Original Paper

Italian Version of the mHealth App Usability Questionnaire (Ita-MAUQ): Translation and Validation Study in People With Multiple Sclerosis

Jessica Podda[1], PhD; Erica Grange[1], OT; Alessia Susini[1], MPsy; Andrea Tacchino[1], PhD; Federica Di Antonio[1], MPhys; Ludovico Pedullà[1], PhD; Giampaolo Brichetto[1,2], MD, PhD; Michela Ponzio[1], PhD

[1]Scientific Research Area, Italian Multiple Sclerosis Foundation, Genoa, Italy
[2]Rehabilitation Service of Genoa, Italian Multiple Sclerosis Society, Genoa, Italy

Corresponding Author:

Jessica Podda, PhD
Scientific Research Area
Italian Multiple Sclerosis Foundation
Via Operai 40
Genoa, 16149
Italy
Phone: 39 0102713306
Email: jessica.podda@aism.it

Abstract

Background: Telemedicine and mobile health (mHealth) apps have emerged as powerful tools in health care, offering convenient access to services and empowering participants in managing their health. Among populations with chronic and progressive disease such as multiple sclerosis (MS), mHealth apps hold promise for enhancing self-management and care. To be used in clinical practice, the validity and usability of mHealth tools should be tested. The most commonly used method for assessing the usability of electronic technologies are questionnaires.

Objective: This study aimed to translate and validate the English version of the mHealth App Usability Questionnaire into Italian (ita-MAUQ) in a sample of people with MS.

Methods: The 18-item mHealth App Usability Questionnaire was forward- and back-translated from English into Italian by an expert panel, following scientific guidelines for translation and cross-cultural adaptation. The ita-MAUQ (patient version for stand-alone apps) comprises 3 subscales, which are ease of use, interface and satisfaction, and usefulness. After interacting with DIGICOG-MS (Digital Assessment of Cognitive Impairment in Multiple Sclerosis), a novel mHealth app for cognitive self-assessment in MS, people completed the ita-MAUQ and the System Usability Scale, included to test construct validity of the translated questionnaire. Confirmatory factor analysis, internal consistency, test-retest reliability, and construct validity were assessed. Known-groups validity was examined based on disability levels as indicated by the Expanded Disability Status Scale (EDSS) score and gender.

Results: In total, 116 people with MS (female n=74; mean age 47.2, SD 14 years; mean EDSS 3.32, SD 1.72) were enrolled. The ita-MAUQ demonstrated acceptable model fit, good internal consistency (Cronbach α=0.92), and moderate test-retest reliability (intraclass coefficient correlation 0.84). Spearman coefficients revealed significant correlations between the ita-MAUQ total score; the ease of use (5 items), interface and satisfaction (7 items), and usefulness subscales; and the System Usability Scale (all *P* values <.05). Known-group analysis found no difference between people with MS with mild and moderate EDSS (all *P* values >.05), suggesting that ambulation ability, mainly detected by the EDSS, did not affect the ita-MAUQ scores. Interestingly, a statistical difference between female and male participants concerning the ease of use ita-MAUQ subscale was found (*P*=.02).

Conclusions: The ita-MAUQ demonstrated high reliability and validity and it might be used to evaluate the usability, utility, and acceptability of mHealth apps in people with MS.

JMIR Hum Factors 2024;11:e58079; doi: 10.2196/58079

Keywords: mHealth; multiple sclerosis; cognitive assessment; questionnaire validation; usability; mHealth app; mHealth application; validation study; MAUQ; app usability; telemedicine; disability; usability questionnaire; mobile health

Introduction

Telemedicine has enabled convenient and effective visits; reduced unnecessary testing and referrals; maintained good perception of care; reduced travel costs and caregiver burden; and, not least, helped health care providers to manage an ever-increasing volume of information and relationships [1,2]. Growing use of smartphones and tablets has made mobile health (mHealth) apps promising tools for empowering and engaging people in the self-management of their own health [3]. mHealth tools create opportunities to deliver new forms of health care and to expand services without the need to increase the existing workforce. For example, medical apps can be used within various domains such as wellness management, behavior change, health data collection, disease management, self-diagnosis, and rehabilitation as well as act as an electronic patient portal and medication reminder [4,5], leading to greater time spent at home and fewer medical visits at the center [6]. Furthermore, they represent useful solutions for participants with chronic and progressive diseases, such as multiple sclerosis (MS), that require continuous assistance and care.

mHealth apps promise to offer alternative methods for enhanced real-time data capture to screen for, monitor, and treat the heterogeneous symptoms in MS, thus favoring a substantial transformation in traditional paradigms of medicine [7-13]. However, as the number of mHealth apps increases, the demand for scientific evaluation of these solutions is strongly recommended as well [14]. Despite the growing popularity of mHealth apps, the amount of usability reports does not correlate with the number of published digital health implementation studies [15]. For instance, while Salimzadeh and colleagues [3] found 104 MS-related apps in iTunes (Apple Inc) and Google Play (Google LLC), they noted that there was no corresponding evidence regarding the usability and utility of these solutions in people with MS. To be used in clinical practice, the validity and usability of mHealth tools should be tested. mHealth apps must be designed to ensure good usability, and they must be easy to use and able to reach their goals efficiently. As indicated by Wilson and Lankton [16], perceived ease of use and usefulness affect people's intention to adopt mHealth devices. Generally, a mobile app is considered to have good usability when (1) it is efficient, (2) users have a positive opinion about the app, (3) it is easy to learn, (4) it is easy to remember even after users have not used it for a while, and (5) it has a low error rate [17].

The mHealth apps can be grouped according to the nature of the interaction between patients and health care providers in the app: interactive and stand-alone mHealth apps. In interactive mHealth apps, users can send and receive information from their health care providers or patients via the app in a synchronous or asynchronous modality. In stand-alone mHealth apps, users enter, collect, or store health information about themselves or other people, which are not directly sent to the user's health care providers [18].

The most commonly used method for assessing the usability of electronic technologies are questionnaires. General and technology-independent questionnaires such as the System Usability Scale (SUS) [19] and the Post-Study System Usability Questionnaire [20] are usually used in usability studies of mHealth apps [15]. However, these questionnaires were created for general software systems and cannot reliably identify mHealth specific problems that may arise, for example, in health self-management or accessing health care services.

In this context, a new specific usability scale for evaluating the validity of mHealth apps was developed, the mHealth App Usability Questionnaire (MAUQ) [18]. The English version of MAUQ has been translated to various languages such as Malay [21], Chinese [22], Spanish [23], German [24] and French [25]. However, no literature was found reporting a translated version of the questionnaire in Italian, although it was used in 1 study on app usability [26]. Thus, this study aimed to translate and validate the English version of the mHealth App Usability Questionnaire into Italian (ita-MAUQ) in a sample of people with MS.

