



Italian translation and psychometric validation of the Manual Ability Measure-36 (MAM-36) and its correlation with an objective measure of upper limb function in patients with multiple sclerosis

Claudio Solaro¹ · Rachele Di Giovanni¹ · Erica Grange¹ · Giampaolo Brichetto² · Margit Mueller² · Andrea Tacchino² · Rita Bertoni³ · Francesco Patti⁴ · Angelo Pappalardo⁴ · Luca Prosperini⁵ · Letizia Castelli³ · Rosalba Rosato⁶ · Davide Cattaneo³ · Davide Marengo⁶ 

Received: 6 July 2019 / Accepted: 18 January 2020 / Published online: 23 January 2020
© Fondazione Società Italiana di Neurologia 2020

Abstract

Background Multiple sclerosis (MS) is a chronic autoimmune disease of the central nervous system with an unpredictable course. During its course, deficits affecting upper limb functions may occur. Hence, there is a need for self-administered scales providing a comprehensive assessment of upper limb functions. The Manual Ability Measure-36 (MAM-36), which investigates patients' performance in activities of daily living requiring upper limb function, has not been adapted and validated in the Italian context.

Objectives We develop an Italian translation and validation of the MAM-36 in a population of people with MS (PwMS), explore its psychometric properties and investigate its associations with clinical data and the Nine Hole Peg Test (9-HPT).

Research plan and methods The multicentre study involved five Italian neurological centres. Subjects were evaluated using EDSS, 9-HPT and the MAM-36 scale. We used confirmatory factor analysis and Rasch analysis to investigate the properties of the MAM-36.

Results We enrolled 218 PwMS. Results supported the unidimensionality of the MAM-36, and adequate functioning of rating scale and items. Additionally, the MAM-36 showed weak negative associations with age and disease duration, and moderate associations with EDSS and 9-HPT scores.

Discussion The adapted MAM-36 showed adequate psychometric properties. However, indications of problematic targeting to PwMS with low disability emerged. For this reason, use of the scale appears to be more suitable among patients with moderate-to-severe disability.

Keywords Multiple sclerosis · MAM-36 · Upper limb function · 9-HPT

Introduction

Multiple sclerosis (MS) is a chronic, autoimmune disease, characterised by inflammatory demyelination, affecting the white and the grey matter in the central nervous system. It typically affects multiple functional systems, resulting in a variety of symptoms, which may include sensory disorders, muscle weakness, ataxia, tremor and especially lower and upper limb impairments [1]. Although hand movements are essential in daily living, as demonstrated by Johansson et al. [2], disability in people with MS (PwMS) is thought to be mainly related to ambulation; as a matter of fact, the upper limb function is often impaired in MS subjects [3]. Moreover, upper limb functioning could impact negatively on quality of life of PwMS [4, 5], especially due to the bimanual nature of several tasks of everyday life, such as

✉ Davide Marengo
davide.marengo@unito.it

¹ Department of Rehabilitation, CRRF “Mons. Luigi Novarese”, Moncrivello, VC, Italy
² Scientific Research Area, Italian Multiple Sclerosis Foundation (FISM), Genoa, Italy
³ IRCCS Fondazione Don Carlo Gnocchi, Milan, Italy
⁴ MS Center Institute of Neurological Sciences, University of Catania, Catania, Italy
⁵ Department of Neurosciences, S. Camillo-Forlanini Hospital, Rome, Italy
⁶ Department of Psychology, University of Turin, Turin, Italy

changing clothes, performing toileting or washing both the hands [6]. As reported by Bertoni et al. [7], 75 out of 110 subjects with moderate disability at the Expanded Disability Status Scale (EDSS) [8] had bilaterally impaired manual dexterity assessed with Nine Hole Peg Test (9-HPT) and correlate with ability in daily living activities.

Recently, as result of a revision on clinical tools to measure upper limb function in MS, the 9-HPT has been considered the gold standard for upper limb assessment [9], especially for its strict correlation with performance in daily living activities, and quality of life [10]. Moreover, the 9-HPT normative data have been recently published [11].

Based on a recent review [12], the most used patient-reported outcome measures for upper limb perceived function in MS are the Disability of Arm, Shoulder and Hand (DASH), the Manual Ability Measure-36 (MAM-36) and Abilhand questionnaires [13]. In spite of the importance of upper limb functions as predictor of patients' outcomes [14], in a recent overview of patient-reported outcome measures of manual ability, Kahraman et al. [15] concluded that no MS-specific scales are available. As regards the MAM-36, the scale has been used to assess subjective upper limb function in several neurological and not neurological diseases such as Charcot–Marie–Tooth [16], rheumatoid arthritis [17], acquired brain injuries [18], spinal cord injuries [18] and upper extremity orthopaedic disease [18, 19]. The psychometric validation of the MAM-36 [18] was performed on 337 subjects affected by variety of neurologic and musculoskeletal disorders including a small sample of 44 subjects with MS. The correlation between MAM-36 and EDSS was reported in 44 PwMS [20].

