $n = 105$ , 63% women,  $50 \pm 10.6$  years), Slovenian ( $n = 48$ , 79% women,  $44 \pm 11.2$  years) and Spanish ( $n = 37$ , 62% women,  $47 \pm 9.0$  years). Two Flemish subsamples were used because of potential cultural differences between rural area (Flemish 1) and more urban area (Flemish 2) and the balance in sample sizes.

Gender distribution, self-perceived cognitive problems and some subscales of fatigue impact and quality of life were similar, whereas other demographic and clinical characteristics differed between the subsamples (Table 1). Due to one extreme outlying LBI score in a Slovenian participant further analyses are based on 47 people from Slovenia.

Mean number of days between test and re-test measurements was  $8.7 \pm 11.3$ , ranging from 3 to 17 days in various samples (Table 1). Changes in the situation between first and second LBI assessment were reported by 35 people (11% of total sample): personal issues (3), acute viral infection respiratory tract (2), feel better (11), feel worse (15), changes in medication use (2), initiation therapeutic interventions (orthopedics, physiotherapy) (2). All below mentioned analyses were repeated excluding this subsample, resulting in similar results (results not shown).

ICCs only reached the required level of 0.70 in one of the samples (Flemish 2) (Table 2). The smallest detectable change score ranged from 0.19 to 0.86. Mean differences between repeated LBI assessment for every subsample ranged from  $-0.01$  to  $-0.16$  with 95% limits of agreement (LoA) from  $-0.89$  to 0.70 (Fig. 1). In one sample (Spanish) the mean difference between first and second LBI total score assessment differed significantly from zero (mean difference of 0.16 with 95% confidence interval [0.03–0.28]), assuming a systematic error. In none of the samples the 95% CI for the slope  $\beta$  exceeded the value 1, assuming no proportional bias.

Correlation coefficients with related constructs did not meet all hypotheses for construct validity (Table 3): LBI was moderately correlated to quality of life in two subdomains of one sample and self-efficacy did not reach the desired correlation coefficients of 0.5 or higher. The expected negative correlations between LBI and fatigue impact

**Table 2**

Test-retest reliability of LBI total score for separate samples.

Sample	n	ICC [95% CI]	SEM	SDC
Dutch	72	0.55 [0.37–0.69]	0.22	0.61
Flemish 1	31	0.65 [0.39–0.82]	0.20	0.55
Flemish 2	95	0.88 [0.83–0.92]	0.07	0.19
Slovenian	43	0.58 [0.35–0.75]	0.19	0.52
Spanish	31	0.45 [0.15–0.67]	0.27	0.75
Total	278	0.71 [0.64–0.76]	0.26	0.52

ICC [95%CI]: intraclass correlation coefficients with their 95% confidence intervals; SEM: standard error of measurement; SDC: smallest detectable change.

(MFIS) were found in two samples and with depression (QIDS) in one sample only.

### 4. Discussion

This study evaluated psychometric properties of four translated versions of the Life Balance Inventory and showed mixed results. Test retest reliability is adequate in one sample only and did not seem sufficient in the other language samples. Random error fell within the a priori set ranges from  $-1$  to  $1$ , indicating that a difference of one LBI answer category would exceed the measurement error. These results are applicable for the mean LBI total score, and cannot be generalized to scores at item level as we did not perform these sub-analyses. The smallest detectable change score ranged from an acceptable value of 0.19 (which is 6% relative to the maximum score of 3) to a rather high value of 0.86 (29%). In a stroke sample using the Flemish LBI, similar intraclass correlation coefficients (0.88) and SDC scores of the total score (0.39, 20%) were found (Van Gils et al., 2019).

It is arguable that results may partly be explained by the cognitive functioning of participants. Although adequate cognitive functioning was an inclusion criterion, not all participants stayed below the desired level of the MSNQ (up to 23) (Benedict et al., 2004). Post-hoc analyses with subgroups based on the cut-off of MSNQ revealed better test-retest reliability in most samples with ICCs above 0.70 (Additional file 2). These results suggest that participants would require a minimum level of (self-reported) cognitive performance to reliably complete the LBI. However, higher self-reported MSNQ scores may also reflect depressive feelings (Benedict et al., 2004). Nearly half of the sample (42.6%) reported moderate to severe depression, which is a higher prevalence than non-population-based samples (30.5%) (Boeschoten et al., 2017). This high depression rate may partly explain the higher MSNQ scores. Recent work also reported difficulties in identifying cognitive impairment in people with fatigue and depressive symptoms based on self-report (Hughes et al., 2019). Self-report may therefore not be the most adequate cognitive screening method, but the online registration hindered any objective testing procedure. Computer-generated cognitive screening procedures may be applied in future research, although in these kinds of research designs, it remains uncertain whether the participant performed the test alone (and not with help from a relative) (Lapshin et al., 2013; Patel et al., 2018).

