Qual Life Res (2015) 24:1587–1596
DOI 10.1007/s11136-014-0908-0

The Patient Activation Measure: a validation study in a neurological population

Tanya L. Packer · George Kephart · Setareh Ghahari · Åsa Audulv · Joan Versnel · Grace Warner

Accepted: 23 December 2014 / Published online: 6 January 2015

Abstract
Purpose To assess the validity of the Patient Activation Measure (PAM13) of patient activation in persons with neurological conditions.
Methods "The Everyday Experience of Living with and Managing a Neurological Condition" (The LINC study) surveyed 948 adults with neurological conditions residing in Canada in 2011 and 2012. Using data for 722 respondents who met coding requirements for the PAM-13, we examined the properties of the measure using principle components analysis, inter-item correlations and Cronbach's alpha to assess unidimensionality and internal consistency. Rasch modeling was used to assess item performance and scaling. Construct validity was assessed by calculating associations between the PAM and known correlates.
Results PAM-13 provides a suitably reliable and valid instrument for research in patients with neurological conditions, but scaling problems may yield measurement error and biases for those with low levels of activation. This is of particular importance when used in clinical settings or for individual client care. Our study also suggests that measurement of activation may benefit from tailoring items and scaling to specific diagnostic groups such as people with neurological conditions, thus allowing the PAM-13 to recognize unique attributes and management challenges in those conditions.
Conclusions The PAM-13 is an internally reliable and valid tool for research purposes. The use of categorical activation "level" in clinical settings should be done with caution.

Keywords Chronic disease management · Measurement · Validation · Neurological conditions · Patient activation · PAM-13

T. L. Packer (✉) · Å. Audulv · J. Versnel · G. Warner
School of Occupational Therapy, Dalhousie University, Forrest Building, Rm 161, 5869 University Av.,
P.O. Box 15000, Halifax, NS B3H 4R2, Canada
e-mail: Tanya.packer@dal.ca

Å. Audulv
e-mail: Asa.audulv@miun.se

J. Versnel
e-mail: jversnel@dal.ca

G. Warner
e-mail: Grace.warner@dal.ca

G. Kephart
Community Health and Epidemiology, Dalhousie University, Halifax, NS, Canada
e-mail: George.kephart@dal.ca

S. Ghahari
School of Rehabilitation Therapy, Queen's University, Kingston, ON, Canada
e-mail: Setareh.ghahari@queensu.ca

S. Ghahari
Department of Occupational Therapy, University of Social Welfare and Rehabilitation Sciences, Tehran, Iran

Å. Audulv
Department of Nursing, Mid Sweden University, Sundsvall, Sweden

Introduction

Patient activation is defined as "understanding one's role in the care process and having the knowledge, skills and confidence to manage one's health and health care" [1]. It is recognized as a critical component of quality chronic condition management [2–4] and is associated with positive health outcomes [5] and health utilization patterns. For example, higher activation levels are associated with better health [6–9], preventive and healthy lifestyle behaviors [7, 10], active engagement in healthcare decision making [11–13] and more efficient health utilization [1, 14]. Education, particularly university completion, and higher income consistently demonstrate positive associations with higher activation levels; however, age, gender and sex are inconsistently associated with activation [6, 9, 15].

The Patient Activation Measure (PAM), developed by Hibbard et al. [16], is the most commonly used tool for assessing activation. The original 22-item PAM was rigorously developed using a four-stage process to identify, operationalize and pilot test the instrument; it was reduced to 13 in 2005 (PAM-13) [9]. The PAM is an interval level, unidimensional, Guttman-like scale with items sequenced by "difficulty" of activation. For both PAM and PAM-13, Rasch modeling was used to select items across a range of difficulty. Items were ranked based on response probabilities, with higher scores recorded on easier items and lower scores reported for items of increasing difficulty. These estimated item response probabilities were applied to raw item scores to assign activation scores ranging from 0 to 100, with higher scores indicating greater activation. Activation scores can be converted to a four-level categorical score [16] defined as: Level 1—believing an active role is important; Level 2—having confidence and knowledge to take action; Level 3—taking action to maintain or improve health; and Level 4—continuing healthy behaviors under stress [17].

Original item selection, scoring and validation of PAM were based on a series of samples. Preliminary analysis was completed on 100 persons recruited through newspaper advertisements. Subsequent work was conducted on a mixed sample of 120 cardiac rehabilitation patients and 366 employees of a health organization. Final validation and scoring of both the PAM and PAM-13 was based on a US national probability sample of 1,515 persons age 45 and over selected via a random digit dialing system (48 % response rate with a minimum of 12 call backs) and screened to determine age eligibility. Respondents ranged in age from 45 to 97, with 66 % under the age of 65. Seventy-nine percent of the sample reported at least one chronic disease (angina/heart problems; arthritis, chronic pain, depression; diabetes; hypertension; lung disease, cancer and/or high cholesterol) [9, 16].

The strength of the tool and its potential clinical application has prompted adaptation, translation and validation in Germany [18, 19], Denmark [20] and the Netherlands [21]. Similarly, validation has been reported in the following client groups and settings: mental health [22], elective lumbar spine surgery [23], multi-morbid older adults [24] and in rural settings [25].