A validated version of the MAM-36 is not currently available for the Italian context. The aim of the study was to translate the MAM-36 to Italian and explore its psychometric properties in a large sample of PwMS, and to investigate construct validity by examining its association with demographic and clinical variables, and the 9-HPT. For the purpose of the present study, the psychometric validation of the MAM-36 was performed using a combination of confirmatory factor analysis (CFA) and Rasch modelling techniques.

Research plan and methods

Translation and adaptation of the MAM-36

The MAM-36 includes 36 items assessing perceived ease or difficulty in performing common tasks (e.g. eating, dressing, button clothes) using one's hands, regardless of which hand is used and excluding the use of adaptive equipment. Items are rated on a 4-point scale from 1 (cannot do it) to 4 (easy); a zero-response option is also included, indicating tasks that are almost never performed, with or without hand impairment. Scores on the 36 items are summed to create a raw total score

with a range from 36 to 144. The MAM-36 was translated into Italian by a professional translator with knowledge of health terminology. The translation was evaluated to ensure semantic equivalence and acceptability. During an initial meeting of MS experts, a list of possible alternatives for the controversial item stems and response choices was developed. Problematic items and response choices were retranslated into Italian from the original version, and a definitive version was determined by consensus. The questionnaire has been fulfilled by 5 MS patients followed by second experts' meeting. Subsequently, the Italian version was back-translated into English and compared with the original one.

Subjects

Study participants were a consecutive sample of patients followed at five MS outpatient clinics from “Department of Rehabilitation, Mons. Luigi Novarese Hospital” of Moncrivello; “Department of Neurology, University of Catania”; “Don Gnocchi Foundation” of Milano, “Rehabilitation Service of Liguria of the Italian Multiple Sclerosis Society (AISM)” of Genova; and “Sant’Andrea Hospital” of Rome. Each ethical committee of each participating centre allowed the ethical approval for this study (P.R.196REG2015). Signed informed consent was obtained from each patient prior to enrolment in the study according to the Declaration of Helsinki.

Inclusion criteria were as follows: a minimum age of 18, diagnosis of MS according to McDonald revised criteria at time of recruitment [21], a stable disease course without worsening more than 1 EDSS point over the last 3 months, relapse free at the time of enrolment into the study, completed MAM-36 and capable of understanding and providing signed informed consent. Exclusion criteria were presence of bilateral plegia, orthopaedic or neurological diseases or diseases other than MS. For each patient, we also collected demographic (age, gender) and clinical characteristics (disease course and duration, EDSS), and the 9-HPT.

Strategy of analyses

First, we computed descriptive statistics for the study measures. Next, we investigated the psychometric properties of the MAM-36.¹ As a first step, we examined the fit of the

¹ Prior to conducting these analyses, we investigated the presence of non-independence among MAM-36 scores due to patients' clustering into different centres. We computed the intraclass correlation coefficient (ICC) for MAM-36 scores via linear mixed model, controlling for demographic (age, gender) and clinical (disease course, disease duration, and clinician-rated EDSS) variables, i.e. we computed the residual ICC [22]. Based on our computation, the residual ICC for MAM-36 scores was 0.012, corresponding to a design effect of 1.61, which compared with the suggested cut-off (a design effect ≥ 2 , [23]), indicated no significant clustering effect was present beyond that due to demographic and clinical variables. For this reason, we did not control for the effect of centres in subsequent analyses.

MAM-36 data to a one-factor confirmatory factor analysis (CFA) model. The CFA was performed using MPLUS 7.3 using the WLSMV estimator. We evaluated model fit using the comparative fit (CFI), the Tucker-Lewis (TLI) and the root mean-square error of approximation (RMSEA) indexes. We considered values of CFI > 0.95, TLI > 0.95 and RMSEA < 0.05 as indication of good model fit, while CFI and TLI > 0.90, and RMSEA < 0.08, as indication of acceptable fit [24]. Significant standardised item loadings > 0.70 were considered indicating good convergent validity, while values equal or below 0.70, but exceeding 0.50, were considered acceptable [25]. Values of variance extracted (AVE) \geq 0.50 and omega (ω) composite reliability coefficient \geq 0.70 were considered further evidences of convergent validity of the item set [25–27].

Next, we implemented the Rasch rating scale model [28] to examine the functioning of the MAM-36 scale. Adequate functioning of the 4-point rating scale was established if results met the following criteria: (1) a minimum of ten responses in each rating category, (2) rating category measures increasing monotonically and (3) Outfit mean-square values for each rating category sitting below 2.0 [29]. Adequate item fit was determined by Infit and Outfit mean-square values sitting in a range of 0.6 to 1.4 [30], while items showing values sitting beyond this range but not exceeding 2.0 were considered unproductive for measurement, but not degrading [31]. Reliability of person scores was determined using the Rasch reliability index, assuming values \geq 0.90 as appropriate for clinical applications [32]. Additionally, we computed the person separation index and used it to determine the number of statistically distinguishable ability groups (i.e. person strata [33]). The dimensionality of the scale was examined by performing a principal component analysis (PCA) on Rasch residuals. Unidimensionality was established if Rasch measures explained \geq 40% variance of the data, and the first contrast had an eigenvalue lower than or equal to 2.0 and accounted for both less than 5% total variance, and less than 10% of unexplained variance [34, 35]. Rasch analyses were performed using Winsteps 3.68.2.