In general, it seems that the accuracy of self-report assessment may be influenced by education level or a reduced self-awareness of participants (Goverover et al., 2009). In our sample, the majority was highly educated ( $>12$  years of education), as about 40% had a lower education. Scoring the LBI items adequately may be difficult for this subgroup, however post-hoc analyses did not reveal any effect of education on the stability of LBI responses. All assessment was based on self-report, therefore the potential self-awareness may have been present in all of the answers and probably does not explain the unexpected results. Besides, constructs based on meaning and value are very hard to assess otherwise.

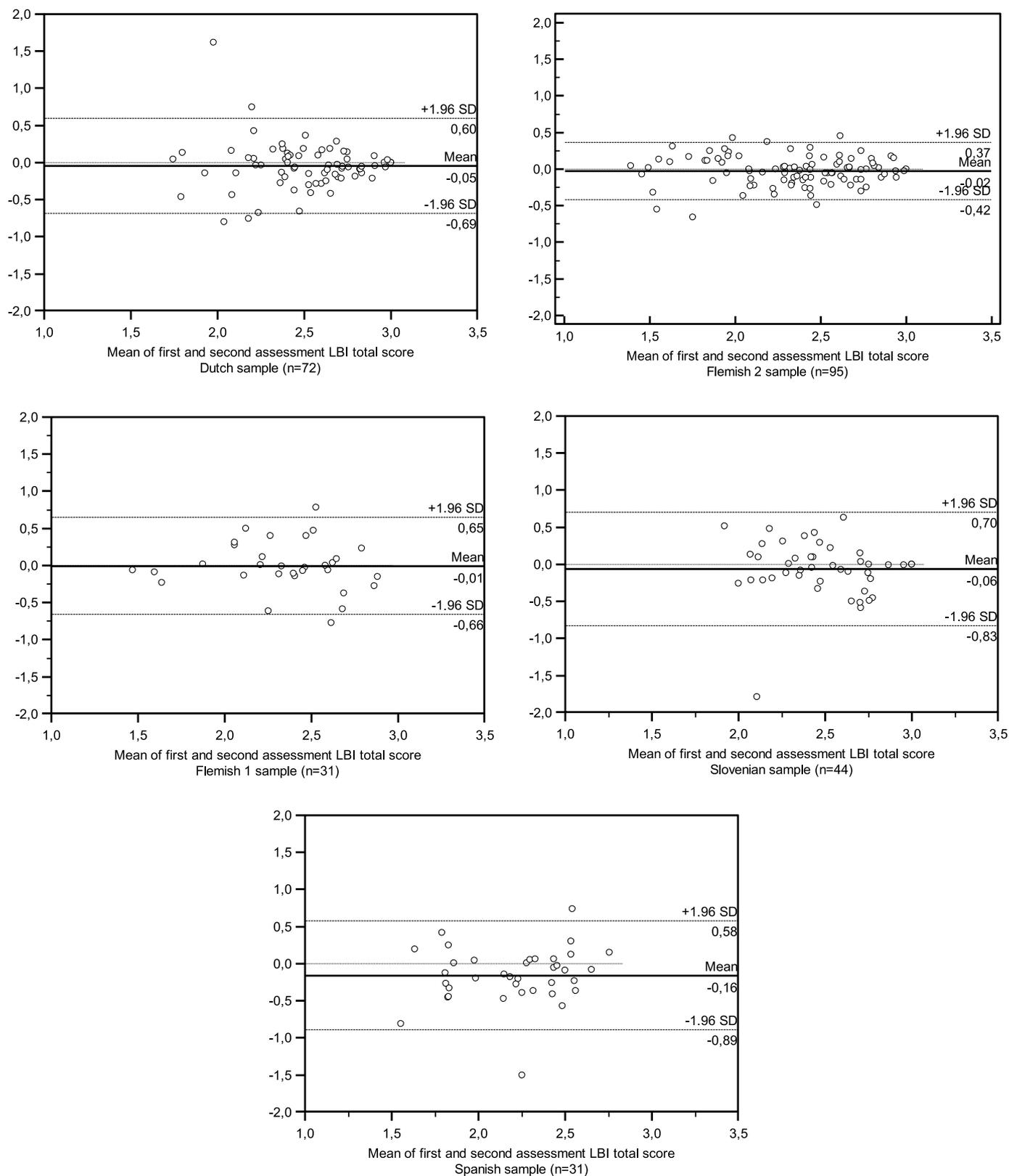


Fig. 1. Bland Altman plots of separate samples.