Recent conceptualizations of validity articulate it as a property of the score and how it is interpreted (rather than a property of the tool), meaning that validity must be established for each use and each specific group [26]. As noted above, the validation of the original PAM included few people with neurological conditions. Unlike conditions where lifestyle changes can alter the disease trajectory (asthma, diabetes, heart disease and even cancer), the prevention of symptoms associated with most neurological conditions is much less possible. Lifestyle changes do not prevent incontinence, gait problems, blurred vision or altered sensation. Additionally, some neurological conditions (MS and Parkinson's) are progressive, and their symptoms are unpredictable. For example, in Stepleman et al.'s study [6] of people with MS, those with relapsing–remitting MS (Mean = 42.29, SD = 5.11) had, on average, higher levels of patient activation than did patients with progressive MS (Mean = 38.44, SD = 6.05). Unpredictability is reported as one of the major challenges of life with MS [27]. To date, Stepleman et al. [6] are the only authors to report use with a neurological population, and their work was restricted to people with multiple sclerosis (MS). The purpose of this paper is to validate and assess the performance of the PAM-13 in a large sample of participants with diverse neurological conditions by examining data quality, internal consistency, unidimensionality, item scaling and construct validity.

Materials and methods

Data

Data were extracted from *The everyday experience of living with and managing a neurological condition* (*the LINC study*) [28], one of the studies in the National Population Health Study of Neurological Conditions (NPHSNC) [29]. The foundational LINC study was a cross-sectional survey of Canadians with neurological conditions between September 14, 2011 and July 1, 2012. The LINC study also included a longitudinal cohort study and a qualitative multiple perspective case study.

The NPHSNC stipulated an emphasis on people with neurological conditions whose charitable organizations held membership in The Neurological Health Charities of Canada (NHCC) [29]. Adults living in the community over

age 17 with one or more of the following neurological conditions were recruited: acquired brain injury, brain tumor, spinal cord injury, hydrocephalus cerebral palsy, epilepsy, spina bifida, MS, dystonia, Tourette syndrome, Parkinson's disease, Huntington's disease, muscular dystrophies, dementia and amyotrophic lateral sclerosis (ALS or Lou Gehrig's disease). People with other neurological conditions were not excluded but made up a smaller percentage of respondents. Recruitment occurred through the charitable organizations who publicized the study, the NHCC and other research mailing lists and registries (see [28] for details).

Ethical approval was received from the Health Canada Research Ethics Board, as well as the appropriate ethics review boards at Dalhousie University, Queen's University, the University of Manitoba, and the University of Prince Edward Island. The province of Newfoundland ethics review board also acknowledged the study within that province.

Participants

A total of 948 eligible adults age 17 and over responded to the on-line survey. For this analysis, participants who completed at least one item on the PAM-13 survey and provided answers to key demographic data were extracted. Participants whose PAM-13 data did not meet the required level of 50 % completed items (i.e., 50 % skipped items) [17] were excluded. Finally, because authors of the PAM-13 indicate that extreme scores should not be considered valid [17] participants with scores of 100 were also excluded. This left 722 eligible respondents, a sample comparable to many other PAM and PAM-13 validation studies and exceeding the sample size of participants with MS ($n = 199$) used by Stepleman et al. [6].

Measures

The PAM-13 [17] was used to measure patient activation. Demographic variables (age, sex, education, marital status and household income) used wording from the Canadian Community Health Survey (CCHS) [30]. The two demographic variables (education and household income) with consistent positive correlations with the PAM-13 were used to assess construct validity (known-groups method). Health status and lifestyle behaviors have also been shown to be correlated with activation [31]. Health status was measured using both a self-reported measure (SF-36) and a preference-based utility measure [Health Utility Index Mark-3 (HUI3)]. The SF-36 is a widely used, well-validated measure of general health status [32]. Reliability and validity have been established in neurological populations such as traumatic brain injury [33] and amyotrophic lateral sclerosis [34]. The two summary scales, the mental health scale (MCS) and physical health scale (PCS), were used to quantify health-related quality of life. The Health Utility Index (HUI) is a preference-based scoring index that can be converted to descriptive health classifications. Unlike a self-reported health status measure, the HUI incorporates community values or judgements regarding the impact of specific health states. The HUI incorporates domains of vision, hearing, speech, ambulation, dexterity, emotion, cognition and pain [35, 36]. Widely used and validated across multiple large population surveys, many studies have tested the responsiveness, reliability and validity (face, content, construct, convergent, discriminative and predictive) of the HUI across various populations and constructs [36]. The Simple Lifestyle Indicator Questionnaire (SLIQ) is a 12-item questionnaire designed to measure five elements of lifestyle behaviors (diet, activity, alcohol consumption, smoking and stress) [37]. Each component is scored individually or weighted to obtain an overall score from 0 to 10, with higher scores indicating a healthier lifestyle. In limited testing, the SLIQ has demonstrated good test–retest reliability ($r = 0.63$–0.97, $p = 0.001$), adequate internal consistency (Cronbach alpha = 0.58 for diet, 0.60 for activity) and adequate external validity ($r = 0.77$, $p = 0.001$) when participants' and blind raters' scores were compared [37].

Analysis and statistical methods

Data were reviewed for completeness prior to statistical analysis. All standardized measures were scored as per standard protocols, where missing data protocols were adhered to where specified (PAM-13, SF-36, HUI, SLIQ). Analysis was performed using Stata 13 (StataCorp 2013) and the statistical program R [38] with the extended Rasch model (eRm) package [39, 40]. Data quality for the PAM-13 items was assessed by examining the distribution of responses to each item, including the percentage of data missing or "not applicable."

Unidimensionality and internal consistency

To assess unidimensionality, item-rest correlations were examined, with the expectation that they should all exceed 0.50, and principle component analysis (PCA) was performed. Casewise deletion of missing data was used for the PCA, resulting in a smaller sample size ($n = 580$) for this analysis compared with others, but sensitivity analyses were run using imputation of missing data, and results were unaffected. We used the percent of variance explained by the first factor, the ratio of percent variance explained between the first and second factors, and a scree plot as evidence of unidimensionality [41]. Internal consistency

was assessed using Cronbach's alpha and item-to-rest correlations (correlation between each item and the average score of the rest of the scale). High item-to-rest correlations indicate strong internal consistency.