We specifically hypothesized that the ita-MAUQ would retain acceptable model fit in confirmatory factor analysis, acceptable levels of internal consistency, and test-retest reliability. We also hypothesized an acceptable construct validity, defined based on relations between the ita-MAUQ and another standardized usability scale, and differences in known-groups.

Methods

mHealth App Usability Questionnaire

The MAUQ was first developed by Zhou and colleagues [18]. MAUQ is designed for different users (patients or health care providers) and different interaction modes (interactive or stand-alone). For study purposes, the patient version for stand-alone mHealth apps was used. It consists of 18 items divided into 3 subscales: ease of use (5 items; MAUQ_E), interface and satisfaction (7 items; MAUQ_I), and usefulness (6 items; MAUQ_U). The overall Cronbach α coefficients of the original questionnaire were 0.85 for MAUQ_E, 0.90 for MAUQ_I, and 0.72 for MAUQ_U, which indicated strong internal consistency of the questionnaire [18]. The questionnaire uses a 7-point Likert scoring system: 1 (strongly disagree), 2 (disagree), 3 (somewhat disagree), 4 (neither agree nor disagree), 5 (somewhat agree), 6 (agree), and 7 (strongly agree). The authors point out that there are no licensing fees for using the questionnaire, and it is not necessary to request permission before using it. The questionnaire is freely accessible on the website.

MAUQ Translation and Cross-Cultural Adaptation

The original MAUQ questionnaire for stand-alone mHealth apps (patient version) was translated into the Italian language using a guideline for the translation, adaptation, and validation of instruments or scales for cross-cultural health care research [27]. The questionnaire was first translated by native Italian speakers proficient in English. The forward-translated versions of the instrument were initially compared by a third independent translator regarding ambiguities and discrepancies of words, sentences, and meanings to generate a preliminary initial translated version of the questionnaire. The back translation was verified by translating the Italian version to English by 2 native English speaker translators with a high Italian proficiency. A multidisciplinary panel with 3 health care professionals with expertise in MS (an occupational therapist, a physiotherapist, and a psychologist), 2 researchers with expertise in scale validation, and 1 expert in app development compared the translated versions with the original questionnaire and made modifications to make the questionnaire more understandable to Italian people with MS. The final ita-MAUQ questionnaire can be viewed in Multimedia Appendix 1.

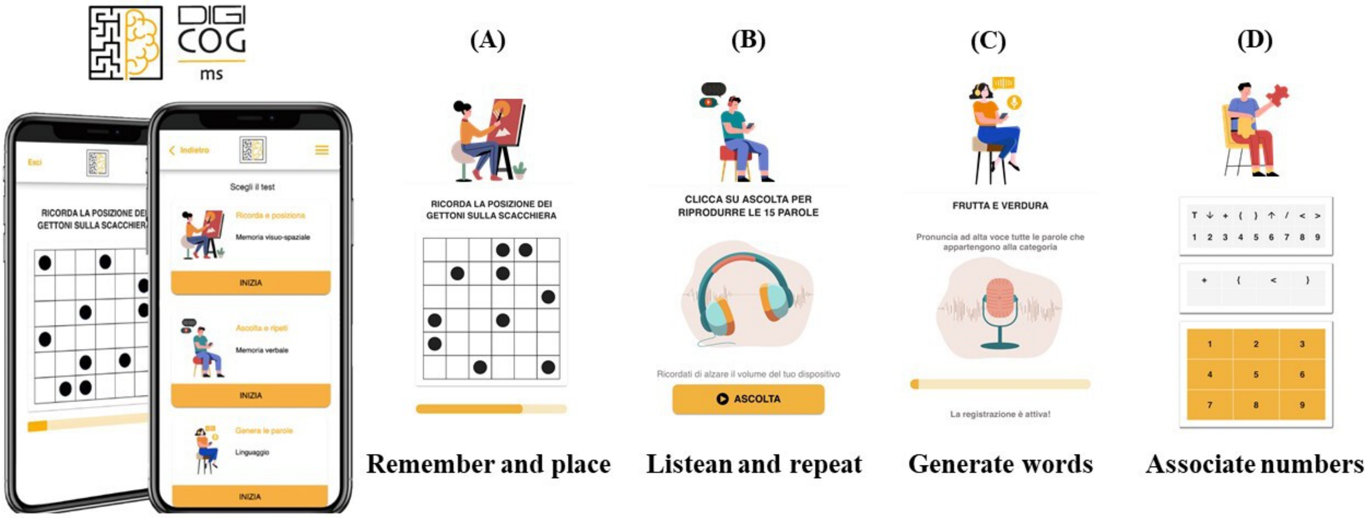
mHealth App Used for Validation of the Ita-MAUQ

In addition to motor and sensory difficulties, cognitive impairment, known as an invisible symptom, affects up to 65% of people with MS. Documented in all MS courses, with more severe deficits in progressive forms, both secondary progressive and primary progressive, compared to relapsing-remitting MS [28], cognitive impairment is recognized as one of the most disturbing disorders in MS, negatively affecting the quality of life and independence of people with MS. Attention, information processing speed, learning and memory, and executive functions seem to be the most commonly affected cognitive domains [29]. Consistent evidence indicates that cognitive functions in people with MS can be grouped into cognitive phenotypes, that is, subgroups of people with MS with a similar pattern of cognitive functioning [30-32]. The validation of the ita-MAUQ was conducted using DIGICOG-MS (Digital Assessment of Cognitive Impairment in Multiple Sclerosis), a smartphone- and tablet-based app for self-assessment of cognitive impairment in people with MS [33] (please see Figure 1 for an overview of the mHealth app). DIGICOG-MS (intellectual property of Italian Multiple Sclerosis Foundation; Italian Society of Authors and Publishers Registration ID: D000018162, 27-12-2022) includes 4 digital tests designed to evaluate the most affected cognitive domains in MS as visuospatial memory, verbal memory, semantic fluency, and information processing speed [34,35]:

- *Remember and place* assesses visuospatial episodic memory. A 36-square grid with 10 black checkers is displayed on the screen for 10 seconds. After the time elapses, the pattern disappears, and participants must reproduce it on a blank checkerboard. This replicates the 10/36 Spatial Recall Test [35], in which a 6 × 6 checkerboard with 10 pieces arranged in a particular pattern is shown to the participant for 10 seconds. Both tests (digital and traditional) include 3 consecutive trials, and the score consists of the total number of correct responses for the 3 trials.
- *Listen and repeat* was developed as an electronic version of the Rey Verbal Learning Test [36] that evaluates verbal memory. Participants listen to a prerecorded list of 15 common nouns and are asked to recall as many words as possible 5 times. Responses are recorded and then scored by the neuropsychologist. In the traditional test, words are read aloud to the participant who is asked to repeat as many words as possible in any order. All pronounced nouns in each of the 5 learning trials are transcribed by the neuropsychologist. For both versions of the test, the total score consists of the number of words recalled across the 5 trials.
- *Generate words* is a digital adaptation of the Word List Generation [35,37] and measures semantic verbal fluency. Participants generate a list of words, typically constrained by a specific semantic category, in 90 seconds. Recordings of pronounced words are processed by the neuropsychologist for scoring. In the traditional test, all words generated within the given semantic category are transcribed by the neuropsychologist. The total score is based on the number of correct words produced.