The LBI assesses the constructs “life balance congruence” (i.e. the match between actual and desired activity configuration in everyday life, expressed in total LBI score) and “life balance equivalence” (i.e. allocated satisfaction with time use across various activities, visible in the subscales) (Matuska, 2012a, 2012b). Convergent validity of LBI

assessing congruence was supported only to a small extent. In most samples, the correlations of LBI total score with related constructs did not reach the hypothesized levels. Only in the Dutch sample LBI showed low to moderate correlations with all of the related constructs. Larivière and Levasseur also found moderate correlations with stress, physical

**Table 3**  
Construct validity of LBI total score.

Correlation LBI total score with:	Dutch (n = 81)	Flemish 1 (n = 42)	Flemish 2 (n = 105)	Slovenian (n = 47)	Spanish (n = 37)	Total (n = 312)
SF36 physical functioning	0.33**	0.14	0.28**	0.15	0.09	0.28**
SF36 role limitation physical	<b>0.53**</b>	-0.18	0.28**	-0.02	-0.09	0.24**
SF36 role limitation emotional	0.41**	0.06	0.07	-0.02	0.24	0.15**
SF36 social functioning	0.38**	0.26	0.33**	-0.07	0.15	0.25**
SF36 bodily pain	0.49**	0.09	0.39**	0.24	0.16	0.32**
SF36 mental health	0.46**	0.34*	0.35**	0.09	0.33*	0.35**
SF36 vitality	<b>0.55**</b>	0.07	0.35**	-0.12	-0.03	0.29**
SF36 general health	0.30**	0.12	0.21*	0.18	0.05	0.26**
UW-SES (self-efficacy)	0.42**	0.28	0.17	0.04	0.03	0.24**
MFIS (fatigue impact)	<b>-0.50**</b>	-0.06	<b>-0.44**</b>	0.04	-0.29	<b>-0.33**</b>
QIDS-SR16 (depression)	<b>-0.44**</b>	-0.17	<b>-0.28**</b>	NA	-0.30	<b>-0.33**</b>

\*  $p < 0.05$ .

\*\*  $p < 0.01$ ; NA: not available; levels of correlation coefficients as hypothesized are marked in bold.

and mental health in an MS sample, but concluded the French version of the LBI is a valid tool to assess time use (Larivière and Levasseur, 2016). Also in Stroke patients, LBI correlated moderately with the participation and emotion subscale of the Stroke Impact Scale (Van Gils et al., 2019). It is unclear whether the construct ‘life balance congruence’ is less related to health-related constructs assessed in all these studies, or the LBI is unable to assess this construct. Studies focusing on the relationship with similar constructs like occupational balance may clarify this further (Wagman et al., 2015).

Self-efficacy includes beliefs about own capabilities to produce desired goals and was therefore expected to be related to life balance (Bandura, 1997). We found self-efficacy scores in line with similar MS samples (overall mean  $49.9 \pm 9.3$ ) (Ammann et al., 2012). However, LBI scores did not relate to self-efficacy as hypothesized. UW-SES scores showed negative moderate correlations with fatigue impact and depression scores (results not displayed), supporting the validity of the UW-SES.

About 65% of the participants experienced high fatigue impact (MFIS score of 39 or higher) (Flachenecker et al., 2002), which is in line with the proportion of people with MS reporting fatigue (56%, based on Fatigue Severity Scale) in a large international cohort (Weiland et al., 2018). Nevertheless, LBI showed negative correlations with fatigue impact as hypothesized in two samples only.

Quality of life scores (SF-36) were not similar in all samples. Obviously, the groups with a high proportion of wheelchair users (Flemish and Spanish samples) showed decreased physical functioning and physically based role limitation scores. Most samples scored quality of life related to bodily pain, mental health, vitality and general health similar or higher than a general MS population (Morley et al., 2018). Correlations between LBI and SF-36 scores were generally low and did not confirm our hypothesis. Although some of the questions in SF-36v1 are very alike (“... Cut down the amount of time you spent on work or other activities”), LBI scores did not seem to be related to health-related quality of life.

