J Rehabil Med 2010; 42: 574–-581

ORIGINAL REPORT

RELIABILITY AND VALIDITY OF TELEPHONE ADMINISTRATION OF THE WHEELCHAIR OUTCOME MEASURE FOR MIDDLE-AGED AND OLDER USERS OF POWER MOBILITY DEVICES

Claudine Auger, OT, PhD[1,2], Louise Demers, OT, PhD[1,2], Isabelle Gélinas, OT, PhD[3,4], François Routhier, PEng, PhD[5], W. Ben Mortenson, OT, PhD[1,2,6] and William C. Miller, OT, PhD[7,8]

From the [1]School of Rehabilitation, University of Montreal, [2]Research Center, Institut Universitaire de Gériatrie de Montréal, [3]School of Physical and Occupational Therapy, McGill University, [4]Center for Interdisciplinary Research in Rehabilitation of Greater Montreal (CRIR), Montreal, [5]Center for Interdisciplinary Research in Rehabilitation and Social Integration (CIRRIS), Quebec, [6]Gerontology Research Centre, Simon Fraser University, [7]GF Strong Rehabilitation Research Laboratory and [8]Department of Occupational Science and Occupational Therapy, University of British Columbia, Vancouver, Canada

Objective: **To examine the measurement properties of the telephone administration of the Wheelchair Outcome Measure (WhOM).**

Subjects: **Power mobility device users aged 50–89 years.**

Methods: **Two independent cohorts were recruited: (*i*) a prospective cohort ($n = 40$) to estimate test-retest reliability and to determine the applicability of the telephone format, and (*ii*) a cross-sectional cohort to examine construct validity with 3 groups: (a) people waiting for a first power mobility device ($n = 44$); (b) initial users ($n = 35$; 1–6 months); and (c) long-term users ($n = 39$; 12–18 months).**

Results: **The tool demonstrated good test-retest reliability (intraclass correlation coefficient 0.77–1.00), took 10.9 min (standard deviation = 5.2) to administer and was practical to use over the telephone. Validity testing showed moderate correlations with the Quebec User Evaluation of Satisfaction with Technology ($r_S = 0.36–0.45$) and the Psychosocial Impact of Assistive Devices Scale ($r_S = 0.31–0.43$). WhOM scores could discriminate non-users from users (wait-list vs initial users; wait-list vs long-term users, $p < 0.001$) and power wheelchair from scooter users (total WhOM scores, $p < 0.05$).**

Conclusion: **The WhOM is a stable, valid and applicable measure for telephone administration with older power mobility device users. It is moderately linked to satisfaction with the device and to the psychosocial impact of the device, and therefore complements rather than replaces those measures.**

Key words: mobility; validation studies; disabled persons; rehabilitation; self-help devices; wheelchairs; scooters; power mobility devices.

J Rehabil Med 2010; 42: 574–581

Correspondence address: Claudine Auger, Research Centre, Institut universitaire de gériatrie de Montréal, 4565 Queen Mary Road, Montréal (Quebec), Canada, H3W 1W5. E-mail: claudine.auger@umontreal.ca

Submitted June 23, 2009; accepted January 20, 2010

INTRODUCTION

The eligibility criteria for power mobility devices (PMDs), such as power wheelchairs and scooters, are an issue of contention in many jurisdictions (1–4) and there is a growing need for sound evidence on the outcomes of these assistive technologies (ATs). Mobility-related subsidy programmes for ATs are being challenged by the ageing of the population (5, 6). Adults aged over 50 years are the most prevalent wheelchairs users (7) and it is estimated that PMD use is 3.5 times more frequent after the age of 65 years (5). However, very few studies have addressed the impact of PMDs on the lives of older adults (8) and those that have generally relied on outcome measures with very little evidence of reliability and validity for this population (9).

One of the fundamental goals of PMD provision is to improve "functioning", that is to maximize the potential use of body functions (body level), to increase the capacity to perform activities (activity level) and, ultimately, to allow an expansion of participation in valued life roles (participation level) (10). A critical review of measurement tools designed for wheelchair users indicated that most of these tools assess only mobility activities, and emphasize a normative evaluation of capacities of the users in standardized settings (9). To appreciate the full impact of PMDs at the participation level, the range of outcome domains has to be expanded beyond mobility activities. Moreover, the input of the user is crucial, since there is no universal standard for successful participation in valued activities and life roles (11, 12). Some wheelchair user-specific questionnaires have been developed to assess self-perceived wheelchair skills (13) and function (14) related to wheelchair/scooter use. Moreover, an exhaustive participation measure for individuals with various mobility limitations is available (15). However, these tools were not designed to take into account the participation aspirations of the wheelchair user, and whether or not significant activities and valued social roles were effectively enabled by the device.

© 2010 The Authors. doi: 10.2340/16501977-0557
Journal Compilation © 2010 Foundation of Rehabilitation Information. ISSN 1650-1977

The Wheelchair Outcome Measure (WhOM) (16) is an individualized, goal-oriented measure of outcome related to wheelchair intervention. Wheelchair intervention is defined in broad terms and covers either a new wheelchair prescription, wheelchair renewal, adjustments to the device/environment or additional training of the user. The WhOM is designed to identify desired outcomes at the participation level, but also some body structure and function items as defined by the International Classification on Functioning, Disability and Health (ICF) (10). The tool measures importance and satisfaction with a range of self-selected, wheelchair-related activities. The WhOM was found to be reliable and valid with a population of young adults with spinal cord injuries who used manual and power wheelchairs (17). Test-retest and inter-rater reliability intraclass correlation coefficients (ICCs) estimates were substantial during face-to-face interviews (ICCs > 0.90), and construct validity was supported by moderate associations ($0.33 < r < 0.66$) with a generic participation measure, as well as with satisfaction with assistive technologies (17).