In a study by Podda and colleagues [33], correlation analysis was performed to determine the strength of the association between digital (ie, remember and place, listen and repeat, generate words, and associate numbers) and traditional (ie, 10/36 Spatial Recall Test, Rey Verbal Learning Test, Word List Generation, and Symbol Digit Modalities Test) tests. Overall, the findings revealed strong correlations between digital and traditional paper-based tests across all cognitive domains, with correlation coefficients (r) ranging from 0.58 to 0.78. Test-retest reliability was excellent for verbal memory and information processing speed (intraclass correlation coefficients [ICCs] ≥0.95) and good for visuospatial memory and semantic fluency (ICCs≥0.83).

Figure 1. Overview of DIGICOG-MS, the mobile health for cognitive assessment of people with multiple sclerosis. The 4 digital tests implemented in DIGICOG-MS that measure visuospatial memory (A), verbal memory (B), semantic fluency (C), and information processing speed (D). DIGICOG-MS: Digital Assessment of Cognitive Impairment in Multiple Sclerosis.



Study Participants

This study's participants were people with MS enrolled by the Italian MS Foundation and followed as outpatients at the Italian Multiple Sclerosis Society Rehabilitation Service in Genoa (Italy). Eligibility criteria were to be aged 18 years or older; have a confirmed MS diagnosis following the McDonald criteria [38]; have any disease course (relapsing-remitting MS, secondary progressive MS, and primary progressive MS); have not relapsed in the last 3 months; have an Expanded Disability Status Scale (EDSS) score [39] ≤7.5; and have adequate visual, hearing, and motor capabilities to work on a tablet. Exclusion criteria were a Montreal Cognitive Assessment score <18, neurological and major psychiatric illness, past serious head trauma, and alcohol or drug abuse. Data were collected from January 2023 to November 2023.

Study Procedure

Participants received a brief explanation of this study's process and goal at the Rehabilitation Service of the Italian Multiple Sclerosis Society in Genoa (Italy). People with MS were invited to use DIGICOG-MS. Particularly, they were required first to perform the 4 digital cognitive tests and then to explore other functionalities of the app (ie, how to log in, insert, and modify personal information; check for historical results of a specific test; consult the app tutorial; and ask for support). While the digital cognitive assessment was supervised by a neuropsychologist, people with MS were invited to navigate through DIGICOG-MS autonomously, and no specific instructions on how to use other functionalities of the app were given. After finishing the tasks, people with MS completed the ita-MAUQ and SUS [19,40]. SUS was included to test construct validity of the translated questionnaire. The 10-item SUS evaluates users' personal perceptions about how to use a given system or device, ranging from "strongly disagree" (1 point) to "strongly agree" (5 points). The SUS score is calculated by adding the individual scores and then multiplying that sum by 2.5. Thus, the SUS score ranges from zero (lowest usability) to 100 (highest usability), with a value of 68 considered above average. Although the SUS has already been translated and validated in many languages [41], individual problems that people may have when using mHealth apps are not specifically identified by the SUS.

Statistical Analysis

Categorical data were summarized by numbers and percentages, while numerical data were indicated by the mean and SD. Degree of education was coded as less than or equal to 12 years (primary school), between 13 and 15 years (high school), and equal to or more than 16 years (university).

Since confirmatory factor analysis (CFA) has emerged as a pivotal technique in such contexts, offering a comprehensive method for comparing the hypothesized measurement model structure with the observed one [42], it was conducted on the 18-item ita-MAUQ using the original 3 higher-order factors structure (ie, MAUQ_E, MAUQ_I, and MAUQ_U) [18]. CFA was used instead of "discovering" or exploring potential relationships between variables, as in exploratory factor analysis, because it is designed to test a predefined model based on consistent theoretical expectations [43]. Goodness-of-fit was tested with the ratio between chi-square (χ^2) and *df* (χ^2/df; good if ≤3), root mean square error of approximation (good if ≤0.08) and comparative fit index >0.9 [44]. We reported these statistics using the Satorra-Bentler adjustment because the scale item distribution was nonnormal [45]; the covariance of error terms was considered to improve the model fit. In the data analysis, missing data were replaced with a value of 4.

The internal consistency of the ita-MAUQ was assessed by calculating Cronbach α coefficient and average interitem correlation. The statistically acceptable Cronbach α coefficient should be >0.7 [46], and average interitem correlations should be between 0.30 and 0.70 [47].

The ICC (2-way analysis of variance random effect model for agreement) was calculated to assess the test-retest reliability. A very small sample size is required for estimating the desired value of ICC (especially when a researcher aims to estimate a very high value of ICC). Using power analysis calculations to test reliability between 2 different observations as described by Bujang and Baharum [48], a minimum of 22 participants is needed to have an acceptable ICC value ≥0.5 (α=.05, power=80%, n=2). The ICC was calculated on subscales and total scores, which are expected to remain stable. An ICC value of 0.70 was recommended as a minimum standard for reliability [49].

Ceiling and floor effects were calculated for the overall ita-MAUQ and its subscales. The floor and ceiling effects were defined as the percentages of respondents who reported the lowest score and the highest score, respectively. Floor and ceiling effects were considered present if >15% of participants achieved either the lowest or highest scores in ita-MAUQ and its subscales [50].

To examine the construct validity, Spearman correlations coefficients (ϱ), used for nonnormally distributed data, were calculated between the ita-MAUQ overall score and subscale scores, and the SUS. Spearman coefficients were considered low for ϱ<0.30, moderate for ϱ 0.30-0.59, and high for ϱ≥0.60 [51].