Considering the limited evidence for construct validity in our samples, it can be questioned whether the LBI assesses what is intended, i.e. satisfaction with time spent in desired activities. The total LBI score is calculated by taking the average of all activities without weighing the importance of specific items. Besides, the total score does not distinct between “too much” or “too less” time, which may refer to complete different situations. For example, when a person indicates to spend too much time in a specific activity, this may refer to the feeling of reduced relevance (I do not want to do this activity, but I should), but it may also reflect the increased time people need to perform a desired activity related to disability. In both cases, people are not in balance, but the construct is not specified in a similar way. However, therapists in clinical practice may use the LBI to detect desired activities in daily life, especially those who are not performed at time of assessment but people with MS would like perform. The therapist and client may

jointly explore reasons of non-performance or dissatisfaction with time and energy use to support goal setting. Further research on internal validity may provide better insight and opportunities to improve the scale for international research purposes.

#### 4.1. Limitations

Participants were recruited through MS centers, therefore results cannot be generalized to the entire MS population. Besides, subsamples showed differences in demographic and clinical variables. A sampling method with extended selection criteria would probably result in more comparable subsamples. However, we did not want to exclude people with decreased mobility or a certain age range in this clinimetric study in order to be able to use the LBI in real-life clinical inpatient and outpatient settings. Research in the general population showed differences in raw life balance scores between different demographic groups (Matuska, 2012a, 2012b; Matuska et al., 2013), but provided no insight on the differences in LBI reliability in subgroups.

Although participants were instructed and reminded to complete the repeated LBI assessment within one week, only 71% did so. Time interval of the second registration varied between samples, however did not explain the limited test-retest reliability.

The convergent validity was evaluated with several measures with a priori formulated hypotheses of the relationships based on literature and clinical experience. Although the instruments are adequate and frequently used in other studies and clinical practice, other measures and constructs (like perceived stress or satisfaction with life) would have provided an additional insight in the validity of the LBI.

In our validity analyses, we only considered total LBI total score and not the LBI subscales. Using the total score of the LBI may be limited in exactly quantifying satisfaction with time spent in desired activity. Therefore, no conclusions can be drawn about life balance equivalence (allocated satisfaction with time use across various activities). Future research may explore other ways of scoring the perceived congruence between desired and actual time use in various activity categories more accurately.

## 5. Conclusions

Our cross-cultural results do not support the use of the LBI as an outcome measure. Due to limited evidence for test-retest reliability and validity, the use of the current LBI translations is limited. Based on the current findings, future research could focus on further development of the questionnaire to ensure a reliable, valid and responsive instrument to assess life balance congruence and equivalence in people with MS. This kind of instrument may support rehabilitation teams in their effort to optimize activity and participation.

## Funding

This study was partly funded by a grant (RiGra 2016) provided by RIMS, European Network for best practice and research in MS Rehabilitation ([www.eurims.org](http://www.eurims.org)).

## Ethics approval and consent to participate

The study protocol was approved by the ethical review boards of all participating centers: The Netherlands: VU University Medical Center Amsterdam; Slovenia: University Hospital Ljubljana; Spain: Multiple Sclerosis Center of Catalonia, Barcelona; Belgium: University Hospital Leuven, National MS Center Melsbroek, Revalidatie & MS Centrum Overpelt. All participants provided informed consent.

## CRedit authorship contribution statement

**Daphne Kos:** Data curation, Investigation, Writing - original draft, Writing - review & editing, Formal analysis. **Sofie Ferdinand:** Data curation, Investigation, Writing - original draft, Writing - review & editing. **Marijke Duportail:** Data curation, Investigation, Writing - original draft, Writing - review & editing. **Isaline Eijssen:** Data curation, Investigation, Writing - original draft, Writing - review & editing. **Sofie Schouteden:** Data curation, Investigation, Writing - original draft, Writing - review & editing. **Lore Kerckhofs:** Data curation, Investigation, Writing - original draft, Writing - review & editing. **Jelka Jansa:** Data curation, Writing - original draft, Writing - review & editing. **Núria Fillo:** Data curation, Investigation, Writing - original draft, Writing - review & editing. **Kathleen Matuska:** Investigation, Writing - original draft, Writing - review & editing. **Heleen Beckerman:** Data curation, Investigation, Writing - original draft, Writing - review & editing, Formal analysis.

## Declaration of Competing Interest

None.