Item difficulty and response scale performance

Rasch modeling [40] was used to examine item scaling and performance. Specifically, we used the rating scale model to assess the scale performance of the items and whether the sequence of item difficulty matched that of the original PAM-13. Item difficulties were hypothesized to follow the same sequence as in PAM-13. Rasch modeling assumes unidimensionality (tested with the PCA noted above) and roughly equal frequencies across the response scale. Estimated location parameters were calculated, such that higher location parameters indicate participants who have a greater difficulty agreeing with the item. Location parameters estimate the sequence of items from easiest to most difficult, and separation distances between items of at least 0.15 logits are expected; smaller distances suggest that the two adjacent items do not individually add significantly to the model and that perhaps one can be eliminated. Infit and outfit statistics were used to assess how well each item contributed to defining the single underlying construct. Infit statistics are sensitive to items that are closest to the person's ability, while outfit statistics are more sensitive to items further away from the person's ability. A well-fitting Rasch model has infit and outfit statistics between 0.6 and 1.4.

Scatter plots were used to assess differences between the original PAM-13 and the rating scale model estimated from our data. Theta (logit) parameters from the rating scale model were linearly converted to a scale having the same minimum and maximum scores as the original PAM-13 scale in LINC. A constant was then subtracted so that the means were equal. PAM-13 activation Levels (Levels 1–4) were plotted on the graph to assess the degree of misclassification resulting from changes in item difficulties and scaling in our sample.

Relationship with other variables (construct validity)

Bivariate analysis was conducted to examine relationships between the PAM-13 and the demographic (education and income), health (HUI-3, SF-36 MCS and PCS) and lifestyle (SLIQ) validation variables. For continuous variables, Pearson's correlation was used; for dichotomous variables (two level categorical), a two-sample *t* test was used and for multilevel categorical variables one-way ANOVA models were estimated. Missing data were handled by casewise deletion.

Table 1 Description of participants

	Number	%
Age ($n = 724$)		
≤25	39	5.39
26–35	97	13.40
36–45	102	14.09
46–55	189	26.10
56–65	172	23.76
66–75	94	12.98
76–85	31	4.14
>85	1	0.14
Marital status ($n = 725$)		
Single, never married	160	22.07
Married, common-law	445	61.38
Widowed, separated, divorced	120	16.55
Education ($n = 697$)		
Less than post-secondary graduation	194	27.83
Post-secondary graduation	503	72.17
Sex ($n = 725$)		
Male	257	35.45
Female	468	64.55
Household income ($n = 507$)		
<$20,000	85	16.77
$20–59,999	198	39.05
$60–89,999	106	20.91
>90,000	118	23.27
Employment at any time in past 3 months ($n = 715$)		
Working	234	32.73
Not working due to health	272	38.04
Not working for other reasons	209	29.23

Results

Participants in this study were mostly working age adults between the ages of 26 and 65 (77 %); 6 % were younger than 26 and only one participant was over 85 (Table 1). They were predominantly woman (65 %) in married or common-law relationships (61 %). Participants were highly educated with 72 % reporting completion of post-secondary education and 23 % reporting a household income of >$90,000. Although predominantly of working age, only 33 % were employed, 38 % were not working due to health reasons, and the remaining 29 % were not working for other reasons.

Data quality

The percentage of missing data was below 1 % for all items except one (item 11) at 1.3 %. The percentage of responses in the "not applicable" category was higher, ranging from

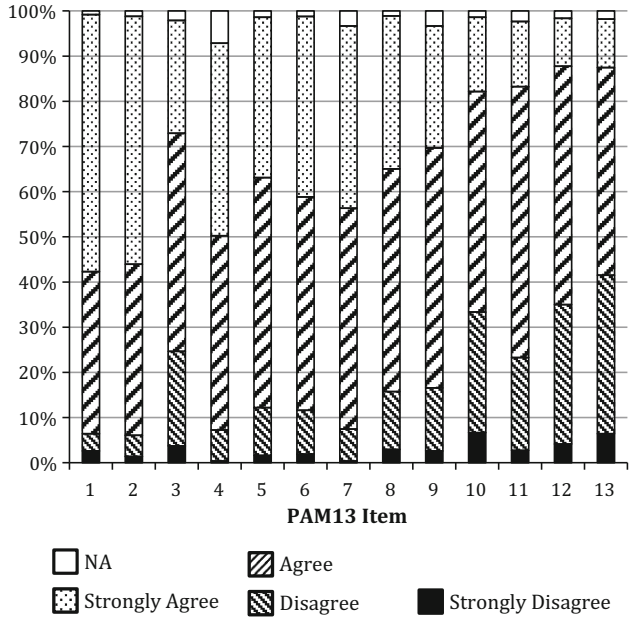

Fig. 1 Distribution of responses by PAM-13-SF item (excluding missing data)

0.3 to 6.5 % (see Fig. 1). Item #4 (*I know what each of my prescribed medications does.*) had the highest number of not applicable responses. Items 7 and 9 were next, both with a not applicable response rate of 2.5 %. Responses across the other four categories (*strongly disagree, disagree, agree, strongly agree*) were not evenly distributed (Fig. 1). The *strongly disagree* category was seldom used with <10 % of responses in this category. By contrast, the *agree* and *strongly agree* categories were highly selected, together representing between approximately 60 and 85 % of responses, depending on the item. Because of the low frequency of use of the *strongly disagree* category in this sample, the *strongly disagree* and *disagree* categories were combined to create a three-category dataset used for the estimation of the rating scale model.