Known-groups validity evaluates whether an instrument can discriminate between known groups of people with MS that are expected to score differently on the measure of interest (ie, EDSS and gender). Here, it was assessed by comparing, with the Mann-Whitney *U* test, the total score and subscale scores of participants' groups with different levels of disability. Groups were defined using an EDSS cutoff value of 3.5, discriminating between people with MS with a mild and moderate disability: able (EDSS≤3.5) or unable (EDSS>3.5) to walk without aid or rest for more than 500 m. Furthermore, since previous research investigating gender differences in users' acceptance for website usability highlights gender as a key variable in understanding usage behavior in information and communication technology [5,52,53], ita-MAUQ total score and subscale scores were also compared with groups divided by participants' gender.

The *P* values <.05 were considered statistically significant. Data were analyzed using Stata (version 17; StataCorp).

Ethical Considerations

This study was approved by Regional Ethics Committee of Azienda Ospedaliera "San Martino" of Genoa (Italy N. 240/2022DB id 12354) and conducted according to the Declaration of Helsinki [54]. Before entering this study, participants had to read, complete, and sign an informed consent. Data collected were stored in an anonymized format to properly protect the privacy and confidentiality of participants, ensuring that no participant can be identified from the data provided. Participants have been informed that data collection could be used only for research purposes. They did not receive any compensation for taking part in the study

Results

In total, 116 people with MS (female: n=74; mean age 47.2, SD 14 years mean EDSS 3.32, SD 1.72) were enrolled (see Table 1).

Results from CFA of the 18 items-MAUQ order structure, defined by Zhou and colleagues [18], indicated acceptable fit for a 3D scale: χ^2/df=2.1, root mean square error of approximation 0.068, and comparative fit index 0.92.

Table 1. Demographic and clinical sample characteristics (N=116).

Characteristics	Value
Gender, n (%)	
Male	42 (36.2)
Female	74 (63.8)
Age (years)	
Mean (SD)	47.2 (14)
Range	19-70
Years of education, mean (SD)	13.3 (2.6)
Education, n (%)	
Primary school	16 (13.8)
High school	80 (69)
University	20 (17.2)
MS[a] duration (years)	
Mean (SD)	11.2 (9.6)
Range	0-32
MS course, n (%)	
RR[b]	91 (78.5)

Characteristics	Value
SP[c]	15 (12.9)
PP[d]	10 (8.6)
EDSS[e]	
Mean (SD)	3.32 (1.72)
Range	1-7.5
EDSS, n (%)	
Mild disability, score ≤3.5	67 (57.8)
Moderate disability, score >3.5	49 (42.2)

[a]MS: multiple sclerosis.
[b]RR: relapsing-remitting multiple sclerosis.
[c]SP: secondary progressive multiple sclerosis.
[d]PP: primary progressive multiple sclerosis.
[e]EDSS: Expanded Disability Status Scale.

The internal consistency of the overall ita-MAUQ and each subscale was good. The Cronbach α for the ita-MAUQ was 0.92, and those for the 3 subscales were 0.78 (ita-MAUQ_E), 0.89 (ita-MAUQ_I), and 0.87 (ita-MAUQ_U). Similarly, the average interitem correlation was between 0.41 and 0.53. These results align with published satisfactory thresholds for scale reliability [23,24].

To assess the test-retest reliability, 25 participants were required to complete the ita-MAUQ, 2 weeks apart. ICC was 0.84 (95% CI 0.66-0.92) for ita-MAUQ total score, 0.66 (95% CI 0.36-0.84) for ita-MAUQ_E, 0.88 (95% CI 0.75-0.94) for ita-MAUQ_I, and 0.67 (95% CI 0.38-0.84) for ita-MAUQ_U, showing good or moderate temporal stability.

There were no floor or ceiling effects for the ita-MAUQ total score (0% scored 18 and 6.9% scored 126). The floor effect for ita-MAUQ_E and ita-MAUQ_I subscales was again null, while for ita-MAUQ_U, only 1 participant had the worst possible score of 6 (0.9%). The ceiling effect for ita-MAUQ subscales was 38.8%, 36.2%, and 9.5%, respectively. No other previous study reported the measure of ceiling and floor effects of this instrument [22-24], and therefore a comparison cannot be established.

Spearman coefficients revealed significant correlations between ita-MAUQ total score; ita-MAUQ_E, ita-MAUQ_I, ita-MAUQ_U subscales; and SUS (85.43, SD 14.3; all *P* values <.05). Table 2 shows the results of the construct validity analysis.

Known-group analysis found no difference between people with MS with mild and moderate EDSS (all *P* values >.05), suggesting that ambulation ability, mainly detected by the EDSS, did not impact the ita-MAUQ scores (Table 3). Interestingly, statistical differences between female and male participants concerning the ita-MAUQ_E was found (*P*=.02; Table 4).

Table 2. Spearman correlation (ϱ) between the ita-MAUQ total[a] score; ita-MAUQ_E[b], ita-MAUQ_I[c], and ita-MAUQ_U[d] subscales; and SUS[e] total score.

Variable	Value, mean (SD)	SUS	ita-MAUQ_E	ita-MAUQ_I	ita-MAUQ_U	ita-MAUQ_tot
SUS	85.43 (14.3)					
r		—[f]				
P value		—				
ita-MAUQ_E	31.7 (4.1)					
r		0.54	—			
P value		<.001	—			
ita-MAUQ_I	44.3 (6.3)					
r		0.60	0.76	—		
P value		<.001	<.001	—		
ita-MAUQ_U	31.8 (6.8)					
r		0.26	0.38	0.48	—	
P value		.005	<.001	<.001	—	
ita-MAUQ_tot	107.8 (14.5)					
r		0.53	0.78	0.85	0.80	—
P value		<.001	<.001	<.001	<.001	—

[a]ita-MAUQ_tot: Italian version of the mHealth App Usability Questionnaire.

Variable	Value, mean (SD)	SUS	ita-MAUQ_E	ita-MAUQ_I	ita-MAUQ_U	ita-MAUQ_tot

[b]ita-MAUQ_E: ease of use subscale of the Italian version of the mHealth App Usability Questionnaire.
[c]ita-MAUQ_I: interface and satisfaction subscale of the Italian version of the mHealth App Usability Questionnaire.
[d]ita-MAUQ_U: usefulness subscale of the Italian version of the mHealth App Usability Questionnaire.
[e]SUS: System Usability Scale.
[f]Not applicable.

Table 3. Comparison of ita-MAUQ total[a] and subscale scores between people with multiple sclerosis with different disability levels.