Next, we examined ceiling and floor effects of the MAM-36 score, which we considered significant if we found more than 15% of patients reporting either minimum or maximum extreme scores. Additionally, we computed 25th, 50th and 75th percentile values for the MAM-36 score, used them to stratify patients in four quartile ability groups and inspect the distribution of scores in each group.

Criterion validity of the MAM-36 score was investigated by examining its association with patients' demographic and clinical variables, and 9-HPT scores for both arms. The association of the MAM-36 score with age, disease duration, EDSS and 9-HPT scores was examined using Spearman's rank correlation coefficient. Gender differences on test scores were examined using Mann–Whitney *U* test. Association between the MAM-36 score and disease course was examined using Kruskal–Wallis test (with Dunn's post hoc test).

Finally, we inspected the distribution of MAM-36 scores among groups of patients showing different level of impairment based on existing normative data for the 9-HPT. Scores were categorised as indicating an “overt” impairment if \geq 2 SDs from the normative 9-HPT values [36]. Based on this threshold, we grouped patients distinguishing between those showing unilateral overt impairment ($N=40$) and bilateral overt impairment ($N=115$), and those for which both arms showed no overt impairment ($N=63$) and investigated differences across these groups on the MAM-36 using the Kruskal–Wallis test (with Dunn's post hoc test). Except where indicated, analyses were performed in SPSS, version 23.

Results

Two hundred and eighteen patients were recruited. Descriptive statistics for patients' demographic, clinical characteristics and the MAM-36 score are reported in Table 1.

Psychometric properties of the MAM-36

Results of the investigation of the psychometric properties of the MAM-36 are reported in Table 2 and 3, and Fig. 1. As regards the CFA, results showed the one-factor model had acceptable model fit based on recommended thresholds (CFI = 0.96, TLI = 0.96, RMSEA = 0.05). Standardised loadings were > 0.70 for all items except for three items—i.e. items 10, 18 and 31—whose loadings ranged from 0.58 to 0.65 (see Table 3). Combined with an AVE of 0.64, and excellent composite reliability ($\omega=0.98$), results of the CFA supported the convergent validity of the item set.

Rasch analyses showed the rating scale diagnostics met all three essential criteria for a functioning rating scale, as responses were greater than 10 for each category; Rasch category measures increased monotonically; and Infit and Outfit mean-square statistics were within suggested thresholds (Table 2). As regards item functioning, results are reported in Table 3. Item difficulty ranged from -1.84 to 1.59 ; item 2 (“Carry a shopping bag with a hand loop”) was the most difficult item, while item 1 (“Eat a sandwich”) was the easiest. Compared with the distribution of person ability (see Fig. 1), which ranged from -1.89 to 5.43 logit with a mean of 2.32 logit, easy to see that most of the items were perceived by patients to be relatively easy, as the average person ability was over 2 logit beyond average item difficulty. The Rasch person separation reliability index was 0.91, indicating excellent reliability. Person separation index was 3.21, resulting in a person strata value of 4.61, which indicate approximately four distinct ability groups could be detected using MAM-36 scores.

Table 1 Descriptive statistics for recruited patients

| Patient characteristics | |
|---|-------------------------------------|
| Number of patients | 218 |
| Gender (female) | 66.66% |
| Age years, mean (SD); median (range) | 48.06 (14.39); 48 (17–85) |
| MS subtype | |
| Relapsing-remitting MS | 64.8% |
| Primary progressive MS | 11.0% |
| Secondary progressive MS | 24.2% |
| Disease duration (years), mean (SD); median (range) | 13.92 (10.47); 11 (0–53) |
| EDSS, mean (SD); median (range) | 4.19 (2.11); 4 (0–8.00) |
| MAM-36 Total score, mean (SD); median (range) | 126.46 (20.72); 131.00 (39–144) |
| NHPT (seconds Right), mean (SD); median (range) | 30.39 (23.96); 23.75 (11.40–180.00) |
| NHPT (Left), mean (SD); median (range) | 34.10 (25.18); 29.41 (12.40–180.00) |

As regards item functioning, four items showed Infit and Outfit mean-square values sitting either above the upper limit (item 18: “Use a remote control”, item 31: “Write 3 to 4 sentences legibly”), or below the lower limit (item 1: “Eat a sandwich”; item 20: “Turn door knob to open a door”) for productive measurement. Interestingly, all the items showing problematic fit assessed prevalently unimanual activities. Because Infit and Outfit statistics for these items did not exceed the 2.0 threshold, we decided not to drop the items from the scale.