As part of a larger study, the WhOM was adapted and translated into French in order to conduct telephone interviews with middle-aged and older adults. Telephone questionnaires are less time-consuming than face-to-face interviews (18) and provide easier access to broad geographical areas at a lower cost. Since reliability and validity are context- and population-specific attributes (19), the translation of a measurement tool and the modification of the assessment format require additional psychometric testing to verify the equivalence of the measure.

The objective of this study was to examine the properties of the French version of the WhOM in relation to reliability and applicability for telephone administration, as well as the construct validity of the WhOM for French- and English-speaking middle-aged and older PMD users. Such information would confirm if the measurement properties of the tool are adequate for telephone interviews with PMD users and extend its use for research and clinical purposes with French- or English-speaking populations as well as older adults.

METHODS

Study design

This study involved 2 independent cohorts of PMD users: (*i*) a French-speaking prospective cohort to estimate reliability, as well as the applicability of the telephone format (reliability sample), and (*ii*) a French- or English-speaking cross-sectional cohort to examine construct validity (validity sample). A test-retest approach was used to address reliability. This approach estimates the stability of the responses by repeating the questionnaire during a period where no changes are expected for the construct under study (19). Applicability refers to context- and population-specific pragmatic qualities of an assessment tool (20). During the telephone interview, applicability was assessed by coding the burden of assessment for the participant and the examiner, and the identification of floor or ceiling effects. Once reliability and content validity was established with the sample of 40 French-speaking users, then the examine validity of the telephone version was examined with a larger sample of French- and English-speaking users. In the absence of a gold standard, construct validity was examined through convergent and discriminant validity testing. Convergent validity estimated the degree of association with the satisfaction with the device and with the psychosocial impact for PMD users. Discriminant validity testing was used to determine whether the WhOM could distinguish between users of 2 device types and 3 durations of use. Differences in frequency of use and perceived environmental barriers have been noted between power wheelchair and scooter users (21) that we anticipated would be captured by the WhOM. Gitlin et al. (22) have found that AT outcomes vary according to the level of experience of the user. They consider the first 6 months as the initial use period, while the period beyond the first year delineates expert use. The study design involved 3 groups (waiting for PMD provision, initial users and long-term users) to determine whether the WhOM discriminated various durations of use.

Participants

The target population included individuals eligible for a power wheelchair funded by the Provincial Health Insurance Agency (*Régie de l'assurance maladie du Québec* (RAMQ)) or for a scooter from the Quebec Health and Social Services Ministry (*Ministère de la santé et des services sociaux du Québec* (MSSS)). The project was approved by the Institutional Review Boards of 4 rehabilitation centres. The rehabilitation centres provided lists of individuals who had received a PMD or were waiting to receive one ($n = 48$ reliability sample; $n = 213$ validity sample). Names on the list were randomized on a per centre basis and individuals were contacted following randomization. Those who were eligible were invited in participate to the study ($n = 46$ reliability sample; $n = 139$ validity sample). Recruitment continued until the planned sample size was reached, i.e. $n = 40$ for the reliability sample and $n = 116$ for the validity sample. Sample size for the reliability analysis was calculated based on requirements reported by Streiner & Norman (19) for consistency between 2 observations, taking into account a previous intra-class coefficient estimation by Garden et al. (17) of 0.93 and a standard error of 0.05. For the validity sample, it was calculated that a sample size of 35 participants per group would enable us to detect a large effect size $f > 0.40$ with a power of 0.80 at a level of significance of 0.05.

Two samples of individuals were drawn from those who met the following criteria : (*i*) eligible for PMD provision from RAMQ or MSSS, (*ii*) age 50 years and over, (*iii*) first PMD provision, (*iv*) ability to communicate by telephone in French (reliability sample), and in French or English (validity sample), (*v*) duration of use was an additional inclusion criterion for the validity sample: group 1 was waiting for the device, group 2 was using the device for 1–6 months and group 3 was using the device for 12–18 months. Exclusion criteria were inability to use the PMD due to mechanical failure or hospitalization for more than 48 h during the month preceding the interview.

The telephone interviews took place between December 2006 and April 2007 for the reliability sample, and between June and October of summers 2007 and 2008 for the validity sample. Those timeframes ensured comparable climatic conditions within each cohort, since climate is known to affect the participation of wheelchair users (21, 23).

Measurement tools

The WhOM is divided into 2 parts. In part I, the examiner asks 2 open-ended questions to allow the client to identify wheelchair-related participation goals at home and in the community. In part II, 3 close-ended questions on body functions relate to comfort, positioning and skin condition. The therapist uses a semi-structured interview guide and prompting questions to collect information. For instance, the question about wheelchair participation in the community asks: "*Some people use their [wheelchairs or scoters] because they want to participate in activities outside of their [home or facility], such as dog walking, going for coffee, to work or to the park. What activities outside of your [home or facility] or in your community would you use your [wheelchair or scoter] to perform?"* When participants mentioned more than 5 goals (at home or in the community), they were asked to narrow down the list to the top 5. Part I rates the importance and satisfaction with participation for each goal on an 11-point scale (0 = not important at all,

to 10 = extremely important; 0 = not satisfied at all, to 10 = extremely satisfied). Two mean scores and 2 total scores are computed: *(*i) mean importance × satisfaction (MeanIMP × SAT): ranging from 0 to 100; (*ii*) mean satisfaction (MeanSAT): ranging from 0 to 10; (*iii*) total importance × satisfaction (TotIMP × SAT): ranging from 0 to 1000; and (*iv*) total satisfaction (TotSAT): ranging from 0 to 100. Part II questions on body function are rated with an 11-point scale (0 = low score, to 10 = high score). A visual 10-point scale helps participants to identify numeric ratings. Higher scores indicate better outcomes. The French version was developed based on a systematic methodology to ensure that the original and adapted versions were equivalent (24, 25). The translation/back-translation process involved professional translators ($n = 4$), bilingual occupational therapists ($n = 4$), and an expert committee composed of 6 members (2 professional translators, 1 bilingual occupational therapist, 1 PMD user, and 2 researchers (CA and FR)). The preliminary version was pretested with 5 bilingual PMD users and 5 bilingual rehabilitation professionals to ensure conceptual equivalence. A bilingual telephone script (French and English) was developed for the WhOM. The script specifies some cueing instructions that can be utilized to probe the answers of the participant over the telephone. For instance, in the absence of the visual cues of the standard 10-point scale, the examiner could use descriptors to narrow down the options of the 0–10 numerical scale, and thereby facilitate the selection of a numerical value by the participant. The manual describes the telephone script, testing procedures and scoring guidelines (freely available from: bcmiller@telus.net).