Unidimensionality and internal consistency

Item-rest correlations ranged from 0.427 to 0.655 with only two items (item 1, 2) having item-rest correlations slightly below the cutoff for strong correlations (≥0.50). The PCA showed that 40 % of the variance was explained by the first component, while the second and third components accounted for 10 and 9 % of the variance. We also ran a factor analysis (principal axis factoring), and the first factor accounted for 61 % of the variance, while the second and third factors accounted for 12 and 10 % of the variance. Scree plots indicated that a one-factor solution was plausible, but suggest that additional factors could be considered. Cronbach's alpha was 0.87, exceeding the minimum criteria of 0.8 and indicating strong internal consistency. Independent of the number of items and sample size, the average inter-item correlations ranged from 0.327 to 0.351, well within the ideal range of 0.15–0.50.

Item difficulty and response scale performance

As per the design and scaling of PAM-13, higher order items were generally more difficult; respondents were less likely to *strongly agree* or *agree* with higher order items (see shift in distributions in Fig. 1 and item means in Table 2). However, these descriptive results also show substantial deviations from the PAM-13 sequence of item difficulty. For example, mean item scores and the distributions of responses indicate that, for this sample, item 3 was more difficult, and item 7 less difficult than specified in the PAM-13 scaling (Fig. 1).

Evidence of deviations in item difficulties between the PAM-13 and the LINC data are clear in the rating scale model. Table 2 shows that difficulty parameters for PAM-13 items followed a clear sequence of increasing difficulty and infit (range from 0.709 to 1.095) and outfit (range 0.718–1.262) statistics fell within the acceptable limits of 0.6 and 1.4, indicating a good fit of the model. However, the ranking of items by difficulty differs from the original PAM-13 ranking. To more clearly illustrate the differences in item sequence between the PAM-13 and the calculated model, difficulty location parameters and item thresholds were graphically displayed using the PAM-13 item sequence and the sequence determined by the rating scale model (Fig. 2a, b). Items 3 and 7 are particularly noticeable "misfit" items. The item *I am confident that I can take actions that will help prevent or minimize some symptoms or problems associated with my health* is ranked as item #3 in difficulty in the PAM-13. For participants in this study, it was the ninth most difficult item. In contrast, item #7 in the PAM-13 (*I am confident that I can follow through on medical treatments I need to do at home*) was ranked as fourth easiest to agree to by participants in the LINC study.

To further examine the deviations in item difficulties and altered estimates of respondent activation level, estimates of patient activation from the PAM-13 were plotted against comparatively scaled estimates of patient activation from our rating scale model (Fig. 3). The PAM-13 manual recommends interventions based on the four levels (Level 1—believing an active role is important; Level 2—having confidence and knowledge to take action; Level 3—taking action to maintain or improve health; and Level 4—continuing healthy behaviors under stress) (Insignia Health 17). Vertical lines demarcating cutoffs for the four activation levels used in PAM-13 and horizontal lines indicating those same activation levels for our model were added. While the correlation between the two estimates of activation is high ($r = 0.98$), it

Table 2 Results of the Rasch rating scale model of PAM-13 in a neurological population with data collapsed to three response categories per item

Item number based on partial credit model	Original PAM-13 item number	Number of participants (*n*) responding to item	Item mean	SD	Not applicable (%)	Difficulty parameter[a]	SE of difficulty parameter	Outfit MSQ	Infit MSQ
1	1	717	3.48	0.698	0.3	Ref		1.262	1.095
2	2	717	3.48	0.655	0.7	−1.308	0.075	1.018	0.999
3	4	671	3.38	0.639	6.5	−0.892	0.073	0.900	0.921
4	7	701	3.34	0.627	2.5	−0.732	0.071	0.899	0.813
5	6	712	3.27	0.714	0.6	−0.532	0.069	1.059	0.972
6	5	711	3.23	0.698	0.8	−0.325	0.068	0.848	0.851
7	8	714	3.16	0.753	0.4	−0.162	0.067	0.984	0.983
8	9	698	3.08	0.726	2.5	0.112	0.067	1.037	0.973
9	3	708	2.96	0.789	1.4	0.474	0.067	1.088	1.080
10	11	703	2.88	0.677	1.4	0.790	0.068	0.718	0.709
11	10	711	2.76	0.806	1.0	1.059	0.069	1.026	1.041
12	12	709	2.71	0.712	1.0	1.317	0.07	0.735	0.747
13	13	712	2.62	0.763	1.4	1.569	0.072	1.019	1.030
					Resp.[b]	2.903	0.063		