Results of the PCA performed on Rasch residuals indicated the Rasch dimensions explained 51.1% of the variance in the data. The eigenvalue for the first contrast was 2.9, thus exceeding the proposed criterion and suggesting the existence of a secondary dimension in the data with roughly the strength of three items. Items 22 (“Carry a shopping bag with a hand loop”) and 34 (“Use a hammer or screwdriver”) revealed positive loadings on the contrast, while items 32 (“Turn pages of a book”) and 36 (“Take a CD/DVD out of its case and put it into a player/drive”) showed negative loadings, possibly suggesting patients had different familiarity with activities involving carrying heavy loads. However, the first contrast accounted for only 3.8% and 8.6% respectively of the total and unexplained variance, which are both below suggested criteria, indicating no substantial violation of scale unidimensionality.

Table 2 MAM-36 response categories: number of responses, estimated Rasch measure and fit statistics

| Response category | Step measure | SE | Number of response | Infit | Outfit |
|-------------------|--------------|------|--------------------|-------|--------|
| 1 Cannot do it* | - | - | 184 | 1.04 | 1.10 |
| 2 Very hard | - 1.33 | 0.09 | 554 | 1.02 | 1.29 |
| 3 A little hard | - 0.20 | 0.05 | 1709 | 0.95 | 0.83 |
| 4 Easy | 1.54 | 0.03 | 3709 | 1.03 | 1.01 |

*Reference category

Distribution of the MAM-36 score

In our sample, the MAM-36 raw score scale had a mean value of 126.46 (SD = 18.75; Skewness = - 1.56), a median value of 131, with an observed range of 39–144 (see Table 1). We found indication of significant ceiling effect (but no floor effect), as more than 15% of the patients (18% of patients, $N = 40$) had the highest possible score on the scale. Coherently, based on 25th, 50th, and 75th percentiles computed for the MAM-36 score (25th pct = 116.75; 50th pct = 131.00; 75th pct = 141.25), most of the patients appear to be clustered in the upper range of the score distribution.

Criterion validity of the MAM-36

Results of correlation between study variables showed the MAM-36 scores had a weak negative correlation with age ($r = - 0.14$, $p < .05$) and disease duration ($r = - 0.25$, $p < .01$), and a moderate negative correlation with EDSS ($r = - 0.47$, $p < .01$), and the 9-HPT scores for both arms (right: $r = - 0.35$, $p < .01$; left: $r = - 0.30$, $p < .01$). There was no significant association between MAM-36 score and gender ($p = .72$). Results indicated significant differences existed in MAM-36 scores among patients with different disease course (relapsing-remitting: $M = 129.05$; secondary progressive: $M = 121.64$; primary progressive: $M = 121.58$, KW [2] = 8.75,

Table 3 MAM-36 items: one-factor CFA loadings, Rasch difficulty parameters and fit statistics