Satisfaction with the PMD was measured with the Quebec User Evaluation Satisfaction with Technology questionnaire (QUEST 2.0) (26), a 12-item questionnaire that generates 3 scores: an average satisfaction with "technology" score based on 8 items, an average satisfaction with "services" score derived from 4 items, and a "total" score averaged across the 12 items. Each item is scored using a 5-point satisfaction scale, with a score of 1 denoting "not satisfied at all" and 5 indicating that the person is "very satisfied".

The 10-item version of the Psychosocial Impact of Assistive Devices Scale (PIADS-10) (27) was used for self-rating perceptions of how assistive devices affect quality of life for aspects such as competence (e.g. feelings of independence), adaptability (e.g. willingness to try new things), and self-esteem (e.g. feelings of emotional wellbeing and happiness). Scores on each item can range from –3 (maximum negative impact) to +3 (maximum positive impact). The short version of the PIADS generates a total score, ranging between –3 and +3, that is averaged across 10 items.

The QUEST and PIADS-10 were chosen based on conceptual appropriateness, potential for telephone use (28) in both French and English and the quality of their measurement properties (26, 28).

Demographic and clinical background information including age, gender, diagnosis, type of PMD, living arrangement, geographical area, accessibility of residence and access to adapted transport were extracted from charts that contained the participant's provincial wheelchair application forms.

Procedure for test-retest reliability and applicability

Telephone interviews were conducted with 40 participants by a rehabilitation clinician (examiner). The delay between the test (T1) and retest (T2) telephone interviews was 7–14 days as suggested by Streiner & Norman (19) and repeated by the same examiner. Both participants and examiners did not have access to the list of participation objectives nominated at T1 when the questionnaire was repeated. The initial interview was taped with a digital recorder and the duration of the assessment was calculated from the mp3 file. The behaviour coding technique of Fowler & Cannell (29) was used systematically to register behaviours that reflect the applicability of the questionnaire, such as the burden for the participant and the examiner (e.g. participants asking for clarifications, examiner reformulating a question, missing data, negative comments). To look at the stability of goals over time, each participation objective was coded by the examiner with detailed 4-level ICF codes. One participant could report only one objective from a single 4-level ICF code.

Procedure for convergent and discriminant validity

The WhOM, QUEST and PIADS-10 were administered randomly to avoid order effects. All questionnaires were completed over 1 or 2 telephone sessions over a period of 7 days, except for 2 participants who needed 3 sessions over the same period. The number of sessions was self-selected by the participants depending on their availability and fatigue.

Analyses

Descriptive statistics were computed and histograms were visually inspected for all variables. The proportion of scores at the top and bottom of the scales were calculated. A floor or ceiling effect was defined as a clustering of > 20% of the responses at the minimal or maximal level of the scale based on recommendations by Andresen et al. (30). Assumptions for homoscedasticity were also verified with normality test results (Kolmogorov-Smirnov Z score; $p < 0.05$).

To evaluate the telephone applicability of the WhOM, the following behaviours were coded with the recorded interviews on a present/absent scale: respondent burden (asks for clarifications, expresses negative comments verbally or non-verbally, interrupts the examiner, does not know the answer) and examiner burden (reformulates the question, simplifies the rating scale, uses probing, adds definitions). The percentage of occurrence (P_{oc}) of each behaviour was computed across all participants and regrouped at the item level. Based on Fowler & Cannell's criteria (29) for behaviour coding, an item was considered problematic when the P_{oc} was above 15% ($n > 6$).

The stability of the participation objectives between T1 and T2 was verified descriptively and statistically. The proportion who mentioned the same participation goals, based on 4-level ICF codes, and the proportions who changed goals were computed. Test-retest reliability was calculated with ICCs from a 2-way mixed effects model ($ICC_{2,1}$) with scores as a between-subjects random effect and assessment session as a within-subjects fixed effect (31). An ICC > 0.75 was considered high, between 0.75 and 0.40 was considered moderate, and < 0.40 was considered low (30). A reflect and square root transformation was applied to the MeanSAT score to correct a skewed distribution towards the higher end of the scale, as recommended by Tabachnik & Fidel (32).

Convergent validity was assessed using Spearman's rho (r_S). Statistically significant correlations were interpreted as large when the coefficient was at least 0.50, moderate between 0.30 and 0.49, and small between 0.10 and 0.29 (33). Based on a previous validation study with the WhOM (17), it was hypothesized that the WhOM scores would be moderately and positively correlated with satisfaction with the AT (QUEST). Higher correlations were expected with psychosocial impacts of the PMD (PIADS-10) as it is centred on the personal experience of the user rather than on the technology. For discriminant validity testing, it was hypothesized that the importance of wheelchair participation goals and satisfaction would be higher for power wheelchair users compared with scooter users, since power wheelchairs are usually provided to those with more severe motor impairments (based on eligibility criteria for RAMQ or MSSS), and thus perform more activities with a PMD. Moreover, lower WhOM scores were expected for the group that was waiting for the device. Due to skewed distributions for some of the scores, non-parametric analyses were used to test the 2 hypothesis including the Kruskal-Wallis test (3 durations) and Mann-Whitney *U* test (2 device types) with a significance level of $p < 0.05$. To contrast the 3 durations, *post-hoc* tests were conducted with the Mann-Whitney *U* test, using a Bonferroni correction ($\alpha = 0.017$).