Variable	ita-MAUQ_E[b], mean (SD)	ita-MAUQ_I[c], mean (SD)	ita-MAUQ_U[d], mean (SD)	ita-MAUQ_tot, mean (SD)
EDSS[e]≤3.5	31.9 (3.6)	44.1 (6.2)	31.0 (6.9)	107.0 (13.9)
EDSS>3.5	31.4 (4.7)	44.7 (6.7)	32.9 (6.6)	109.0 (15.5)
P value	.76	.28	.10	.24

[a]ita-MAUQ_tot: Italian version of the mHealth App Usability Questionnaire.
[b]ita-MAUQ_E: ease of use subscale of the Italian version of the mHealth App Usability Questionnaire.
[c]ita-MAUQ_I: interface and satisfaction subscale of the Italian version of the mHealth App Usability Questionnaire.
[d]ita-MAUQ_U: usefulness subscale of the Italian version of the mHealth App Usability Questionnaire.
[e]EDSS: Expanded Disability Status Scale.

Table 4. Comparison of ita-MAUQ total[a] and subscale scores between people with multiple sclerosis by gender.

Variable	ita-MAUQ_E[b] mean (SD)	ita-MAUQ_I[c] mean (SD)	ita-MAUQ_U[d] mean (SD)	ita-MAUQ_tot mean (SD)
Female	32.4 (3.6)	45.1 (6.0)	32.0 (7.2)	109.5 (14.3)
Male	30.5 (4.6)	43.1 (6.9)	31.4 (6.1)	104.9 (14.8)
P value	.02	.09	.51	.06

[a]ita-MAUQ_tot: Italian version of the mHealth App Usability Questionnaire.
[b]ita-MAUQ_E: ease of use subscale of the Italian version of the mHealth App Usability Questionnaire.
[c]ita-MAUQ_I: interface and satisfaction subscale of the Italian version of the mHealth App Usability Questionnaire.
[d]ita-MAUQ_U: usefulness subscale of the Italian version of the mHealth App Usability Questionnaire.

Discussion

Digital solutions as mHealth apps promise to offer alternative methods for enhanced real-time data capture to screen for, monitor, and treat symptoms in MS. These solutions may fundamentally shift traditional paradigms of medicine [7-12]. To be used in clinical practice, the validity and usability of mHealth tools should be tested. Questionnaires are the well-known methods for usability testing, but developing a new one might require concerted effort by the members of a research team, extra cost, and a lot of time [55]. Thus, adaptation of established, appropriate, and available questionnaires with documented validity in other languages is recommended [55].

Thus, the aim of this study was to translate and validate the English version of the ita-MAUQ in a sample of people with MS. Overall, findings demonstrated that the novel translated questionnaire ita-MAUQ is a reliable and valid measurement tool to assess the usability of mHealth apps for people with MS. In this context, people with MS self-administered the ita-MAUQ after interacting with DIGICOG-MS, a novel mHealth app for cognitive self-assessment in MS.

Results indicated that the ita-MAUQ had good internal consistency and stability, as indicated by the Cronbach α coefficient of 0.92. This is in line with the original version of the MAUQ [18] and with other translations of the same questionnaire [23,24,56].

Worldwide, Spearman coefficients between ita-MAUQ total score, subscales, and SUS were statistically significant, proving good criterion and construct validity. However, correlation between the SUS and ita-MAUQ_U was found to be low (0.26). This was in line with the study by Zhou et al [18], in which correlation between MAUQ_U and the SUS was 0.383, reflecting that MAUQ_U is mainly about the usefulness of apps for health care, which is an aspect not covered by the SUS.

Compared to another previous study on MAUQ that did not perform test-retest [24], results revealed good or moderate temporal stability. Given that mHealth apps may allow continuous health care services over time, identifying valid methods to test whether a digital tool is reliable in different measurements is crucial.

The ceiling effect of the ita-MAUQ subscores could indicate that the proposed mHealth app was indeed easy to use for many people with MS. Even though the interaction with DIGICOG-MS was supervised by a neuropsychologist to provide adequate responses to any questions from participants, they were invited to navigate through the mHealth app autonomously, since no specific instructions on how to use other app functionalities were given. In this way, we interpreted this ceiling effect positively as a successful interaction.

In general, the known-groups validity of the instrument was shown by the comparison of ita-MAUQ total and subscale scores between people with MS with different disability levels. The results indicate no significant difference in people with MS with mild and moderate EDSS. Since EDSS mainly measured ambulation capacity, this suggests that such mHealth apps could be found usable for both people that need assistance during ambulation and those who are able to walk without any aid or support.

Concerning gender, earlier studies found that perceived ease of use and usefulness technology may differ by gender [5,52]. Our results are in line with a previous study that demonstrated that gender was associated with higher usability scores in female than male participants [57]. Interestingly, male and female people with MS had different scores in ita-MAUQ_E, suggesting that women are more likely than men to be influenced by effort expectancy and facilitating conditions [53]. These results can help developers to enhance the usability of their services for all users with different personal and clinical characteristics, since men and women still have different traits and societal roles, which may affect their perceptions and usage of technologies.

Our study has several limitations. First, as a single-center study, participants' characteristics may limit the interpretation of our results. The study sample may be considered representative of those clinic-attending people with MS followed as outpatients in rehabilitation centers (ie, middle-age or older adults and with a longer disease duration) [58]. Thus, results may not generalize to other populations of participants with MS (eg, young and neo-diagnosed people). Second, in this validation study, we used DIGICOG-MS that was designed and developed to assess a specific symptom in MS, that is cognitive impairment. Given the high frequency of cognitive impairment in people with MS, it is reasonable to conclude that study participants who have experienced such a disturbing symptom tended to appreciate the mHealth app more compared to people with other neurological diseases. People with MS with cognitive impairment might find such digital tools particularly beneficial due to their potential to offer structured cognitive training and monitoring, which can enhance their daily functioning and quality of life [59]. Thus, it cannot be ruled out that the results of the ita-MAUQ validation would have been different with a clinical population with dissimilar characteristics.

Third, this study was conducted in a controlled clinical setting, allowing participants to familiarize themselves with both the novel technology and usability questionnaires' items. While this approach may have some limitations in terms of generalizability to real-life scenarios and may influence participants' engagement and perceived usability, having a facilitator available to assist, if needed, people with MS that could have problems with reading and interpreting the questionnaire items ensured they understand each question before responding. Furthermore, in this study we overlooked other key factors that may influence usability, such as assessing the importance of providing adequate training and continuous support to users or the role of previous experience with a similar technology; investigating how interface design that considering layout, navigation, and accessibility features makes the app user-friendly for participants with cognitive impairments; evaluating the concerns and preferences of users regarding data security and privacy, which are crucial for building trust and ensuring compliance with health regulations data security and privacy. Here, we did not collect additional feedback after the completion of both digital assessment and usability questionnaires from people with MS. Further study should include feedback sections where participants can indicate if they found any items difficult to understand, allowing for continuous improvement of the novel tools. Worldwide, incorporating these factors into future research can lead to the development of more effective, user-friendly, and impactful mHealth apps for people with MS.