| Item | One-factor CFA | | Rasch model | | | |
|--|---------------------------|------|-------------------------|------|-------|--------|
| | Std. loading ^a | SE | Difficulty ^b | SE | Infit | Outfit |
| 22 Carry a shopping bag with a hand loop. <i>Portare una borsa della spesa con i manici.</i> | 0.74 | 0.04 | 1.59 | 0.11 | 1.30 | 1.31 |
| 16 Cut nails with a nail clipper. <i>Tagliarsi le unghie con un tagliaunghie.</i> | 0.84 | 0.02 | 1.38 | 0.11 | 1.04 | 0.93 |
| 34 Use a hammer or screwdriver. <i>Usare un martello o un cacciavite.</i> | 0.82 | 0.03 | 1.20 | 0.12 | 0.97 | 0.99 |
| 26 Open a medication bottle with child-proof top. <i>Aprire una bottiglia medicinale con chiusura a prova di bambino</i> | 0.77 | 0.03 | 1.17 | 0.12 | 1.04 | 1.20 |
| 17 Tie shoes with laces. <i>Allacciarsi le scarpe.</i> | 0.84 | 0.03 | 0.96 | 0.12 | 1.22 | 1.29 |
| 23 Open a previously opened wide-mouth jar (jam, pickle). <i>Aprire un barattolo grande precedentemente aperto (marmellata, sottaceti)</i> | 0.76 | 0.04 | 0.86 | 0.12 | 1.00 | 1.05 |
| 11 Wring a towel. <i>Strizzare un asciugamano.</i> | 0.73 | 0.04 | 0.82 | 0.12 | 1.16 | 1.21 |
| 28 Peel vegetables or fruits. <i>Pelare le verdure o la frutta.</i> | 0.83 | 0.03 | 0.75 | 0.12 | 1.00 | 1.05 |
| 3 Pick up a half full water pitcher. <i>Sollevarre una brocca d'acqua piena per metà.</i> | 0.73 | 0.04 | 0.57 | 0.13 | 1.09 | 1.26 |
| 31 Write 3 to 4 sentences legibly. <i>Scrivere 3 o 4 frasi leggibili.</i> | 0.65 | 0.05 | 0.53 | 0.13 | 1.50 | 1.88 |
| 24 Open a previously unopened carton box (milk or cereal). <i>Aprire un contenitore di cartone mai aperto (latte o cereali).</i> | 0.86 | 0.02 | 0.51 | 0.13 | 0.73 | 0.73 |
| 6 Cut meat on a plate with a knife. <i>Tagliare la carne in un piatto con il coltello.</i> | 0.84 | 0.03 | 0.48 | 0.13 | 0.96 | 0.84 |
| 15 Fasten a clothes snap or hook. <i>Chiudere il bottone automatico o il gancio di un vestito.</i> | 0.85 | 0.02 | 0.48 | 0.13 | 0.82 | 0.74 |
| 33 Shuffle and deal cards. <i>Distribuire le carte.</i> | 0.82 | 0.03 | 0.44 | 0.13 | 0.90 | 1.03 |
| 27 Open an envelop without a letter opener <i>Aprire una busta senza il tagliacarte.</i> | 0.72 | 0.04 | 0.31 | 0.13 | 1.15 | 1.37 |
| 35 Fold clothes after laundering. <i>Piegare i vestiti dopo il lavaggio.</i> | 0.85 | 0.03 | 0.28 | 0.14 | 0.93 | 0.94 |
| 25 Pour liquid from a bottle into a glass. <i>Versare un liquido da una bottiglia in un bicchiere.</i> | 0.82 | 0.03 | 0.27 | 0.13 | 0.92 | 0.83 |
| 29 Count/handle money (bills and coins). <i>Contare i soldi (banconote o monete).</i> | 0.85 | 0.03 | 0.22 | 0.13 | 0.90 | 0.77 |
| 14 Button clothes. <i>Abbottonarsi i vestiti.</i> | 0.88 | 0.02 | 0.20 | 0.14 | 0.86 | 0.71 |
| 12 Zip pants. <i>Allacciarsi i pantaloni.</i> | 0.89 | 0.02 | 0.07 | 0.14 | 0.86 | 0.70 |
| 13 Zip a jacket. <i>Chiudere la cerniera di una giacca.</i> | 0.91 | 0.02 | 0.07 | 0.14 | 0.88 | 0.72 |
| 36 Take a CD/DVD out of its case and put it into a player/drive. <i>Prendere un CD/DVD dalla sua custodia e inserirlo nel lettore.</i> | 0.90 | 0.02 | 0.05 | 0.14 | 0.80 | 0.61 |
| 30 Take things out of a wallet (bills, papers, credit cards). <i>Prendere cose dal portafogli (banconote, documenti, carte di credito).</i> | 0.89 | 0.02 | -0.12 | 0.14 | 0.68 | 0.72 |
| 32 Turn pages of a book. <i>Girare le pagine di un libro.</i> | 0.78 | 0.04 | -0.29 | 0.15 | 0.98 | 1.02 |
| 21 Turn key to open a lock. <i>Girare una chiave per aprire una serratura.</i> | 0.76 | 0.04 | -0.31 | 0.15 | 1.08 | 1.19 |
| 5 Butter bread (Put butter or jam on the bread). <i>Spalmare sul pane (mettere burro o marmellata sul pane).</i> | 0.84 | 0.03 | -0.40 | 0.15 | 0.85 | 0.76 |
| 7 Squeeze toothpaste. <i>Spremere un tubetto di dentifricio.</i> | 0.78 | 0.04 | -0.72 | 0.16 | 0.83 | 0.84 |
| 9 Brush or comb hair. <i>Pettinarsi o spazzolarsi i capelli.</i> | 0.75 | 0.05 | -0.81 | 0.17 | 0.94 | 0.87 |
| 4 Use a spoon or fork. <i>Usare un cucchiaino o una forchetta.</i> | 0.76 | 0.05 | -0.82 | 0.17 | 0.95 | 0.82 |
| 20 Turn door knob to open a door. <i>Girare la maniglia per aprire la porta.</i> | 0.87 | 0.03 | -1.06 | 0.18 | 0.85 | 0.57 |

Table 3 (continued)

| Item | One-factor CFA | | Rasch model | | | |
|---|---------------------------|------|-------------------------|------|-------|--------|
| | Std. loading ^a | SE | Difficulty ^b | SE | Infit | Outfit |
| 19 Key in telephone numbers. <i>Digitare i numeri di telefono.</i> | 0.76 | 0.05 | −1.23 | 0.19 | 1.01 | 0.83 |
| 18 Use a remote control. <i>Usare un telecomando</i> | 0.63 | 0.06 | −1.30 | 0.19 | 1.49 | 1.97 |
| 10 Wash hands. <i>Lavarsi le mani.</i> | 0.58 | 0.07 | −1.39 | 0.20 | 1.40 | 1.11 |
| 8 Brush teeth. <i>Lavarsi i denti.</i> | 0.77 | 0.05 | −1.43 | 0.20 | 1.00 | 0.84 |
| 2 Drink a glass of water. <i>Bere un bicchiere d'acqua.</i> | 0.74 | 0.05 | −1.50 | 0.20 | 1.07 | 0.89 |
| 1 Eat a sandwich. <i>Mangiare un panino.</i> | 0.82 | 0.05 | −1.84 | 0.23 | 1.01 | 0.56 |

For each item, Italian translations are reported in italic

^aStandardised loadings are reported

^bItems are reported in decreasing difficulty order

$p = .01$); post hoc analyses showed that only relapsing-remitting and secondary progressive PwMS were statistically different ($p = .01$).