Statistical analyses were performed with the Statistical Package for the Social Sciences (SPSS) version 16.0.

RESULTS

Descriptive results

The characteristics of 2 independent reliability and validity samples are presented in Table I. In both samples, mean age was

Table I. *Characteristics of the 2 samples*

Variable	Reliability sample ($n=40$) *n* (%)	Validity sample ($n=116$) *n* (%)
Age, years		
50–64	25 (62.5)	59 (50.8)
65–74	8 (20.0)	37 (31.9)
≥ 75	7 (17.5)	20 (17.3)
Gender		
Male	17 (42.5)	47 (40.5)
Female	23 (57.5)	69 (59.5)
Main diagnosis		
Neurological	21 (52.5)	58 (50.0)
Musculoskeletal	10 (25.0)	30 (25.9)
Medically complex	9 (22.5)	28 (24.1)
Type of PMD		
Scooter	26 (65.0)	60 (51.7)
Power wheelchair	14 (35.0)	56 (48.3)
Living arrangement		
Alone	15 (43.0)	32 (27.6)
Not alone	25 (57.0)	84 (72.4)
Geographical area		
Urban	28 (70.0)	100 (86.2)
Rural	12 (30.0)	16 (13.8)
Accessibility of residence		
Fully accessible	30 (75.0)	60 (51.7)
Partial or not accessible	10 (25.0)	56 (48.3)
Adapted transport		
Privately owned vehicle	2 (5.0)	4 (3.4)
Public transportation	20 (50.0)	77 (66.4)
No adapted transport	18 (45.0)	35 (30.2)
Duration of PMD use		
Group 1: waiting	–	42 (36.2)
Group 2: 1–6 months	–	35 (30.2)
Group 3: 12–18 months	–	39 (33.6)

PMD: power mobility devices.

65 years (SD = 10) and a majority were women. The reliability sample was comprised of 40 persons who had been using their PMD for 2–15 months (mean = 7.7 months, SD = 3.1). Most of these participants lived in the community (87.5%; $n=35$), while some lived in long-term care facilities (10%; $n=4$). The validity sample comprised 116 persons either waiting for ($n=42$) or using ($n=74$) a PMD. This sample was similar to the reliability sample, except that part of the sample completed the questionnaires in English (4.3%, $n=5$), the representation of rural areas was much lower (13.8%; $n=16$) and the proportion of users with access to adapted transportation was higher (69.8%; $n=81$).

Test-retest reliability

The participation objectives mentioned at test and retest by the 40 participants are listed in Table II in order of descending frequency for each ICF chapter. The proportion of community participation goals was 4 times higher than participation goals at home, with a marked predominance of recreation and leisure activities. Community goals predominated, since a large proportion of participants had no participation goal at home ($n=29$; 73%) whereas only 1 did not mention any goal in the community ($n=1$; 2.5%). The proportion of participants who spontaneously mentioned the same participation objectives on T1 and T2 (based on the 4-level ICF code) or changed one of them was 90% ($n=36$) at home and 77.5% ($n=31$) in the community. The rest of the participants changed 2 objectives (2.5% ($n=1$) at home; 15.0% ($n=6$ community)) or 3 objectives (2.5% ($n=1$) at home; 7.5% ($n=3$) community) between T1 and T2. Descriptive data (mean and ranges) at T1 and T2, as well as reliability estimates are presented at Table III. Com-

Table II. *Participation objectives at home and in the community coded with the ICF (reliability sample)*

ICF code	ICF chapter	Number of participation objectives: At home	Number of participation objectives: Community	Examples of participation objectives
9200	Recreation and leisure	12	56	Play bowling once a week all year round
6200	Acquiring goods and services	0	29	Shop once a week with partner all year round
4600	Moving around in different locations	6	13	Take daily rides to the park depending on climatic conditions
7500	Informal social relationships	2	9	Have a coffee with people at the shopping mall once a week
7600	Family relationships	0	8	Visit sister/brother/children once or twice a month depending on climatic conditions
6500	Caring for household objects	3	3	Shop once a week with partner all year round
9300	Religion and spirituality	1	3	Attend church activities once a week during spring, summer and fall
5700	Looking after one's health	1	2	Go to doctor's appointment when needed
8600	Economic transactions	0	3	Go to the bank without help once a week all year round
3600	Using communication devices	2	0	Get to the computer room to use internet twice per week, all year round
6300	Preparing meals	2	0	Cook all meals while using the PMD
8500	Employment (remunerative or not)	0	2	Use the PMD at work twice a week all year round
9100	Community life	0	2	Participate in the local fund-raising committee once a month
5500	Eating	1	0	Have lunch at the cafeteria every day all year round (long-term care resident)

ICF: International Classification of Functioning, Disability and Health. Note: Listed from most frequent to least frequent ICF codes. The figures correspond to the total number of participation goals mentioned by the 40 participants. A 2-level ICF code aggregates up to 7 4-level ICF codes, and therefore one participant can report more than one objective within a single category; PMD: power mobility devices.