[a] Difficulty relative to previous item

[b] Difficulty between response levels in items

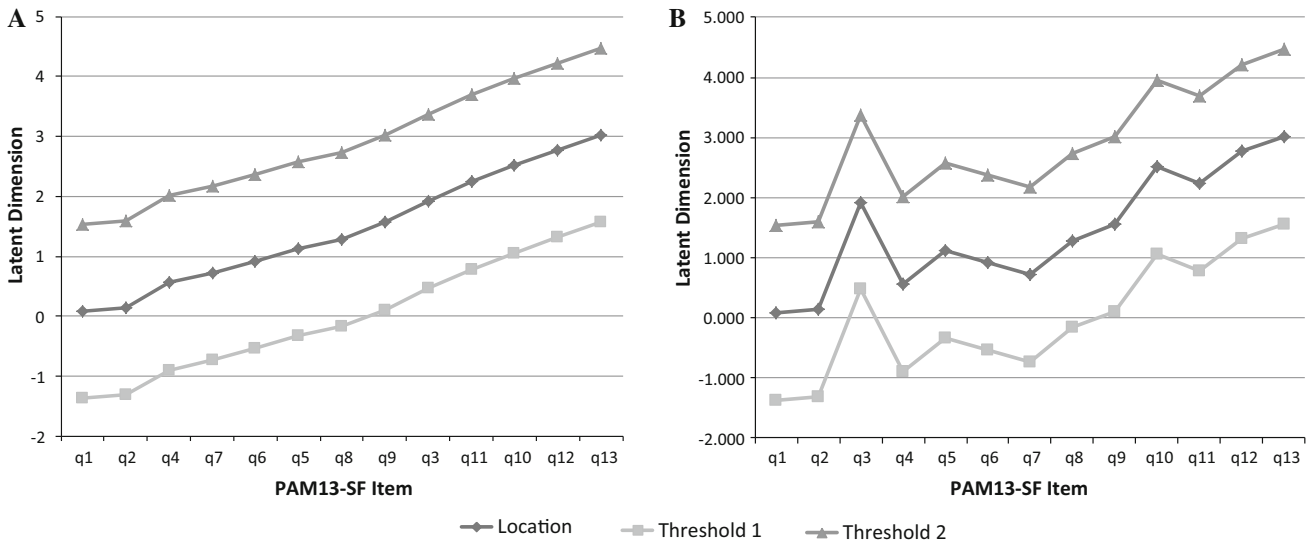

Fig. 2 Location and threshold parameters from the rating scale model. **a** Item sequence in rating scale model. **b** Item sequence as per original version of PAM-13

can be seen that differences in item difficulties from PAM-13 have the biggest impact at low levels of activation. Agreement is lower for PAM-13 scores <50, where estimates of activation levels from our rating scale model are generally higher than the PAM-13. The figure demonstrates that differences in activation may be particularly significant when activation is categorized as levels as many subjects scored as Level 1 by PAM-13 would be classified as Level 2 by our rating scale model.

Relationship with other variables (construct validity)

As expected, differences between groups based on education and household income were found (Table 3). Participants who had completed post-secondary education had higher PAM-13 scores than those with lower levels of education. Significant differences were also found based on income. People with household incomes of >$90,000 had the highest PAM-13 scores. Also as expected, health

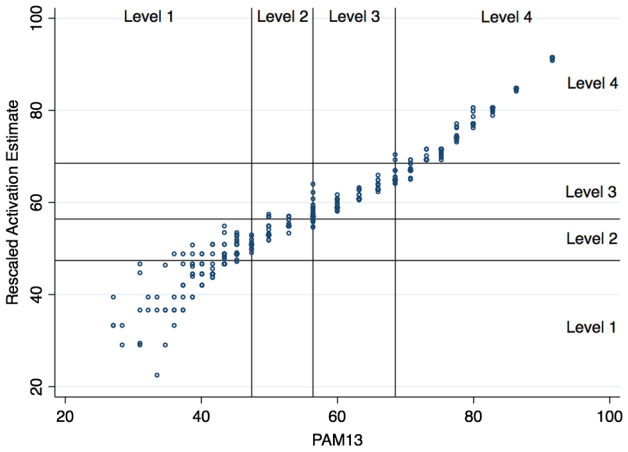

Fig. 3 Comparison of respondent activation level using the PAM-13 conversion tables and activation estimates from our Rasch rating scale model

Table 3 Construct validity—comparison with validation variables

Variables	Pearson *r*	*p* value
Health status[a]		
HUI-3	0.32	<0.0001
MCS	0.35	<0.0001
PCS	0.20	<0.0001
SLIQ	0.29	<0.0001
Variables	**Mean PAM (SD)**	***p* value**
Education ($n = 692$)[b]		0.024
Less than post-secondary graduation	59.1 (15.74)	
Post-secondary graduation	61.8 (14.01)	
Family income ($n = 506$)[c]		0.004
<$20,000	57.3 (15.75)	
$20–59,999	60.2 (14.63)	
$60–89,999	61.3 (13.33)	
>90,000	64.5 (13.55)	

[a] Pearson correlation

[b] Two-sample *t* test

[c] Analysis of variance (ANOVA)

variables were strongly correlated ($p \leq 0.0001$) with the PAM-13 (Table 3) using the self-reported measures of quality of life (SF-36 MCS and PCS), healthy lifestyles (SLIQ) and the preference-based HUI-3.

Discussion

Overall, the PAM-13 performed reasonably well in our sample of people with neurological conditions, but some scaling concerns were identified. Our analysis found evidence of good internal consistency. Principal component analysis and factor analysis suggest that assumed unidimensionality of PAM-13 is reasonable, but also suggest that consideration of additional dimensions in an activation measure should be considered in future research.

Construct validity, through comparison to three other measures and two demographic variables was also very strong ($p = 0.024$ to <0.0001) with strong evidence of associations with patient activation. These findings are consistent with studies from the USA where the PAM was developed [9, 16] and in studies validating the tool internationally [42] in which authors have similarly concluded that it is reliable, valid tool that retains the probabilistic, Guttman-like scale properties. These conclusions are mirrored by studies using clinical samples such as MS [6], mental health conditions [22] people awaiting spine surgery [23] and multi-morbid older adults living in the community [24].

While a Rasch model had a good overall fit to our data, concerns about PAM-13 scaling arise from our results. The scaled difficulty of items in our sample of people with neurological conditions deviated significantly from the scaling of PAM-13. Notably, participants in our study found it more difficult to agree with the item #3 *I am confident that I can help prevent or reduce problems associated with my health*. This may reflect the progressive and unpredictable nature of symptoms as experienced by many people with neurological conditions. Participants found it less difficult to agree with item #7, *I am confident that I can follow through on medical treatments I need to do at home*. This item was ranked as fourth easiest (rather than seventh) by participants in our study. This result is not surprising as medication adherence rates are high in people with neurological conditions. While WHO reports an adherence rate of 50 % [43] for people with chronic illnesses, adherence rates for disease modifying therapies by people with MS is 61–87 % [44] and a multicentre European study on people with Parkinson's disease showed 97 % total median adherence to anti-Parkinson medication [45].