We also found a significant difference that existed in MAM-36 score when comparing patients with different levels of upper limb impairment (no overt impairment: $M = 131.22$; unilateral overt impairment: $M = 127.55$; bilateral overt impairment: $M = 123.28$; KW [2] = 10.92, $p < .01$). However, only patients with bilateral overt impairment showed significantly different scores than patients reporting no overt impairment on both arms ($p < .01$).

Discussion

The present study had multiple aims. First, the adaptation to Italian of the MAM-36 and the investigation of its psychometric properties, including dimensionality, rating scale and item functioning, and targeting to a sample of MS patients. Next, we examined construct validity of the MAM-36 by investigating its association with theoretically correlated demographic and clinical criteria.

As regards the investigation of the psychometric properties of the instrument, the CFA and Rasch model analyses supported the unidimensionality of the scale, score reliability, and adequate functioning of response categories. Still, Rasch analyses showed a few items did not fully comply with proposed standards, thus potentially introducing undesired noise in the measurement process. Interestingly, all items showing problematic functioning assessed unimanual activities, as opposed to bimanual activities, which instead represent the majority of tasks

assessed by MAM-36 items. Overall, given the relatively low number of items showing problematic functioning, and low item bias, we decided not to remove the items from the scale. Still, future studies employing large samples of MS patients should consider examining the feasibility of producing independent scores for unimanual and bimanual MAM-36 tasks, as well as exploring the use of alternative models including secondary ability dimensions [37].

As regards the targeting to the sample of the scale, both the Rasch analysis and the examination of score distribution showed the scale failed to target patients with low disability, a finding that has been already reported by other authors [18]. Based on the comparison of the Rasch item difficulty and person ability measures, most of the items seemed to cluster at the lower end of the ability spectrum, and none of the included items targeted individuals showing low disability. Coherently, the scale also showed indications of ceiling effect. Combination of these findings suggests that the scale might be more suitable to assess upper limb impairment in patients with moderate-to-high disability. For use in unselected MS population, a revision of the scale is advisable.

Lastly, criterion validity of MAM-36 scores was examined by investigating its association with clinician-rated EDSS score and 9-HPT scores for both arms, and results indicated a moderate convergence between the MAM-36 and these measures. As expected, the MAM-36 score also negatively correlated with age and disease duration and was significantly higher among patients with secondary progressive MS than among relapsing-remitting patients.

Overall, findings from the present study indicate that MAM-36 shows adequate fit to Rasch assumptions and

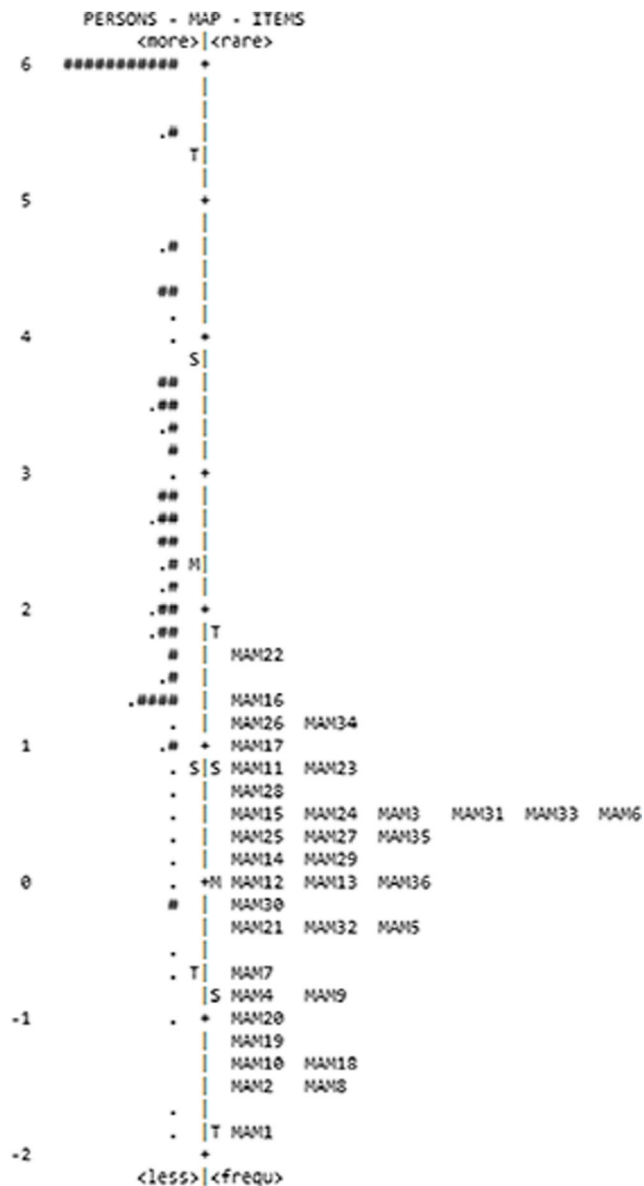


Fig. 1 Rasch item-person map for the MAM-36

provides a reliable assessment for upper limb disability for MS patients, in particular in the medium-to-high area of the disability continuum.