Table III. *Test/retest reliability estimates of the Wheelchair Outcome Measure (WhOM) scores (*n = *40)*

WhOM scores (theoretical range)	Test mean score (range)	Retest mean score (range)	ICC	$ICC_{0.95}$
Part I: Mean scores				
MeanIMP × SAT (0–100)	79.8 (22.4–100)	77.5 (28–100)	0.89	0.78–0.94
Mean SAT (0–10)	8.9 (4.8–10)	8.6 (4–10)	0.79[a]	0.59–0.89
Part I: Total scores				
TotIMP × SAT (0–1000)	312 (64–900)	293 (28–900)	0.92	0.84–0.96
TotSAT (0–100)	35 (10–90)	33 (4–90)	0.91	0.83–0.95
Part II: Body functions				
Comfort (0–10)	8.7 (4–10)	8.8 (4–10)	0.85	0.71–0.92
Position (0–10)	9.2 (6–10)	9.1 (4–10)	0.77	0.57–0.88
Skin condition (0–10)	10 (0)	10 (0)	1.00	1.00–1.00

[a] ICC calculated on transformed score (reflect and square root transformation).
ICC: intra-class coefficient; $ICC_{0.95}$: ICC 95% confidence interval.

puted test-retest reliability estimates were high for all WhOM scores (ICC = 0.77–1.00).

Applicability

The WhOM interview was administered in 3–25 min (mean = 10.9, SD = 5.2). The behaviour coding revealed that respondents could answer all items and that the questionnaire was well accepted, since negative comments ($n = 1$) and interruptions ($n = 2$) were minimal. The examiner burden coding did not reveal any administration difficulties. The total scores of the WhOM at T1 and T2 were well distributed, but the mean scores were negatively skewed (Table IV). A ceiling effect was identified for the MeanSAT and MeanIMP × SAT, since more than 20% of the responses were clustered at the top of the scale. Statistically significant Kolmogorov-Smirnov tests confirmed the distribution biases ($p < 0.05$). Regarding the distribution of body function scores, the observed range was restricted to between 8 and 20 on a theoretical scale of 0–30, since none of the participants reported any skin breakdown.

Convergent and discriminant validity

Convergent validity estimates between the QUEST, PIADS-10 and WhOM are shown in Table V. At a descriptive level, the 74 PMD users were generally "Quite satisfied (4/5)" (66.2%; $n = 49$) with their PMD as measured with the QUEST (mean = 4.4/5; SD = 0.6; range = 2.5–5.0). Positive psychosocial impacts were reported on average on the PIADS-10 (mean = 1.9/3; SD = 0.8; range = –0.2 to 3.0). The convergent validity analyses estimated moderate coefficients ranging from 0.36 to 0.46 between all QUEST scores and mean WhOM scores (MeanIMP × SAT and MeanSAT). Correlations ranging from 0.31 to 0.43 were moderate with the PIADS-10 for all WhOM scores, except for TotSAT, which was not significant. The strength of associations were slightly larger for the PIADS-10 when both importance and satisfaction were included in the WhOM scoring (MeanIMP × SAT and TotIMP × SAT) compared with corresponding satisfaction scores (MeanSAT and TotSAT). Total satisfaction (TotSAT) was not correlated significantly with any other variable, and TotIMP*SAT had a small correlation with QUEST services.

Discriminant validity compared participation outcomes for different durations of use and device types. The Kruskal-Wallis test estimated statistically different WhOM scores across duration of use ($p < 0.001$) for all the WhOM scores (Table VI). Significant contrasts were found between wait-list and both groups of users (Mann-Whitney *U* test, wait-list vs initial users $p < 0.001$; wait-list vs long-term users $p < 0.001$) for all WhOM scores. There was no difference between initial and long-term users (Mann-Whitney *U* test; $p = 0.58$–0.83). When the results were analysed separately for the home and community environments, duration was not statistically significant at home (Kruskal-Wallis, $p = 0.23$–0.66), but remained significant in the community (Kruskal-Wallis, $p < 0.001$) (data not shown).

Discriminant validity based on device type was significant for the WhOM total scores (TotIMP × SAT and TotSAT; Mann-Whitney *U* test, $p < 0.05$), with higher scores for power wheelchair users, as shown at Table VII. When results were analysed

Table IV. *Distribution of the Wheelchair Outcome Measure (WhOM) subscores at test and retest (*n = *40)*

	Test (T1)			Retest (T2)		
	Actual range	bottom of scale % (*n*)	top of scale % (*n*)	Actual range	bottom of scale % (*n*)	top of scale % (*n*)
Part I: Mean scores						
MeanIMP × SAT (0–100)	22–100**	0 (0)	22.5 (9)	28–100**	0 (0)	15 (6)
MeanSAT (0–10)	4.8–10*	0 (0)	27.5 (11)	4–10	0 (0)	25 (10)
Part I: Total scores						
TotIMP × SAT (0–10009	64–900	0 (0)	0 (0)	28–900	0 (0)	0 (0)
TotSAT (0–100)	10–90	0 (0)	0 (0)	4–90	0 (0)	0 (0)
Part II: Body functions (0–30)	11–20*	0 (0)	0 (0)	8–20**	0 (0)	0 (0)

*$p < 0.05$; **$p < 0.001$.
Statistically significant Kolmogorov-Smirnov test.

Table V. *Convergent validity for initial and long-term users (*n = *74)*

	WhOM Part I: Participation			
	Mean IMP × SAT	Mean SAT	Tot IMP × SAT	Tot SAT
QUEST total score	0.37**	0.45**	0.10	0.03
QUEST technology	0.41**	0.46**	0.10	0.01
QUEST services	0.36**	0.42**	0.24*	0.19
PIADS-10	0.43**	0.33**	0.31**	0.21

Spearman's rho; *$p < 0.05$; **$p < 0.001$.
WhOM: Wheelchair Outcome Measure; QUEST: Quebec User Evaluation Satisfaction with Technology questionnaire.

separately for the home and community environments, both total (TotIMP × SAT, TotSAT) and mean satisfaction scores (MeanIMP × SAT and MeanSAT) discriminated between power wheelchair and scooter users at home ($p < 0.001$). In the community, none of the scores discriminated between device types (Kruskal-Wallis, $p = 0.10$–0.19).