Interestingly, similar findings showing variation in item difficulties between samples clearly appear in the results, but not the conclusions of other studies. This suggests that this problem is not specific to our sample, and that scaling variation may be a more generalized problem with the PAM-13. Reported location parameters indicate that item #3 was rated as more difficult (i.e., the item sequence ranking was higher than in PAM-13) by people with MS [6], people with chronic conditions in the Netherlands and Germany [19, 21], and in rural areas [25]. Item #7 was ranked as the first, second and fourth easiest item in the MS [6], Dutch [21], and Danish [42] studies, respectively, but was perceived as much more difficult by participants in the German study who rated it as tenth out of 13th [18].

Viewed together, these results suggest that consideration of differences in the self-management tasks between diseases and populations may be an important consideration in the measurement of activation.

One possible explanation for differences in response scaling may be the disease composition of samples. The PAM and PAM-13 were rigorously designed using a series of studies, building from smaller, convenience samples during design and item selection, to a large, national sample to determine the final scaling for items and cutoffs for levels. The low prevalence of neurological conditions in these samples [9, 16] potentially resulted in a lack of items relevant to this population and estimates of item difficulties that do not reflect the experiences of persons with neurological conditions. The benefits of a national sample may have inadvertently resulted in items of limited relevance to groups with a lower prevalence. Differing "response processes" may provide a second explanation. Response process refers to the fit between the construct being measured and respondents' thought processes as he or she answers the questions i.e., how the person understands the meaning of the questions [26]. Response processes may be influenced by cultural beliefs (for example, in European vs. Canadian samples), access to and expectations of their healthcare system (for example, in Canadian vs. American samples) alone or in combination with the nature, prognosis and trajectory of the condition (for example, in samples with different diagnosis). Future studies could investigate response process using "think aloud" or cognitive interviewing [46].

The higher rate of *not applicable* responses for item #4 *I know what each of my prescribed medications do* also warrants comment. Greater than 6 % of participants indicated that this statement did not apply to them, inferring that they do not take medications. Unlike diabetes, hypertension and many other common chronic conditions, some neurological conditions (spinal cord injury, cerebral palsy etc.) do not warrant pharmacological management or symptom control. At least one other study has shown that a considerable number of people with MS have found medication-related questions "not applicable" [47]. Interestingly, the percentage of responses in the "not applicable" category was lower in our sample than in other studies; higher than 10 % for items 2, 7 and 10 in the Dutch study [21] and items 4, 9, 11, 12 and 13 in Danish version [42]. Rewording this question to include understanding of other prescribed treatments (e.g., the prevention of secondary symptoms such of urinary tract infections, falls and/or bed sores) may increase the applicability to additional clinical populations.

Our results sound a caution in the use of PAM-13 as a clinical assessment tool for persons with neurological conditions. The PAM-13 manual itself promotes it as a useful clinical tool providing "Guidelines for Using the PAM-13 to Tailor Care" based on the four levels of patient activation [17]. However, in our sample differences in item difficulties resulted in individuals' activation levels being underestimated at low levels of activation. We estimate that 64 of 156 (41 %) participants would be incorrectly rated as being as Level 1 (may not yet believe the patient role is important) rather than Level 2 (lacks confidence and knowledge to take action). For a person in activation Level 1, Insignia Health 17 recommends that action planning should focus on monitoring choices and outcomes and encouraging simple strategies to develop confidence such as preparing three questions to ask their healthcare professional. However, a person in activation Level 2 is assumed to have some confidence and healthcare professionals are expected to focus on helping people understand their health condition or related regimen, supporting small behavior changes and developing problem-solving skills [17]. While potentially counter to the recommendation by Insignia Health, one solution to overcome the apparent loss of precision at the lower end of activation scale would be to design interventions tailored to the combined needs of clients at both levels 1 and 2.

Our results lend credence to cautions by others regarding the use of PAM-13 as a clinical assessment tool. While some authors are suggesting that PAM-13 is sufficiently precise to justify its use with individual clients [31, 48], others have concluded that the PAM-13 has potential as a clinical tool, but requires further testing of the scale to identify which items are most applicable in different clinical settings [49]. Ledford et al. [50] were less optimistic, reporting that some people (with diabetes) with high PAM scores still responded passively in a patient–provider communication.

The LINC survey is not a probability sample of persons with neurological conditions; however, representativeness is less critical in validation studies than in other designs. Subjects appear more likely to be female, younger and better educated than the general population and some diseases are probably overrepresented because of the recruitment process. The sample size was not sufficiently large for us to assess whether item difficulties differed by neurological condition. This should be considered in the interpretation of our results and explored in future research.

In conclusion, our results suggest that PAM-13 provides a suitably reliable and valid instrument for research in patients with neurological conditions. However, our study suggests potential for measurement error and bias at low levels of activation. The implications of this error and bias may be of particular importance in clinical applications of the tool. Our study also suggests that measurement of activation may benefit from items and scaling tailored to specific diagnostic groups, recognizing their unique attributes and management challenges.

Acknowledgments This study was part of the National Population Health Study of Neurological Conditions. We wish to acknowledge the membership of Neurological Health Charities Canada and the Public Health Agency of Canada for their contribution to the success of this initiative. Funding for the study was provided by the Public Health Agency of Canada. The opinions expressed in this publication are those of the authors/researchers and do not necessarily reflect the official views of the Public Health Agency of Canada. The authors wish to thank the many participants in the LINC study as well as the large research team who made the study possible.