## Acknowledgments

All participants are acknowledged for their time and efforts. All collaborating centers are acknowledged for providing time for therapists to be part of this study group.

## Supplementary materials

Supplementary material associated with this article can be found, in the online version, at [doi:10.1016/j.msard.2019.101879](https://doi.org/10.1016/j.msard.2019.101879).

## References

- Altman, D.G., Bland, J.M., 1983. Measurement in medicine : the analysis of method comparison studies. *Statistician* 32, 307–317.
- Amtmann, D., Bamer, A.M., Cook, K.F., Askew, R.L., Noonan, V.K., Brockway, J.A., 2012. University of Washington self-efficacy scale: a new self-efficacy scale for people with disabilities. *Arch. Phys. Med. Rehabil.* 93, 1757–1765.
- Bandura, A., 1997. *Self-efficacy*. Springer Reference. Springer-Verlag, Berlin/Heidelberg.
- Barin, L., Salmen, A., Disanto, G., Babačić, H., Calabrese, P., Chan, A., et al., 2018. The disease burden of multiple sclerosis from the individual and population perspective: which symptoms matter most? *Mult. Scler. Relat. Disord* 25, 112–121.
- Beckerman, H., Roebroek, M.E., Lankhorst, G.J., Becher, J.G., Bezemer, P.D., Verbeek, A.L.M., 2001. Smallest real difference, a link between reproducibility and responsiveness. *Qual. Life Res.* 10, 571.
- Benedict, R.H.B., Cox, D., Thompson, L.L., Foley, F., Weinstock-Guttman, B., Munschauer, F., 2004. Reliable screening for neuropsychological impairment in multiple sclerosis. *Mult. Scler. J.* 10, 675–678.
- Benedict, R.H.B., Munschauer, F., Linn, R., Miller, C., Murphy, E., Foley, F., et al., 2003. Screening for multiple sclerosis cognitive impairment using a self-administered 15-item questionnaire. *Mult. Scler.* 9, 95–101.
- Boeschoten, R.E., Braamse, A.M.J., Beekman, A.T.F., Cuijpers, P., van Oppen, P., Dekker, J., et al., 2017. Prevalence of depression and anxiety in multiple sclerosis: a systematic review and meta-analysis. *J. Neurol. Sci.* 372, 331–341.
- de Vet, H., Terwee, C., Mokkink, L., Knol, D., 2011. *Measurement in Medicine*. Cambridge University Press, New York.
- Elbers, R.G., Rietberg, M.B., van Wegen, E.E.H., Verhoef, J., Kramer, S.F., Terwee, C.B., et al., 2012. Self-report fatigue questionnaires in multiple sclerosis, Parkinson's disease and stroke: a systematic review of measurement properties. *Qual. Life Res.* 21, 925–944.
- Fischer, A., Fischer, M., Nicholls, R.A., Lau, S., Poettgen, J., Patas, K., et al., 2015. Diagnostic accuracy for major depression in multiple sclerosis using self-report questionnaires. *Brain Behav* 5 (9), e00365.
- Flachenecker, P., Kämpfel, T., Kallmann, B., Gottschalk, M., Grauer, O., Rieckmann, P., et al., 2002. Fatigue in multiple sclerosis: a comparison of different rating scales and correlation to clinical parameters. *Mult. Scler.* 8 (6), 523–526.
- Glaser, A., Stahmann, A., Meissner, T., Flachenecker, P., Horáková, D., Zaratini, P., et al., 2019. Multiple sclerosis registries in Europe - An updated mapping survey. *Mult. Scler. Relat. Disord.* 27, 171–178.
- Goverover, Y., Chiaravalloti, N., Gaudino-Goering, E., Moore, N., DeLuca, J., 2009. The relationship among performance of instrumental activities of daily living, self-report of quality of life, and self-awareness of functional status in individuals with multiple sclerosis. *Rehabil. Psychol.* 54 (1), 60–68.
- Hughes, A.J., Bhattarai, J., Paul, S., Beier, M., 2019. Depressive symptoms and fatigue as predictors of objective-subjective discrepancies in cognitive function in multiple sclerosis. *Mult. Scler. Relat. Disord. Elsevier B.V.* 30, 192–197.
- Kos, D., Kerckhofs, E., Carrea, I., Verza, R., Ramos, M., Jansa, J., 2005. Evaluation of the modified fatigue impact scale in four different European countries. *Mult. Scler.* 11, 76–81.