DISCUSSION

This study examined the measurement properties of the telephone version of the WhOM with middle-aged and older PMD users with respect to reliability, applicability and validity. One of the strengths of the present study was the random selection of PMD users through institutions that held provincial or regional mandates for the provision of subsidized power wheelchairs and scooters. This design ensured that participants representing various levels of participation and satisfaction were included. Another strength was the inclusion of users with different durations of use. Research on the impacts of PMDs in older adults is scarce and typically does not control for duration of use and device type (8).

All the 2 week test-retest reliability estimates were above the threshold of 0.75 suggested by Andresen (30). Moreover, a qualitative verification of the stability of the nominated objectives was performed and revealed that 90% spontaneously repeated the same participation objectives at home or modified one between the 2 assessments. This proportion was slightly lower (77.5%), meaning less stable, regarding the participation objectives in the community. In both environments, it must be noted that changes in the nominated objectives generally remained in the same ICF chapters at test and retest (e.g. different leisure activities within the recreation chapter). These results support the WhOM administration guidelines that suggest reminding the user of his or her initial participation objectives when performing the follow-up assessment. However, the assessment should not be deferred if the initial list of participation objectives is unavailable, since the WhOM score remains stable.

We found that modifying the assessment format from face-to-face to telephone did not alter its applicability for middle-aged and older PMD users. The WhOM was well accepted and simple to administer in less than 15 min. It was well understood by middle-aged and older PMD users. The delimitation of home vs community had to be clarified occasionally when the participation goals took place in the backyard or on the terrace of the person's residence. The telephone script comprised descriptors to narrow down the options of the numerical scale, which proved to be useful for participants who had difficulty selecting a numerical value. The appraisal of the properties of a questionnaire must also consider the distribution of scores to rule out floor or ceiling effects. A clustering of responses at the maximal level of the scale was found for up to 27.5% of respondents on the MeanSAT score. Total scores might be a better alternative than mean scores to track change over time, especially if the same number of initial goals are rerated, since total scores displayed normal distributions and comparable reliability results to the original version tested by Garden et al. (ICC = 0.93; $ICC_{0.95}$ = 0.88–0.96) (17).

This study was the first to examine the properties of the total scores of the WhOM. The discriminant validity of the total scores (TotIMP × SAT and TotSAT) across duration of use was confirmed partly, as these scores could distinguish non-users from users, and they highlighted significant differences between device types. Interestingly, however, the total satisfaction score (TotSAT) did not meet the convergent validity hypothesis with the QUEST and PIADS-10. Moreover, TotIMP × SAT was only significantly correlated with the

Table VI. *Discriminant validity results for duration of use (*n = *116)*

	Group mean				
WhOM scores	W	I	LT	K-W[1]	Statistically significant pairwise comparisons[2]
MeanIMP × SAT (/100)	29.6	72.9	77.5	0.00*	W-I*; W-LT*
MeanSAT (/10)	3.4	8.3	8.9	0.00*	W-I*; W-LT*
TotIMP × SAT (/1000)	133.6	365.3	344.4	0.00*	W-I*; W-LT*
TotSAT (/100)	15.7	41.0	38.7	0.00*	W-I*; W-LT*

[1]Kruskal-Wallis test; [2]Satistically significant pairwise comparisons with Mann-Whitney *U* test.
*$p < 0.001$.
WhOM: Wheelchair Outcome Measure; W: Wait-list group; I: Initial user group (1–6 months); LT: Long-term user group (12–18 months).

Table VII. *Discriminant validity results for device type (*n = *74)*

	Group mean		
WhOM scores	PWC	Scooter	M-W[1]
Home/Community			
MeanIMP × SAT (/100)	71.5	78.5	0.15
MeanSAT (/10)	8.1	8.8	0.12
TotIMP × SAT (/1000)	426.0	296.5	0.02*
TotSAT (/100)	48.2	33.0	0.00**
Home			
MeanIMP × SAT (/100)	57.9	15.7	0.00***
MeanSAT (/10)	6.6	1.9	0.00***
TotIMP × SAT (/1000)	198.8	29.2	0.00***
TotSAT (/100)	22.5	3.3	0.00***
Community			
MeanIMP × SAT (/100)	69.9	79.5	0.10
MeanSAT (/10)	7.9	8.8	0.16
TotIMP × SAT (/1000)	227.4	267.3	0.13
TotSAT (/100)	25.7	29.8	0.19

[1]Mann-Whitney *U* test.
*$p < 0.05$; **$p < 0.01$; ***$p < 0.001$.
PWC: power wheelchair; WhOM: Wheelchair Outcome Measure.

QUEST services and PIADS-10 scores. It is possible that total scores, through a reflection of the number of participation objectives nominated by PMD users, are actually more closely related to how much the device is used than to the satisfaction with the device or its psychosocial impact. Future studies should verify the utility of the WhOM total scores to track how much the device is used when transitioning from part-time use to full-time PMD use, and examine the relationships of the different types of WhOM summary scores with the severity of mobility limitations.

As expected, our results support a moderate positive link with satisfaction with the device, as measured by the QUEST. However, the coefficients of 0.42–0.46 obtained with the present sample are slightly lower than estimated by the validation study of Garden et al. (17). That study found an association of 0.58 between mean WhOM and QUEST in younger long-term users of power and manual wheelchairs with spinal cord injury. A question remains whether the discrepancies between the samples can be explained by differences in eligibility criteria, diagnosis, age, duration of use, and types of devices.

A stronger positive correlation was expected with the PIADS-10 than with the QUEST, since the former is centred on the impacts on the user rather than on the characteristics of the device. The associations in our validity sample were slightly stronger, though moderate, with the PIADS-10 when both importance and satisfaction were considered. That would indicate a better psychological reaction to the assistive devices when users value the activities they are performing with it. The correlations in the present study were stronger than estimated by Buning et al. (34), who identified a small correlation of 0.21 between the PIADS and a generic occupational performance measure for PMD users.