Limitations

The present study is not without limitations. Primarily, at 218 patients, sample size for the study was small, limiting the robustness of our findings. However, our sample size sits just above the 200 observations threshold typically suggested for conducting CFA analyses [38], and for stable parameter estimates in Rasch analyses [39]. Further, it is worthy to note that Rasch mean-square fit statistics are expected to provide reliable information even at low sample sizes [40]. Still, studies

performed on larger samples may help clarify further the functioning of the scale among PwMS.

Conclusions

In summary, this study indicates that the Italian adaptation of the MAM-36 provides an assessment of upper limb function in MS that shows good measurement properties. The scale demonstrated excellent score reliability, scale unidimensionality and a well-functioning rating scale. Still, in line with previous findings, the scale showed indications of problematic targeting to patients with low disability. For this reason, use of the scale appears to be more suitable among patients with moderate-to-severe disability.

Authors' contributions CS conceived the study and designed the study protocol. CS, DM, DC, RDG, EG, BG, MM, AT, RB, FP, AP, LP, LC and RR recruited the subjects and carried out the experiments. DM, DC and RR analysed the results. CS, DM, DC, RDG and EG wrote the manuscript. All authors have read and approved the final version of the manuscript.

Funding information This work was partly supported by a Grant 2014/R20 from the Italian Multiple Sclerosis Foundation (FISM).

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflicts of interest.

Ethical approval None.

References

- McDonald I, Compston A (2006) The symptoms and signs of multiple sclerosis. In: Compston A, Ebers G, Lassmann H (eds) *McAlpine's multiple sclerosis*, 4th edn. Curchill Livingstone, London, pp 287–346
- Johansson S, Ytterberg C, Claesson IM, Lindberg J, Hillert J, Andersson M, Widén Holmqvist L, von Koch L (2007) High concurrent presence of disability in multiple sclerosis. Associations with perceived health. *J Neurol* 254(6):767–773
- Lamers I, Feys P (2014) Assessing upper limb function in multiple sclerosis. *Mult Scler* 20:775–784
- Yalachkov Y, Soydaş D, Bergmann J, Frisch S, Behrens M, Foerch C, Gehrig J (2019) Determinants of quality of life in relapsing-remitting and progressive multiple sclerosis. *Mult Scler Relat Disord* May 30:33–37
- Kratz AL, Braley TJ, Foxen-Craft E, Scott E, Murphy JF 3rd, Murphy SL (2017) How do pain, fatigue, depressive, and cognitive symptoms relate to well-being and social and physical functioning in the daily lives of individuals with multiple sclerosis? *Arch Phys Med Rehabil*. Nov;98(11):2160–2166
- Yozbatiran N, Baskurt F, Baskurt Z, Ozakbas S, Idiman E (2006) Motor assessment of upper extremity function and its relation with fatigue, cognitive function and quality of life in multiple sclerosis patients. *J Neurol Sci* 246(1-2):117–122