- Lapshin, H., Lanctôt, K.L., O'Connor, P., Feinstein, A., 2013. Assessing the validity of a computer-generated cognitive screening instrument for patients with multiple sclerosis. *Mult. Scler. J.* 19, 1905–1912.
- Larivière, N., Levasseur, M., 2016. Traduction et validation du questionnaire ergothérapique l'Inventaire de l'équilibre de vie. *Can. J. Occup. Ther.* 83, 103–114.
- López-góngora, M., Querol, L., Escartín, A., 2015. 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.* 15, 2–9.
- Ludbrook, J., 1997. Comparing methods of measurement. *Clin. Exp. Pharmacol. Physiol.* 24, 193–203.
- Matuska, K., 2012a. Description and development of the life balance inventory. *OTJR Occup. Particip. Heal.* 32, 220–228.
- Matuska, K., Bass, J., Schmitt, J.S., 2013. Life balance and perceived stress: predictors and demographic profile. *OTJR (Thorofare, NJ)*. 33, 146–158.
- Matuska, K., 2012b. Validity evidence of a model and measure of life balance. *OTJR Occup. Particip. Heal.* 32, 229–237.
- Matuska, K.M., Christiansen, C.H., 2008. A proposed model of lifestyle balance. *J. Occup. Sci.* 15, 9–19.
- Matuska, K.M., Erickson, B., 2008. Lifestyle balance: how it is described and experienced by women with multiple sclerosis. *J. Occup. Sci.* 15, 20–26.
- Matuska, K.M., 2010. Validity evidence for a model and measure of life balance. University of Minnesota.
- Morley, D., Dummett, S., Kelly, L., Fitzpatrick, R., Jenkinson, C., 2018. Predictors of activity and participation across neurodegenerative conditions: a comparison of people with motor neurone disease, multiple sclerosis and Parkinson's disease. *BMC Neurol.* 18, 1–6.
- Multiple Sclerosis Council for Clinical Practice Guidelines, 1998. *Fatigue and Multiple Sclerosis: Evidence-based Management Strategies for Fatigue in Multiple Sclerosis*. Paralyzed Veterans America, Washington.
- Parmenter, Ba, B., Weinstock-Guttman, Garg, N., Munschauer, F., Benedict, R.H.B., 2007. Screening for cognitive impairment in multiple sclerosis using the Symbol Digit Modalities Test. *Mult. Scler.* 13, 52–57.
- Patel, V.P., Shen, L., Rose, J., Feinstein, A., 2018. Taking the tester out of the SDMT: a proof of concept fully automated approach to assessing processing speed in people with MS. *Mult. Scler. J.* 1–8.
- Rush, A., Gullion, C., Basco, M., Jarrett, R., Trivedi, M., 1996. The Inventory of Depressive Symptomatology (IDS): psychometric properties. *Psychol. Med.* 26, 477–486.
- Van Gils, A., Meyer, S., Beyens, H., Schillebeeckx, F., Verheyden, G., Kos, D., 2019. Psychometric properties of the Flemish version of the life balance inventory: a patient-reported outcome measure in the chronic stage after stroke. *Disabil. Rehabil.* 1–8. <https://doi.org/10.1080/09638288.2019.1575921>. Mar 4.
- Vickrey, B., Hays, R., Harooni, R., Myers, L., Ellison, G., 1995. A health-related quality of life measure for multiple sclerosis. *Qual. Life Res.* 4, 187–206.
- Wagman, P., Håkansson, C., Jonsson, H., 2015. Occupational balance: a scoping review of current research and identified knowledge gaps. *J. Occup. Sci.* 22, 160–169.
- Ware, J., Snow, K., Kosinski, M., Gandek, B., 1993. *SF-36 Health Survey: Manual and Interpretation Guide*. Health Institute, New England Medical Center, University of Michigan.
- Weiland, T.J., De Livera, A.M., Brown, C.R., Jelinek, G.A., Aitken, Z., Simpson, S.L., et al., 2018. Health outcomes and lifestyle in a sample of people with multiple sclerosis (HOLISM): longitudinal and validation cohorts. *Front. Neurol.* 9, 1–17.
- Wild, D., Grove, A., Martin, M., Eremenco, S., McElroy, S., Verjee-Lorenz, A., et al., 2005. Principles of good practice for the translation and cultural adaptation process for patient-reported outcomes (PRO) measures: report of the ISPOR Task Force for Translation and Cultural Adaptation. *Value Heal* 8 (2), 94–104.