The moderate associations between the WhOM, and both the QUEST and PIADS-10, suggest that each brings distinct contributions to the assessment of the outcomes of the PMD intervention. Higher levels of satisfaction with the device or positive psychosocial impacts do not necessarily imply that the participation objectives of the users are fully met and vice versa. This observation is in accordance with theoretical models of the interaction between the user, the device and the activity in a specific context (35, 36). For example, the user might indicate that the device is technically adequate, but still experience dissatisfaction when performing specific activities or life roles with the device.

Differentiating outcomes according to device types has been identified as a fundamental gap in the field of wheeled mobility (37). To our knowledge, very few tools can distinguish participation levels for various types of mobility-related assistive technology, besides the Facilitators And Barriers Survey of environmental influences on participation among people with lower limb Mobility impairments (FABS-M) (21). The present study showed that the WhOM captures participation differences between power wheelchair users and scooter users on their total satisfaction with participation, possibly because the number of objectives identified by power wheelchair users was higher at home. When the environments were analysed separately, all WhOM scores differentiated power wheelchair users from scooter users in the home environment, but none in the community. The environment was a key aspect in differentiating outcomes. Participation in the community was most valued and appeared to be a key dimension for users of both device types, while participation at home was an aspect that distinguished power wheelchair from scooter users. This result is supported with previous reports on powered mobility indicating that domestic roles, such as shopping (23, 38), and community life, such as social activities (23, 38, 39), are highly valued. The validity results have implications for the use of the WhOM. Mean scores are recommended for group comparisons because the scores are not affected by the number of goals, and separate analyses for the home and community environments are indicated to discriminate device type. If the aim is to quantify PMD participation at the individual level, total scores are indicated to reflect changes in the number of participation goals across time.

There are limitations to this study. First, future studies should address other psychometric properties, such as the responsiveness of the different WhOM scores and the inter-rater reliability of the telephone format. Secondly, the representation of English-speaking PMD users was very low in the validity sample, therefore culture-specific evidence for these users would require larger samples. Thirdly, a cross-sectional design does not allow one to state with confidence that the differences observed between the waiting group, and the initial and long-term users truly reflect differences due to duration of use. As we attempted to control for climatic conditions by selecting a time-window limiting the overlay of summer and winter seasons, the amount of experience post-delivery might differ for users who receive their device at the beginning of winter or in the middle of the summer. Future studies tracking PMD use with repeated measures across time could provide firmer evidence. Fourthly, subsidy guidelines were based on a medical prescription with specific eligibility criteria and the mandatory involvement of an occupational or physical therapist in the assessment and follow-up process. The generalization of the present results applies to populations with neurological, musculoskeletal or medically complex conditions requiring medically prescribed devices.

In conclusion, our results show that the WhOM is a stable, valid and applicable measure that is practical to use with a population of PMD users of age 50 years and over. The results of this study are important because there is a need for wheeled device-specific tools that are applicable to older adults. The tool could discriminate non-users from users and device types (power wheelchair vs scooter). Moreover, this client-centred measure is moderately linked to the perception of the client about his or her satisfaction with the device and its psychosocial impact, and therefore complements rather than replaces those measures.

ACKNOWLEDGEMENTS

Special thanks are extended to Malgorzata Depa, Sophie Turgeon Londei (research assistants), and Francine Giroux (statistician). The authors gratefully acknowledge the study participants for their generosity and the

personnel from the Quebec City Rehabilitation Institute, Gingras-Lindsay Rehabilitation Institute of Montreal, Lucie-Bruneau Rehabilitation Center and Constance-Lethbridge Rehabilitation Center.

This project was supported by Grant MIA85858 from the Institute of Aging of the Canadian Institutes of Health Research (CIHR). Claudine Auger's work was supported by a doctoral award for health professionals from the *Fonds de recherche en santé du Québec* and by a CIHR Institute of Aging Fellowship.