References

1. Hibbard, J. H., & Greene, J. (2013). What the evidence shows about patient activation: Better health outcomes and care experiences; fewer data on costs. *Health Affairs (Millwood), 32*(2), 207–214.
2. Wagner, E. H. (1998). Chronic disease management: What will it take to improve care for chronic illness? *Effective clinical practice, 1*(1), 2–4.
3. Wagner, E. E. H. (2001). Meeting the needs of chronically ill people. *British Medical Journal, 323*(7319), 945–946.
4. Barr, V. J., Robinson, S., Marin-Link, B., Underhill, L., Dotts, A., Ravensdale, D., et al. (2003). The expanded chronic care model: An integration of concepts and strategies from population health promotion and the chronic care model. *Hospital Quarterly, 7*(1), 73–82.
5. Greene, J., & Hibbard, J. H. (2012). Why does patient activation matter? An examination of the relationships between patient activation and health-related outcomes. *Journal of General Internal Medicine, 27*(5), 520–526.
6. Stepleman, L., Rutter, M., Hibbard, J., Johns, L., Wright, D., & Hughes, M. (2010). Validation of the patient activation measure in a multiple sclerosis clinic sample and implications for care. *Disability and Rehabilitation, 32*(19), 1558–1567.
7. Fowles, J. B., Terry, P., Xi, M., Hibbard, J., Bloom, C. T., & Harvey, L. (2009). Measuring self-management of patients' and employees' health: Further validation of the patient activation measure (PAM) based on its relation to employee characteristics. *Patient Education and Counseling, 77*(1), 116–122.
8. Hibbard, J. H., Collins, P. A., Mahoney, E., & Baker, L. H. (2010). The development and testing of a measure assessing clinician beliefs about patient self-management. *Health Expectations, 13*(1), 65–72.
9. Hibbard, J. H., Mahoney, E. R., Stockard, J., & Tusler, M. (2005). Development and testing of a short form of the patient activation measure. *Health Services Research, 40*(6 Pt 1), 1918–1930.
10. Hibbard, J. H., Mahoney, E. R., Stock, R., & Tusler, M. (2007). Do increases in patient activation result in improved self-management behaviors? *Health Services Research, 42*(4), 1443–1463.
11. Kaplan, R. M., Anderson, J. P., Wu, A. W., Mathews, W. C., Kozin, F., & Orenstein, D. (1989). The quality of well-being scale. Applications in AIDS, cystic fibrosis, and arthritis. *Medical Care, 27*(3 Suppl.), 27–43.
12. Alexander, J. A., Hearld, L. R., Mittler, J. N., & Harvey, J. (2012). Patient–physician role relationships and patient activation among individuals with chronic illness. *Health Services Research, 47*(3 Pt 1), 1201–1223.
13. Glasgow, R. E., Toobert, D. J., Hampson, S. E., & Strycker, L. A. (2002). Implementation, generalization and long-term results of the "choosing well" diabetes self-management intervention. *Patient Education Counseling, 48*(2), 115–122.
14. Begum, N., Donald, M., Ozolins, I. Z., & Dower, J. (2011). Hospital admissions, emergency department utilisation and patient activation for self-management among people with diabetes. *Diabetes Research and Clinical Practice, 93*(2), 260–267.
15. Hibbard, J. H., & Cunningham, P. J. (2008). How engaged are consumers in their health and health care, and why does it matter? *Res Briefs, 8*, 1–9.
16. Hibbard, J. H., Stockard, J., Mahoney, E. R., & Tusler, M. (2004). Development of the patient activation measure (PAM): Conceptualizing and measuring activation in patients and consumers. *Health Services Research, 39*(4 Pt 1), 1005–1026.
17. Insignia Health. (2011). *Patient activation measure* (*PAM*) 13™ *License Materials copyright Insignia Health, LLC*.
18. Brenk-Franz, K., Hibbard, J. H., Herrmann, W. J., Freund, T., Szecsenyi, J., Djalali, S., et al. (2013). Validation of the German version of the patient activation measure 13 (PAM13-D) in an international multicentre study of primary care patients. *PLoS One, 8*(9), e74786.
19. Zill, J. M., Dwinger, S., Kriston, L., Rohenkohl, A., Harter, M., & Dirmaier, J. (2013). Psychometric evaluation of the German version of the patient activation measure (PAM13). *BMC Public Health, 13*, 1027-2458-13-1027.
20. Maindal, H. T., Sandbaek, A., Kirkevold, M., & Lauritzen, T. (2011). Effect on motivation, perceived competence, and activation after participation in the "Ready to Act" programme for people with screen-detected dysglycaemia: A 1-year randomised controlled trial, Addition-DK. *Scandinavian Journal of Public Health, 39*(3), 262–271.
21. Rademakers, J., Nijman, J., van der Hoek, L., Heijmans, M., & Rijken, M. (2012). Measuring patient activation in The Netherlands: Translation and validation of the American short form patient activation measure (PAM13). *BMC Public Health, 12*, 577.
22. Green, C. A., Perrin, N. A., Polen, M. R., Leo, M. C., Hibbard, J. H., & Tusler, M. (2010). Development of the patient activation measure for mental health. *Administration and Policy in Mental Health, 37*(4), 327–333.
23. Skolasky, R. L., MacKenzie, E. J., Riley, L. H, I. I. I., & Wegener, S. T. (2009). Psychometric properties of the patient activation measure among individuals presenting for elective lumbar spine surgery. *Quality of Life Research, 18*(10), 1357–1366.
24. Skolasky, R. L., Green, A. F., Scharfstein, D., Boult, C., Reider, L., & Wegener, S. T. (2011). Psychometric properties of the patient activation measure among multimorbid older adults. *Health Services Research, 46*(2), 457–478.
25. Hung, M., Carter, M., Hayden, C., Dzierzon, R., Morales, J., Snow, L., et al. (2013). Psychometric assessment of the patient activation measure short form (PAM-13) in rural settings. *Quality of Life Research, 22*(3), 521–529.
26. Cook, D. A., & Beckman, T. J. (2006). Current concepts in validity and reliability for psychometric instruments: Theory and application. *The American Journal of Medicine, 119*(2), 166.e7-166.16.
27. Malcomson, K. S., Lowe-strong, A. S., & Dunwoody, L. (2008). What can we learn from the personal insights of individuals living and coping with multiple sclerosis? *Disability and Rehabilitation, 30*(9), 662–674.
28. Versnel, J., Packer, T., Weeks, L. E., Brown, J., Godwin, M., Hutchinson, S., et al. (2013). The everyday experience of living with and managing a neurological condition (the LINC study): Study design. *BMC Neurology, 13*, 30-2377-13-30.
29. Caesar-Chavannes, C. R., & MacDonald, S. (2013). Cross-Canada forum—National Population Health Study of neurological conditions in Canada. *Chronic Diseases Injuries in Canada, 33*(3), 188–191.
30. Statistics Canada. (2009). *User guide 2007–2008 microdata files. Canadian Community Health Survey (CCHS)—Annual Component*. Ottawa: Statistics Canada.