In conclusion, the ita-MAUQ demonstrated high reliability and validity, and it might be used to evaluate the usability, utility, and acceptability of mHealth apps in people with MS. This finding is in line with previous validation of the MAUQ in different languages as Malay [21], Chinese [22], Spanish [23], German [24], and French [25], further confirming the cross-cultural validity, reliability, and adaptability of the MAUQ.

Acknowledgments

We would like to thank Maria Madera and Giulia Bignone for their help in participants' enrollment and all the people with MS that took part in our study. This work has been funded as Progetto Ricerca Speciale by Italian Multiple Sclerosis Foundation (2021/Special/005).

Data Availability

All data produced in this study are available upon reasonable request to the corresponding author.
The questionnaire can be used in clinical trials and is available free of charge for non-commercial parties, provided the authors are cited, or by contacting the corresponding author.

Authors' Contributions

JP conceived this study, interpreted results, and drafted this paper. MP performed statistical analysis, created tables, and contributed to the initial draft of this paper. AS performed data collection. FDA, AT, LP, and EG revised this paper. All authors read and approved the final paper. This paper was written entirely by humans. All authors are responsible and accountable for the originality, accuracy, and integrity of the work.

Conflicts of Interest

None declared.

Multimedia Appendix 1

The Italian version of the mHealth App Usability Questionnaire (ita-MAUQ; for stand-alone mHealth apps used by patients).
[DOCX File (Microsoft Word File), 24 KB-Multimedia Appendix 1]

References

1. Schwamm LH. Telehealth: seven strategies to successfully implement disruptive technology and transform health care. Health Aff. Feb 2014;33(2):200-206. [doi: 10.1377/hlthaff.2013.1021]
2. Lavorgna L, Brigo F, Moccia M, et al. e-Health and multiple sclerosis: an update. Mult Scler J. Nov 2018;24(13):1657-1664. [doi: 10.1177/1352458518799629] [Medline: 30231004]
3. Salimzadeh Z, Damanabi S, Kalankesh LR, Ferdousi R. Mobile applications for multiple sclerosis: a focus on self-management. Acta Inform Med. Mar 2019;27(1):12-18. [doi: 10.5455/aim.2019.27.12-18] [Medline: 31213737]
4. Kao CK, Liebovitz DM. Consumer mobile health apps: current state, barriers, and future directions. PM R. May 2017;9(5S):S106-S115. [doi: 10.1016/j.pmrj.2017.02.018] [Medline: 28527495]
5. Zhang X, Guo X, Lai KH, Guo F, Li C. Understanding gender differences in m-Health adoption: a modified theory of reasoned action model. Telemed J E Health. Jan 2014;20(1):39-46. [doi: 10.1089/tmj.2013.0092] [Medline: 24161004]
6. van Elburg FRT, Klaver NS, Nieboer AP, Askari M. Gender differences regarding intention to use mHealth applications in the Dutch elderly population: a cross-sectional study. BMC Geriatr. May 24, 2022;22(1):449. [doi: 10.1186/s12877-022-03130-3] [Medline: 35610577]
7. Beier M, Alschuler K, Amtmann D, Hughes A, Madathil R, Ehde D. ICAMS: assessing the reliability of a Brief International Cognitive Assessment for Multiple Sclerosis (BICAMS) tablet application. Int J MS Care. 2020;22(2):67-74. [doi: 10.7224/1537-2073.2018-108] [Medline: 32410901]
8. van Dongen L, Westerik B, van der Hiele K, et al. Introducing multiple screener: an unsupervised digital screening tool for cognitive deficits in MS. Mult Scler Relat Disord. Feb 2020;38:101479. [doi: 10.1016/j.msard.2019.101479] [Medline: 31760365]
9. Montalban X, Graves J, Midaglia L, et al. A smartphone sensor-based digital outcome assessment of multiple sclerosis. Mult Scler. Apr 2022;28(4):654-664. [doi: 10.1177/13524585211028561] [Medline: 34259588]
10. Hochstrasser C, Rieder S, Jufer-Riedi U, et al. Computerized symbol digit modalities test in a swiss pediatric cohort part 1: validation. Front Psychol. Apr 22, 2021;12:631536. [doi: 10.3389/fpsyg.2021.631536] [Medline: 33967898]
11. Costabile T, Signoriello E, Lauro F, et al. Validation of an iPad version of the Brief International Cognitive Assessment for Multiple Sclerosis (BICAMS). Mult Scler Relat Disord. Jun 2023;74:104723. [doi: 10.1016/j.msard.2023.104723] [Medline: 37086633]
12. Hsu WY, Rowles W, Anguera JA, et al. Assessing cognitive function in multiple sclerosis with digital tools: observational study. J Med Internet Res. Dec 30, 2021;23(12):e25748. [doi: 10.2196/25748] [Medline: 34967751]
13. Brichetto G, Pedullà L, Podda J, Tacchino A. Beyond center-based testing: understanding and improving functioning with wearable technology in MS. Mult Scler J. Sep 2019;25(10):1402-1411. [doi: 10.1177/1352458519857075] [Medline: 31502913]
14. McKay FH, Cheng C, Wright A, Shill J, Stephens H, Uccellini M. Evaluating mobile phone applications for health behaviour change: a systematic review. J Telemed Telecare. Jan 2018;24(1):22-30. [doi: 10.1177/1357633X16673538] [Medline: 27760883]
15. Maramba I, Chatterjee A, Newman C. Methods of usability testing in the development of eHealth applications: a scoping review. Int J Med Inform. Jun 2019;126:95-104. [doi: 10.1016/j.ijmedinf.2019.03.018] [Medline: 31029270]
16. Wilson EV, Lankton NK. Modeling patients’ acceptance of provider-delivered e-Health. J Am Med Inform Assoc. 2004;11(4):241-248. [doi: 10.1197/jamia.M1475] [Medline: 15064290]
17. Zhou L, Bao J, Parmanto B. Systematic review protocol to assess the effectiveness of usability questionnaires in mHealth app studies. JMIR Res Protoc. Aug 1, 2017;6(8):e151. [doi: 10.2196/resprot.7826] [Medline: 28765101]
18. Zhou L, Bao J, Setiawan IMA, Saptono A, Parmanto B. The mHealth App Usability Questionnaire (MAUQ): development and validation study. JMIR mHealth uHealth. Apr 11, 2019;7(4):e11500. [doi: 10.2196/11500] [Medline: 30973342]
19. Brooke J. SUS: a quick and dirty usability. In: Usability Evaluation in Industry. Taylor & Francis; 1996:189-194. URL: https://digital.ahrq.gov/sites/default/files/docs/survey/systemusabilityscale%2528sus%2529_comp%255B1%255D.pdf [Accessed 2024-08-20]
20. Lewis JR. Psychometric evaluation of the PSSUQ using data from five years of usability studies. Int J Human Comp Interaction. Sep 1, 2002;14(3):463-488. [doi: 10.1207/S15327590IJHC143&4_11]