REFERENCES

1. Roy CW. Common criteria for providing powered wheelchairs should be agreed by wheelchair service centres. BMJ 1997; 315: 605–606.
2. Jörg F, Boeije HR, Schrijvers AJP. Professionals assessing clients' needs and eligibility for electric scooters in the Netherlands: both gatekeepers and client's advocates. Br J Soc Work 2005; 35: 823–842.
3. Agence d'évaluation des technologies et des modes d'intervention en santé (AETMIS). Triporteurs et quadriporteurs: solutions de rechange aux fauteuils roulants à propulsion motorisée? Montréal: AETMIS; 2007. Report No.: AETMIS 07–05.
4. Dicianno BE, Tovey E. Power mobility device provision: understanding Medicare guidelines and advocating for clients. Arch Phys Med Rehabil 2007; 88: 807–816.
5. Kaye H, Kang T, LaPlante M. Mobility device use in the United States. Disability Statistics Report, (14). Washington, DC: US Department of Education, National Institute on Disability and Rehabilitation Research; 2000.
6. Vignier N, Ravaud JF, Winance M, Lepoutre FX, Ville I. Demographics of wheelchair users in France: results of national community-based handicaps-incapacites-dependance surveys. J Rehabil Med 2008; 40: 231–239.
7. Clarke P, Colantonio A. Wheelchair use among community-dwelling older adults: prevalence and risk factors in a national sample. Can J Aging 2005; 24: 191–198.
8. Auger C, Demers L, Gelinas I, Jutai J, Fuhrer M, DeRuyter F. Powered mobility for middle-aged and older adults: systematic review of outcomes and appraisal of published evidence. Am J Phys Med Rehabil 2008; 87: 666–680.
9. Mortenson WB, Miller WC, Auger C. Issues for the selection of wheelchair-specific activity and participation outcome measures: a review. Arch Phys Med Rehabil 2008; 89: 1177–1186.
10. World Health Organization. International Classification of Functioning, Disability and Health: short version. Geneva: World Health Organization; 2001.
11. Rochette A, Korner-Bitensky N, Levasseur M. 'Optimal' participation: a reflective look. Disabil Rehabil 2006; 28: 1231–1235.
12. Noreau L, Fougeyrollas P, Post M, Asano M. Participation after spinal cord injury: the evolution of conceptualization and measurement. J Neurol Phys Ther 2005; 29: 147–156.
13. Mountain AD, Kirby RL, Smith C. The wheelchair skills test, version 2.4: Validity of an algorithm-based questionnaire version. Arch Phys Med Rehabil 2004; 85: 416–423.
14. Mills T, Holm MB, Trefler E, Schmeler M, Fitzgerald S, Boninger M. Development and consumer validation of the Functional Evaluation in a Wheelchair (FEW) instrument. Disabil Rehabil 2002; 24: 38–46.
15. Gray DB, Hollingsworth HH, Stark SL, Morgan KA. Participation survey/mobility: psychometric properties of a measure of participation for people with mobility impairments and limitations. Arch Phys Med Rehabil 2006; 87: 189–197.
16. Mortenson B, Miller W, Miller Polgar J. Measuring wheelchair intervention outcomes: development of the Wheelchair Outcome Measure. Dis Rehabil Assist Technol 2007; 2: 265–285.
17. Garden J, Miller W, Mortenson B. A client-specific outcome measure of wheelchair intervention. Can J Occup Ther suppl 2006; 73: 59.
18. Holbrook A, Green M, Krosnick J. Telephone versus face-to-face interviewing of national probability samples with long questionnaires: comparisons of respondent satisficing and social desirability response bias. Public Opin Q 2003; 67: 79–125.
19. Streiner DL, Norman GR. Health measurement scales: a practical guide to their development and use. 3rd ed. New York: Oxford University Press; 2003.
20. Auger C, Demers L, Swaine B. Making sense of pragmatic criteria for the selection of geriatric rehabilitation measurement tools. Arch Gerontol Geriatr 2006; 43: 65–83.
21. Gray DB, Hollingsworth HH, Stark S, Morgan KA. A subjective measure of environmental facilitators and barriers to participation for people with mobility limitations. Disabil Rehabil 2008; 30: 434–457.
22. Gitlin LN. From hospital to home: individual variations in experience with assistive devices among older adults. In: Gray D, Quatrano L, Lieberman M, editors. Designing and using assistive technology: the human perspective. Baltimore: Paul H. Brookes; 1998, p. 117–136.
23. Brandt A, Iwarsson S, Stahle A. Older people's use of powered wheelchairs for activity and participation. J Rehabil Med 2004; 36: 70–77.
24. Beaton DE, Bombardier C, Guillemin F, Ferraz MB. Guidelines for the process of cross-cultural adaptation of self-report measures. Spine 2000; 25: 3186–3191.
25. Vallerand R. Vers une méthodologie de validation trans-culturelle de questionnaires psychologiques: implications pour la recherche en langue française. Can Psychol 1989; 30: 662–680.
26. Demers L, Monette M, Lapierre Y, Arnold D, Wolfson C. Reliability, validity, and applicability of the Quebec User Evaluation of Satisfaction with assistive Technology (QUEST 2.0) for adults with multiple sclerosis. Disabil Rehabil 2002; 24: 21–30.
27. Jutai J, Day H, Coulson S, Demers L, Fuhrer M, Lenker JA, et al. Developing a short form of the Psychosocial Impact of Assistive Devices Scale (PIADS). Rehabilitation Engineering and assistive technology Society of North America RESNA 2007 Annual conference. Phoenix, Arizona; 2007.
28. Day H, Jutai J, Campbell KA. Development of a scale to measure the psychosocial impact of assistive devices: lessons learned and the road ahead. Disabil Rehabil 2002; 24: 31–37.
29. Fowler F, Cannell C. Using behavioral coding to identify cognitive problems with survey questions. In: Schwarz N, Sudman S, editors. Answering questions: methodology for determining cognitive and communicative processes in survey research. San Francisco: Josey-Bass; 1996, p. 15–36.
30. Andresen EM. Criteria for assessing the tools of disability outcomes research. Arch Phys Med Rehabil 2000; 81: S15–S20.
31. Shrout P, Fleiss J. Intraclass correlations: uses in assessing rater reliability. Psychol Bull 1979; 86: 420–428.
32. Tabachnik B, Fidel L. Using multivariate statistics. 5th edn. Boston: Pearson Education; 2007.
33. Cohen J. Statistical power analysis for the behavioral sciences. 2nd edn. Hillsdale, NJ: Lawrence Erlbaum Associates; 1988.
34. Buning ME, Angelo JA, Schmeler MR. Occupational performance and the transition to powered mobility: a pilot study. Am J Occup Ther 2001; 55: 339–344.
35. Cook A, Polgar J. Cook and Hussey's assistive technologies: principles and practice. 3rd edn. St Louis: Mosby; 2008.
36. Scherer MJ, editor. The Matching person & Technology. (MPT) ModelManuel. Webster, NY: The Institute for Matching Person & Technology Rochester, NY; 1991.
37. Harris F. Conceptual issues in the measurement of participation among wheeled mobility device users. Dis Rehabil Assist Technol 2007; 2: 137–148.
38. Hoenig H, Pieper C, Branch LG, Cohen HJ. Effect of motorized scooters on physical performance and mobility: a randomized clinical trial. Arch Phys Med Rehabil 2007; 88: 279–286.
39. Pettersson I, Tornquist K, Ahlstrom G. The effect of an outdoor powered wheelchair on activity and participation in users with stroke. Dis Rehabil Assist Technol 2006; 1: 235–243.