31. Harvey, L., Fowles, J. B., Xi, M., & Terry, P. (2012). When activation changes, what else changes? The relationship between change in patient activation measure (PAM) and employees' health status and health behaviors. *Patient Education and Counseling, 88*(2), 338–343.
32. Hawthorne, G., Osborne, R. H., Taylor, A., & Sansoni, J. (2007). The SF36 version 2: Critical analyses of population weights, scoring algorithms and population norms. *Quality of Life Research, 16*(4), 661–673.
33. Guilfoyle, M. R., Seeley, H. M., Corteen, E., Harkin, C., Richards, H., Menon, D. K., et al. (2010). Assessing quality of life after traumatic brain injury: Examination of the short form 36 health survey. *Journal of Neurotrauma, 27*(12), 2173–2181.
34. Jenkinson, C., Hobart, J., Chandola, T., Fitzpatrick, R., Peto, V., Swash, M., et al. (2002). Use of the short form health survey (SF-36) in patients with amyotrophic lateral sclerosis: Tests of data quality, score reliability, response rate and scaling assumptions. *Journal of Neurology, 249*(2), 178–183.
35. Marra, C. A., Woolcott, J. C., Kopec, J. A., Shojania, K., Offer, R., Brazier, J. E., et al. (2005). A comparison of generic, indirect utility measures (the HUI2, HUI3, SF-6D, and the EQ-5D) and disease-specific instruments (the RAQoL and the HAQ) in rheumatoid arthritis. *Social Science and Medicine, 60*(7), 1571–1582.
36. Feng, Y., Bernier, J., McIntosh, C., & Orpana, H. (2009). Validation of disability categories derived from Health Utilities Index Mark 3 scores. *Health Reports, 20*(2), 43–50.
37. Godwin, M., Streight, S., Dyachuk, E., van den Hooven, E. C., Ploemacher, J., Seguin, R., et al. (2008). Testing the Simple Lifestyle Indicator Questionnaire: Initial psychometric study. *Canadian Family Physician, 54*(1), 76–77.
38. R Development Core Team. (2006). *R: A language and environment for statistical computing*. Vienna: R Development Core Team.
39. Mair, P., & Hatzinger, R. (2006). *eRm; extended Rasch models. R package version 0.3.2.* http://CRAN.R-project.org/. (Accessed October 2013).
40. Mair, P., & Hatzinger, R. (2007). CML based estimation of extended Rasch models with the eRm package in R. *Psychology Science, 49*(1), 26–43.
41. Hattie, J. (1985). Methodology review: Assessing unidimensionality of tests and items. *Applied Psychological Measurement, 9*(2), 139–164.
42. Maindal, H. T., Sokolowski, I., & Vedsted, P. (2009). Translation, adaptation and validation of the American short form patient activation measure (PAM13) in a Danish version. *BMC Public Health, 9*, 209.
43. World Health Organization. (2003). *Adherence to long-term therapies: Evidence for action* (p. 12). Geneva: World Health Organization.
44. Remington, G., Rodriguez, Y., Logan, D., Williamson, C., & Treadaway, K. (2013). Facilitating medication adherence in patients with multiple sclerosis. *International Journal of MS Care, 15*(1), 36–45.
45. Grosset, D., Antonini, A., Canesi, M., Pezzoli, G., Lees, A., Shaw, K., et al. (2009). Adherence to antiparkinson medication in a multicenter European study. *Movement Disorders, 24*(6), 826–832.
46. Drennan, J. (2003). Cognitive interviewing: Verbal data in the design and pretesting of questionnaires. *Journal of Advanced Nursing, 42*(1), 57–63.
47. Ghahari, S., Forwell, S., & Khoshbin, L. (2014). The multiple sclerosis self-management scale: Clinicometric testing. *International Journal of MS Care,*. doi:10.7224/1537-2073.2013-019.
48. Harvey, L., Fowles, J. B., Xi, M., & Terry, P. (2012). When activation changes, what else changes? The relationship between change in patient activation measure (PAM) and employees' health status and health behaviors. *Patient Education and Counseling, 88*(2), 338–343.
49. Mosen, D. M., Schmittdiel, J., Hibbard, J., Sobel, D., Remmers, C., & Bellows, J. (2007). Is patient activation associated with outcomes of care for adults with chronic conditions? *Journal of Ambulatory Care Management, 30*(1), 21–29.
50. Ledford, C. J. W., Ledford, C. C., & Childress, M. A. (2013). Exploring patient activation in the clinic: Measurement from three perspectives. *Health Education and Behavior, 40*(3), 339–345.