21. Zubairi AM, Kassim NLA. Classical and Rasch analyses of dichotomously scored reading comprehension test items. Malaysian J ELT Res. Mar 2006;2:1-20. URL: https://meltajournals.com/index.php/majer/article/view/663/643 [Accessed 2024-08-24]
22. Shan Y, Ji M, Xie W, et al. Chinese version of the Mobile Health App Usability Questionnaire: translation, adaptation, and validation study. JMIR Form Res. Jul 6, 2022;6(7):e37933. [doi: 10.2196/37933] [Medline: 35793132]
23. Quifer-Rada P, Aguilar-Camprubí L, Gómez-Sebastià I, Padró-Arocas A, Mena-Tudela D. Spanish version of the mHealth App Usability Questionnaire (MAUQ) and adaptation to breastfeeding support apps. Int J Med Inform. Jun 2023;174:105062. [doi: 10.1016/j.ijmedinf.2023.105062] [Medline: 37037124]
24. Tacke T, Nohl-Deryk P, Lingwal N, et al. The German version of the mHealth App Usability Questionnaire (GER-MAUQ): translation and validation study in patients with cardiovascular disease. D Health. Jan 31, 2024;10:20552076231225168. [doi: 10.1177/20552076231225168] [Medline: 38303970]
25. Gagnon J, Probst S, Chartrand J, Lalonde M. mHealth App Usability Questionnaire for stand-alone mHealth apps used by health care providers: Canadian French translation, cross-cultural adaptation, and validation (part 1). JMIR Form Res. Feb 13, 2024;8:e50839. [doi: 10.2196/50839] [Medline: 38349710]
26. Renati R, Bonfiglio NS, Rollo D. Dealing with loved ones' addiction: development of an app to cope with caregivers' stress. Int J Environ Res Public Health. Nov 30, 2022;19(23):15950. [doi: 10.3390/ijerph192315950] [Medline: 36498025]
27. Sousa VD, Rojjanasrirat W. Translation, adaptation and validation of instruments or scales for use in cross-cultural health care research: a clear and user-friendly guideline. J Eval Clin Pract. Apr 2011;17(2):268-274. [doi: 10.1111/j.1365-2753.2010.01434.x] [Medline: 20874835]
28. Amato MP, Prestipino E, Bellinvia A. Identifying risk factors for cognitive issues in multiple sclerosis. Expert Rev Neurother. Apr 2019;19(4):333-347. [doi: 10.1080/14737175.2019.1590199] [Medline: 30829076]
29. Benedict RHB, Amato MP, DeLuca J, Geurts JJG. Cognitive impairment in multiple sclerosis: clinical management, MRI, and therapeutic avenues. Lancet Neurol. Oct 2020;19(10):860-871. [doi: 10.1016/S1474-4422(20)30277-5] [Medline: 32949546]
30. De Meo E, Portaccio E, Giorgio A, et al. Identifying the distinct cognitive phenotypes in multiple sclerosis. JAMA Neurol. Apr 1, 2021;78(4):414-425. [doi: 10.1001/jamaneurol.2020.4920] [Medline: 33393981]
31. Leavitt VM, Tosto G, Riley CS. Cognitive phenotypes in multiple sclerosis. J Neurol. Mar 2018;265(3):562-566. [doi: 10.1007/s00415-018-8747-5] [Medline: 29356970]
32. Podda J, Ponzio M, Pedullà L, et al. Predominant cognitive phenotypes in multiple sclerosis: insights from patient-centered outcomes. Mult Scler Relat Disord. Jun 2021;51:102919. [doi: 10.1016/j.msard.2021.102919] [Medline: 33799285]
33. Podda J, Tacchino A, Ponzio M, et al. Mobile health app (DIGICOG-MS®) for self-assessment of cognitive impairment in people with multiple sclerosis: instrument validation and usability study. JMIR Form Res. Jun 20, 2024;8:e56074. [doi: 10.2196/56074] [Medline: 38900535]
34. Goretti B, Patti F, Cilia S, et al. The Rao's Brief Repeatable Battery version B: normative values with age, education and gender corrections in an Italian population. Neurol Sci. Jan 2014;35(1):79-82. [doi: 10.1007/s10072-013-1558-7] [Medline: 24101117]
35. Amato MP, Portaccio E, Goretti B, et al. The Rao's Brief Repeatable Battery and Stroop Test: normative values with age, education and gender corrections in an Italian population. Mult Scler J. Dec 2006;12(6):787-793. [doi: 10.1177/1352458506070933] [Medline: 17263008]
36. Carlesimo GA, Caltagirone C, Gainotti G. The mental deterioration battery: normative data, diagnostic reliability and qualitative analyses of cognitive impairment. The group for the standardization of the mental deterioration battery. Eur Neurol. 1996;36(6):378-384. [doi: 10.1159/000117297] [Medline: 8954307]
37. Goretti B, Niccolai C, Hakiki B, et al. The Brief International Cognitive Assessment for Multiple Sclerosis (BICAMS): normative values with gender, age and education corrections in the Italian population. BMC Neurol. Sep 10, 2014;14:171. [doi: 10.1186/s12883-014-0171-6] [Medline: 25204350]
38. Polman CH, Reingold SC, Banwell B, et al. Diagnostic criteria for multiple sclerosis: 2010 revisions to the McDonald criteria. Ann Neurol. Feb 2011;69(2):292-302. [doi: 10.1002/ana.22366] [Medline: 21387374]
39. Kurtzke JF. Rating neurologic impairment in multiple sclerosis: an expanded disability status scale (EDSS). Neurology. Nov 1983;33(11):1444-1452. [doi: 10.1212/wnl.33.11.1444] [Medline: 6685237]
40. 5 ways to interpret a SUS score. MeasuringU. 2018. URL: https://measuringu.com/interpret-sus-score/ [Accessed 2024-08-17]
41. Gao M, Kortum P, Oswald FL. Multi-language toolkit for the System Usability Scale. Int J Hum Comput Interact. Dec 13, 2020;36(20):1883-1901. [doi: 10.1080/10447318.2020.1801173]