7. Bertoni R, Lamers I, Chen CC, Feys P, Cattaneo D (2015) Unilateral and bilateral upper limb dysfunction at body functions, activity and participation levels in people with multiple sclerosis. *Mult Scler* 21(12):1566–1574
8. Kurtzke JF (1983) Rating neurologic impairment in multiple sclerosis: an expanded disability status scale (EDSS). *Neurology* 33(11):1444–1452
9. Feys P, Lamers I, Francis G et al (2017) The Nine-Hole Peg Test as a manual dexterity performance measure for multiple sclerosis. *Mult Scler* 23(5):711–720
10. Kierkegaard M, Einarsson U, Gottberg K, von Koch L, Holmqvist LW (2012) The relationship between walking, manual dexterity, cognition and activity/participation in persons with multiple sclerosis. *Mult Scler* 18(5):639–646
11. Solaro C, Cattaneo D, Bricchetto G, Castelli L, Tacchino A, Gervasoni E, Prosperini L (2019) Clinical correlates of 9-Hole Peg Test in a large population of people with multiple sclerosis. *Mult Scler Relat Disord* 30:1–8
12. Lamers I, Kelchtermans S, Baert I, Feys P (2014) Upper limb assessment in multiple sclerosis: a systematic review of outcome measures and their psychometric properties. *Arch Phys Med Rehabil* 95(6):1184–1200
13. Penta M, Thonnard JL, Tesio L (1998) ABILHAND: a Rasch-built measure of manual ability. *Arch Phys Med Rehabil* 79(9):1038–1042
14. Bricchetto G, Bragadin MM, Fiorini S, Battaglia MA, Konrad G, Ponzio M, Pedullà L, Verri A, Barla A, Tacchino A (2019) The hidden information in patient-reported outcomes and clinician-assessed outcomes: multiple sclerosis as a proof of concept of a machine learning approach. *Neurol Sci*:1–4
15. Kahraman T (2018) Performance measures for upper extremity functions in persons with multiple sclerosis. *Arch Neuropsychiatry* 55(Suppl 1):S41–S45
16. Poole JL, Huffman M, Hunter A, Mares C, Siegel P (2015) Reliability and validity of the Manual Ability Measure-36 in persons with Charcot-Marie-Tooth disease. *J Hand Ther* 28(4):364
17. Tsai CL, Lin CF, Lin HT, Liu MF, Chiu HY, Hsu HY, Kuo LC (2017) How kinematic disturbance in the deformed rheumatoid thumb impacts on hand function: a biomechanical and functional perspective. *Disabil Rehabil* 39(4):338–345
18. Chen CC, Bode RK (2010) Psychometric validation of the Manual Ability Measure-36 (MAM-36) in patients with neurologic and musculoskeletal disorders. *Arch Phys Med Rehabil* 91:414–420
19. Rallou CR, Chen CC (2008) Relationship between performance-based and self-reported assessment of hand function. *Am J Occup Ther*:62
20. Chen CC, Kasven N, Karpatkin HI, Sylvester A (2007) Hand strength and perceived manual ability among patients with multiple sclerosis. *Arch Phys Med Rehabil* 88(6):794–797
21. Polman CH, Reingold SC, Banwell B, Clanet M, Cohen JA, Filippi M et al (2011) Diagnostic criteria for multiple sclerosis: 2010 revisions to the McDonald criteria. *Ann Neurol* 69(2):292–302
22. Snijders T, Bosker R (2012) *Multilevel analysis: an introduction to basic and advanced multilevel modeling*, 2nd edn. Sage, Thousand Oaks
23. Lai MHC, Kwok O (2015) Examining the rule of thumb of not using multilevel modeling: the “design effect smaller than two” rule. *J Exp Educ* 83:423–438
24. Hu L, Bentler PM (1999) Cutoff criteria for fit indexes in covariance structure analysis: conventional criteria versus new alternatives. *Struct Equ Model Multidiscip J* 6(1):1–55
25. Hair JF Jr, Black WC, Babin BJ, Anderson RE (2009) *Multivariate data analysis*, 7th edn. Prentice Hall, Upper Saddle River, p 761
26. McDonald RP (1985) *Factor analysis and related methods*. Lawrence Erlbaum Associates, Inc, Hillsdale
27. Fornell C, Larcker DF (1981) Evaluating structural equation models with unobservable variables and measurement error. *J Mark Res* 18(1):39–50
28. Andrich D (1978) A rating formulation for ordered response categories. *Psychometrika* 43:561–573
29. Linacre JM (2002a) Optimizing rating scale category effectiveness. *J Appl Meas* 3(1):85–106
30. Bond TG, Fox CM (2007) *Applying the Rasch model: fundamental measurement in the human sciences*. 2nd. Lawrence Erlbaum Associates Publishers. Chapter 12
31. Linacre JM (2002b) What do Infit and Outfit, mean-square and standardized mean? *Rasch Measurement Transactions* 16:878
32. Bland JM, Altman DG (1997) Cronbach’s alpha. *BMJ* 314(7080):572
33. Fisher WJ (1992) Reliability, separation, strata statistics. *Rasch Meas Trans* 6:328
34. Linacre JM (2006) Data variance explained by Rasch measures. *Rasch Measurement Transactions* 20(1):1045
35. Smith EV Jr (2002) Detecting and evaluating the impact of multidimensionality using item fit statistics and principal component analysis of residuals. *J Appl Meas* 3(2):205–231
36. Erasmus LP, Sarno S, Albrecht H, Schwecht M, Pöllmann W, König N (2001) Measurement of ataxic symptoms with a graphic tablet: standard values in controls and validity in multiple sclerosis patients. *J Neurosci Methods* 108(1):25–37
37. Marengo D, Miceli R, Settanni M (2016) Do mixed item formats threaten test unidimensionality? Results from a standardized math achievement Test. *TPM: Testing, Psychometrics, Methodology in Applied Psychology* 23:25–36
38. Bandalos DL (2014) Relative performance of categorical diagonally weighted least squares and robust maximum likelihood estimation. *Struct Equ Model* 21:102–116
39. Suen HK (1990) *Principles of test theories*. Lawrence Erlbaum, Hillsdale
40. Smith AB, Rush R, Fallowfield LJ, Velikova G, Sharpe M (2008) Rasch fit statistics and sample size considerations for polytomous data. *BMC Med Res Methodol* 8(1):33

Publisher’s note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.