42. Rios J, Wells C. Evidencia de validez basada en la estructura interna. Psicothema. 2014;26:108-116. URL: https://go.gale.com/ps/i.do?id=GALE%7CA361242341&sid=googleScholar&v=2.1&it=r&linkaccess=abs&issn=02149915&p=IFME&sw=w&userGroupName=anon%7E4f35d9db&aty=open-web-entry [Accessed 2024-08-24]
43. Rogers P. Best practices for your confirmatory factor analysis: a JASP and lavaan tutorial. Behav Res Methods. Mar 13, 2024. [doi: 10.3758/s13428-024-02375-7] [Medline: 38480677]
44. Hu L, Bentler PM. Cutoff criteria for fit indexes in covariance structure analysis: conventional criteria versus new alternatives. Struct Equ Model Multidiscip J. Jan 1999;6(1):1-55. [doi: 10.1080/10705519909540118]
45. Satorra A, Bentler PM. Corrections to test statistics and standard errors in covariance structure analysis. In: Latent Variables Analysis: Applications for Developmental Research. Sage Publications; 1994:399-419. URL: https://psycnet.apa.org/record/1996-97111-016 [Accessed 2024-08-17]
46. Streiner DL, Norman GR, Cairney J. Health Measurement Scales: A Practical Guide to Their Development and Use. Oxford University Press; 2015. [doi: 10.1093/med/9780192869487.001.0001]
47. Nunnally JC, Bernstein IH. Psychometric Theory. McGraw-Hill; 1994. URL: https://books.google.co.in/books/about/Psychometric_Theory.html?id=r0fuAAAAMAAJ&redir_esc=y [Accessed 2024-08-20]
48. Bujang MA, Baharum N. A simplified guide to determination of sample size requirements for estimating the value of intraclass correlation coefficient: a review. Arch Orofac Sci. 2017;12:1-11. URL: https://www.researchgate.net/profile/Telmo-Firmino/post/Could_anyone_show_me_where_to_look_to_determine_the_sample_size_of_a_test-retest_reliability_research/attachment/5b14158a4cde260d15e2cede/AS%3A633471949090817%401528042890819/download/A+simplified+guide+to+determination+of+sample+size.pdf [Accessed 2024-08-20]
49. de Vet HCW, Terwee CB, Knol DL, Bouter LM. When to use agreement versus reliability measures. J Clin Epidemiol. Oct 2006;59(10):1033-1039. [doi: 10.1016/j.jclinepi.2005.10.015] [Medline: 16980142]
50. Terwee CB, Bot SDM, de Boer MR, et al. Quality criteria were proposed for measurement properties of health status questionnaires. J Clin Epidemiol. Jan 2007;60(1):34-42. [doi: 10.1016/j.jclinepi.2006.03.012] [Medline: 17161752]
51. Cohen J. Set correlation and contingency tables. Appl Psychol Meas. Dec 1988;12(4):425-434. [doi: 10.1177/014662168801200410]
52. Huang Z, Mou J. Gender differences in user perception of usability and performance of online travel agency websites. Technol Soc. Aug 2021;66:101671. [doi: 10.1016/j.techsoc.2021.101671]
53. Nunes A, Limpo T, Castro SL. Acceptance of mobile health applications: examining key determinants and moderators. Front Psychol. Dec 10, 2019;10:2791. [doi: 10.3389/fpsyg.2019.02791] [Medline: 31920836]
54. World Medical Association General Assembly. Declaration of Helsinki. Ethical principles for medical research involving human subjects. World Med J. 2008;54. URL: https://www.wma.net/policies-post/wma-declaration-of-helsinki-ethical-principles-for-medical-research-involving-human-subjects/ [Accessed 2024-08-24]
55. Mohamad Marzuki MF, Yaacob NA, Yaacob NM. Translation, cross-cultural adaptation, and validation of the Malay version of the System Usability Scale Questionnaire for the assessment of mobile apps. JMIR Hum Factors. May 14, 2018;5(2):e10308. [doi: 10.2196/10308] [Medline: 29759955]
56. Mustafa N, Safii NS, Jaffar A, et al. Malay version of the mHealth App Usability Questionnaire (M-MAUQ): translation, adaptation, and validation study. JMIR Mhealth Uhealth. Feb 4, 2021;9(2):e24457. [doi: 10.2196/24457] [Medline: 33538704]
57. Albrink K, Schröder D, Joos C, Müller F, Noack EM. Usability of an app for medical history taking in general practice from the patients' perspective: cross-sectional study. JMIR Hum Factors. Jan 5, 2024;11:e47755. [doi: 10.2196/47755] [Medline: 38180798]
58. Tacchino A, Brichetto G, Zaratin P, Battaglia MA, Ponzio M. Multiple sclerosis and rehabilitation: an overview of the different rehabilitation settings. Neurol Sci. Dec 2017;38(12):2131-2138. [doi: 10.1007/s10072-017-3110-7] [Medline: 28914385]
59. Waskowiak PT, de Jong BA, Uitdehaag BMJ, et al. Don't be late! Timely identification of cognitive impairment in people with multiple sclerosis: a study protocol. BMC Neurol. Jan 13, 2024;24(1):26. [doi: 10.1186/s12883-023-03495-x] [Medline: 38218777]

Abbreviations

CFA: confirmatory factor analysis
DIGICOG-MS: Digital Assessment of Cognitive Impairment in Multiple Sclerosis
EDSS: Expanded Disability Status Scale
ICC: intraclass correlation coefficient
ita-MAUQ: Italian version of the mHealth App Usability Questionnaire
ita-MAUQ_E: ease of use subscale of the Italian version of the mHealth App Usability Questionnaire
ita-MAUQ_I: interface and satisfaction subscale of the Italian version of the mHealth App Usability Questionnaire

ita-MAUQ_U: usefulness subscale of the Italian version of the mHealth App Usability Questionnaire
MAUQ: mHealth App Usability Questionnaire
MAUQ_E: ease of use subscale
MAUQ_I: interface and satisfaction subscale
MAUQ_U: usefulness subscale
mHealth: mobile health
MS: multiple sclerosis
SUS: System Usability Scale

Edited by Christine Jacob; peer-reviewed by Angelina Müller, Ozgul Ekmekci, Preetha Moorthy; submitted 05.03.2024; final revised version received 04.06.2024; accepted 07.06.2024; published 30.09.2024

Please cite as:
Podda J, Grange E, Susini A, Tacchino A, Di Antonio F, Pedullà L, Brichetto G, Ponzio M
Italian Version of the mHealth App Usability Questionnaire (Ita-MAUQ): Translation and Validation Study in People With Multiple Sclerosis
JMIR Hum Factors 2024;11:e58079
URL: https://humanfactors.jmir.org/2024/1/e58079
doi: 10.2196/58079